ΠΟΛΥΤΕΧΝΕΙΟ ΚΡΗΤΗΣ

ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ ΠΑΡΑΓΩΓΗΣ ΚΑΙ ΔΙΟΙΚΗΣΗΣ

ΔΥΝΑΜΙΚΟΣ ΠΡΟΓΡΑΜΜΑΤΙΣΜΟΣ

Καθ. Μάρκος Παπαγεωργίου

ΧΑΝΙΑ, 2020

ΠΙΝΑΚΑΣ ΠΕΡΙΕΧΟΜΕΝΩΝ

1. ΕΙΣΑΓΩΓΗ 1.1

2. ΣΥΝΔΥΑΣΤΙΚΑ ΠΡΟΒΛΗΜΑΤΑ ΚΑΙ ΠΑΙΓΝΙΑ 2.1

2.1 Προβλήματα Βραχύτατων Μονοπατιών 2.1

2.2 Παίγνια Δύο Προσώπων 2.4

2.3 Άλλα Συνδυαστικά Προβλήματα 2.12

Προβλήματα 2.16

3. ΒΕΛΤΙΣΤΟΣ ΕΛΕΓΧΟΣ ΔΙΑΚΡΙΤΟΥ ΧΡΟΝΟΥ 3.1

3.1 Διατύπωση του Προβλήματος 3.1

3.2 Ο Αναδρομικός Τύπος του Bellman 3.2

3.3 Διακριτός Δυναμικός Προγραμματισμός 3.9

(α) Διακριτοποίηση 3.9

(β) Αλγόριθμος Διακριτού Δυναμικού Προγραμματισμού 3.11

(γ) Υπολογιστικός Φόρτος 3.13

(δ) Γραμμική Παρεμβολή 3.15

(ε) Τελικά Συμπεράσματα 3.18

3.4. Επεκτάσεις του Προβλήματος 3.20

(α) Διαχρονικοί Περιορισμοί 3.22

(β) Προβλήματα Ελεύθερου Τελικού Χρόνου 3.24

(γ) Μη Μεταβλητός Κανόνας Ελέγχου 3.25

Προβλήματα 3.30

4. ΒΕΛΤΙΣΤΟΣ ΕΛΕΓΧΟΣ ΣΥΝΕΧΟΥΣ ΧΡΟΝΟΥ 4.1

4.1 Διατύπωση του Προβλήματος 4.1

4.2 Η Εξίσωση Hamilton-Jacobi-Bellman 4.1

Προβλήματα 4.6

5. ΕΦΑΡΜΟΓΕΣ 5.1

5.1 Το Πρόβλημα Αποθήκευσης 5.1

5.2 Το Πρόβλημα Αντικατάστασης 5.3

5.3 Στατικά Προβλήματα Κατανομής 5.6

α) Το Πρόβλημα Αξιοπιστίας 5.8
β) Το Πρόβλημα Κατανομής Εργασίας 5.8
γ) Το Πρόβλημα Κατανομής Φορτίου 5.9
δ) Γενική Επίλυση 5.9
5.4 Γραμμικός-Τετραγωνικός (ΓΤ) Έλεγχος 5.11
α) ΓΤ Έλεγχος Συνεχούς Χρόνου 5.11
β) ΓΤ Έλεγχος Διακριτού Χρόνου 5.15
Προβλήματα 5.19

6. ΠΑΡΑΛΛΑΓΜΕΝΟΙ ΑΛΓΟΡΙΘΜΟΙ ΔΥΝΑΜΙΚΟΥ ΠΡΟΓΡΑΜΜΑΤΙΣΜΟΥ 6.1
6.1 Εισαγωγή 6.1
6.2 Διακριτός Διαφορικός Δυναμικός Προγραμματισμός (ΔΔΔΠ) 6.1
6.3 Αυξάνων Δυναμικός Προγραμματισμός (ΑΔΠ) 6.3
6.4 Διαφορικός Δυναμικός Προγραμματισμός (ΔΙΑΔΠ) 6.7
6.5 Συγκριτικό Παράδειγμα 6.10
Προβλήματα 6.11

7. ΣΤΟΧΑΣΤΙΚΟΣ ΔΥΝΑΜΙΚΟΣ ΠΡΟΓΡΑΜΜΑΤΙΣΜΟΣ 7.1
7.1 Εισαγωγή 7.1
7.2 Το Πρόβλημα Στοχαστικού Βέλτιστου Ελέγχου (ΠΣΒΕ) 7.3
7.3 Μέθοδος Στοχαστικού Δυναμικού Προγραμματισμού 7.6
7.4 Διακριτός Στοχαστικός Δυναμικός Προγραμματισμός 7.13
7.5 Επεκτάσεις του Προβλήματος 7.17
α) Προβλήματα με Νεκρούς Χρόνους 7.17
β) Συσχετιζόμενες, Μετρήσιμες Διαταραχές 7.17
γ) Χρήση Προγνώσεων 7.17
δ) Μη Μεταβλητός Κανόνας Ελέγχου 7.18
ε) Προβλήματα με Ελλιπή Πληροφορία 7.23
Προβλήματα 7.23

8. ΕΦΑΡΜΟΓΕΣ ΣΤΟΧΑΣΤΙΚΟΥ ΔΥΝΑΜΙΚΟΥ ΠΡΟΓΡΑΜΜΑΤΙΣΜΟΥ 8.1
8.1 Η Έννοια του Ρίσκου 8.1
8.2 Το Πρόβλημα του Εισοδηματία 8.4
8.3 Το Πρόβλημα Προγραμματισμού Επενδύσεων 8.5
8.4 Προβλήματα Βέλτιστης Στάσης 8.12

8.5 Στοχαστικός Γραμμικός-Τετραγωνικός Έλεγχος 8.14
Προβλήματα 8.16

9. ΕΠΙΛΟΓΟΣ Η ΕΠΙΛΟΓΕΣ ΠΡΑΚΤΙΚΗΣ ΕΦΑΡΜΟΓΗΣ 9.1

ΠΡΟΤΕΙΝΟΜΕΝΗ ΒΙΒΛΙΟΓΡΑΦΙΑ .....

ΠΑΡΑΡΤΗΜΑ Α: ΔΥΝΑΜΙΚΑ ΣΥΣΤΗΜΑΤΑ Α.1
Α.1 Γενικοί Ορισμοί Α.1
Α.2 Δυναμικά Συστήματα Συνεχούς Χρόνου Α.1
Α.3 Δυναμικά Συστήματα Διακριτού Χρόνου Α.2
Α.4 Ελεγξιμότητα και Παρατηρισιμότητα Γραμμικών Συστημάτων Α.3
Προβλήματα Α.4

ΠΑΡΑΡΤΗΜΑ Β: ΒΕΛΤΙΣΤΟΣ ΕΛΕΓΧΟΣ Β.1
Β.1 Διατύπωση του Προβλήματος Βέλτιστου Ελέγχου Β.1
Β.2 Προσεγγίσεις στο Πρόβλημα Βέλτιστου Ελέγχου Β.7
Β.3 Απόλυτο και Σχετικό Ελάχιστο Β.9

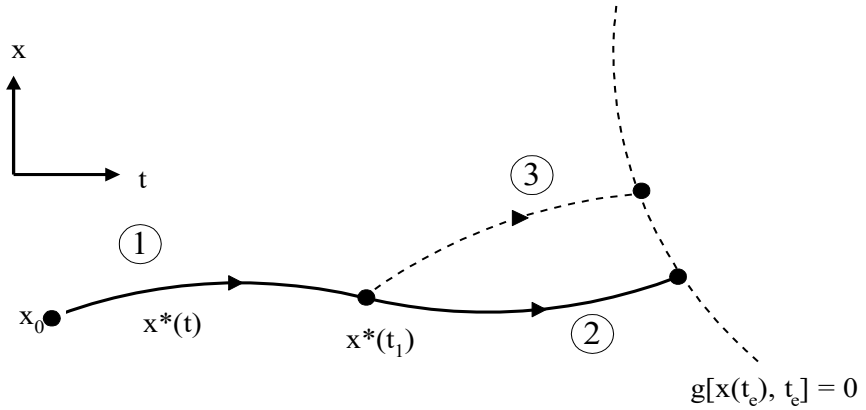
1. ΕΙΣΑΓΩΓΗ

Η μέθοδος του **Δυναμικού Προγραμματισμού** αναπτύχθηκε στις αρχές της δεκαετίας 1950 από τον εξέχοντα αμερικανό μαθηματικό **R.E. Bellman**. Ο Δυναμικός Προγραμματισμός βασίζεται στην λεγόμενη **Αρχή του Βέλτιστου** που δεν είναι παρά μια απλούστατη και εύκολα κατανοητή ιδιότητα της λύσης προβλημάτων βέλτιστου ελέγχου. Οι συνέπειες όμως της ιδιότητας αυτής είναι σημαντικές και οδήγησαν σε μια οικογένεια αλγορίθμων για την επίλυση προβλημάτων δυναμικής, αλλά και συνδυαστικής βελτιστοποίησης, τόσο αιτιοκρατικών, όσο και στοχαστικών. Οι συγκεκριμένες εφαρμογές του Δυναμικού Προγραμματισμού είναι πάμπολλες και εκτείνονται σε διάφορες περιοχές (Επιχειρησιακή Έρευνα, Τεχνικές Επιστήμες, Οικονομικές Επιστήμες) και σε διάφορους τύπους προβλημάτων (οργάνωση, σχεδιασμός, αυτόματος έλεγχος).

Για την κατανόηση της Αρχής του Βέλτιστου, θεωρούμε το πρόβλημα της βέλτιστης μεταφοράς της αρχικής κατάστασης $\mathbf{x}(0) = \mathbf{x}_0$ ενός δυναμικού συστήματος (βλ. Παράρτημα Α) σε μια τελική τροχιά $\mathbf{g}[\mathbf{x}(t_e), t_e] = \mathbf{0}$ λαμβάνοντας υπόψη μια σειρά περιορισμών[1]. Η ακριβής διατύπωση αυτού του **Προβλήματος Βέλτιστου Ελέγχου (ΠΒΕ*)*** δίδεται στο Παράρτημα Β. Έστω $\mathbf{u}^*(t)$ και $\mathbf{x}^*(t)$, $t \in [0, t_e]$, οι βέλτιστες τροχιές των μεταβλητών ελέγχου και κατάστασης, αντίστοιχα, του δυναμικού συστήματος. Ο R. E. Bellman διατύπωσε το 1952 την εξής **Αρχή του Βέλτιστου**:

> Κάθε **υπόλοιπο** $\mathbf{u}^*(t)$, $t \in [t_1, t_e]$, $0 \le t_1 \le t_e$, της βέλτιστης τροχιάς ελέγχου $\mathbf{u}^*(t)$, $t \in [0, t_e]$, είναι μια βέλτιστη τροχιά για την μεταφορά της αντίστοιχης ενδιάμεσης κατάστασης $\mathbf{x}^*(t_1)$ στην τελική τροχιά $\mathbf{g}[\mathbf{x}(t_e), t_e] = \mathbf{0}$.

Για την επεξήγηση της Αρχής του Βέλτιστου θεωρούμε το Σχήμα 1.1, όπου διαιρούμε την βέλτιστη τροχιά κατάστασης $\mathbf{x}^*(t)$ σε δύο μέρη, 1 για $t \in [0, t_1]$ και 2 για $t \in [t_1, t_e]$. Η Αρχή του Βέλτιστου αποφαίνεται ότι το δεύτερο μέρος είναι συγχρόνως η (πλήρης) βέλτιστη τροχιά ενός άλλου προβλήματος, που επιδιώκει την βέλτιστη μεταφορά της κατάστασης $\mathbf{x}^*(t_1)$ στην τελική τροχιά $\mathbf{g}[\mathbf{x}(t_e),t_e] = \mathbf{0}$ υπό τους ίδιους περιορισμούς του αρχικού προβλήματος. Πράγματι, αν δεν ίσχυε η διαπίστωση αυτή, τότε θα έπρεπε να υπάρχει μια άλλη, τρίτη, τροχιά, που οδηγεί από το σημείο $\mathbf{x}^*(t_1)$ στην τελική τροχιά $\mathbf{g}[\mathbf{x}(t_e), t_e] = \mathbf{0}$ με μικρότερο κόστος από την δεύτερη



Σχήμα 1.1: Επεξήγηση της Αρχής του Βέλτιστου.

[1] Η Αρχή του Βέλτιστου μπορεί να αποδοθεί κατά παρόμοιο τρόπο και για ΠΒΕ διακριτού χρόνου.

τροχιά. Τότε όμως και η τροχιά 1+3 θα έπρεπε να έχει μικρότερο κόστος από την τροχιά 1+2, πράγμα που αντιτίθεται στην αρχική μας υπόθεση πώς η τροχιά 1+2 είναι βέλτιστη.

Αυτή η απλή ιδιότητα βέλτιστων τροχιών έχει μια σειρά από ενδιαφέρουσες συνέπειες που οδηγούν σε χρήσιμα θεωρητικά αποτελέσματα καθώς και σε μεθόδους επίλυσης του προβλήματος βέλτιστου ελέγχου και άλλων προβλημάτων βελτιστοποίησης που φέρουν τη γενική ονομασία **Δυναμικός Προγραμματισμός**. Το ιστορικό της ονομασίας αυτής, που σήμερα μπορεί να παρερμηνευθεί σαν προγραμματισμός υπολογιστού, έδωσε ο Bellman σε ένα άρθρο που δημοσιεύθηκε λίγο μετά τον ξαφνικό του θάνατο το 1984 :

> *"An interesting question is, "Where did the name dynamic programming (DP) come from?" The technique really is for solving multistage stochastic decision processes[2]. The word* ***stochastic*** *is not used lightly. DP was developed for stochastic processes. It was realized fairly late that DP can also be used to solve deterministic problems.*
>
> *The 1950s were not good years for mathematical research. We had a very interesting gentleman in Washington as the Secretary of Defense, and he actually had a pathological fear and hatred of the word* ***research****. His face would suffuse, he would turn red, and he would get violent if people used the term research in his presence. One can imagine how he felt, then, about the term* ***mathematical****. The RAND Corporation[3] was employed by the Air Force, and the Air Force had the Secretary of Defense as its boss, essentially. Hence, Dr. Bellman felt that he had to do something to shield the Air Force from the fact that he was really doing mathematics inside the RAND Corporation. What title, what name, could one choose? In the first place, one was interested in planning, in decision making, in thinking. But planning is not a good word for various reasons. He decided, therefore, to use the word programming. He wanted to get across the idea that this was dynamic, this was multistage, this was time-varying; let's kill two birds with one stone. Let's take a word that has an absolutely precise meaning, namely* ***dynamic****, in the classical, physical sense. It also has a very interesting property as an adjective, that is, it's impossible to use the word* ***dynamic*** *in the pejorative sense. Try thinking of some combination that will possibly give it a pejorative meaning. It's impossible. Thus,* ***dynamic programming*** *was a good name."*

[2] βλ. Κεφάλαιο 7.

[3] Ο Bellman εργάζονταν την εποχή εκείνη για αυτή την εταιρεία.

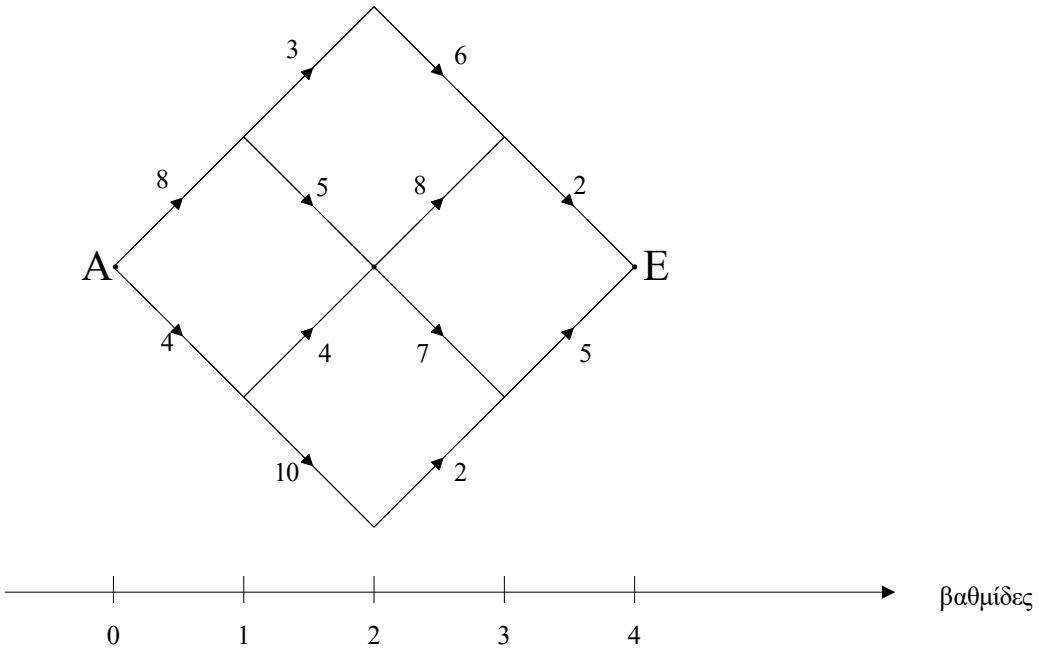
2. ΣΥΝΔΥΑΣΤΙΚΑ ΠΡΟΒΛΗΜΑΤΑ ΚΑΙ ΠΑΙΓΝΙΑ

2.1 Προβλήματα Βραχύτατων Μονοπατιών

Η Αρχή του Βέλτιστου βρίσκει άμεση εφαρμογή στην επίλυση μιας κατηγορίας συνδυαστικών προβλημάτων βελτιστοποίησης, τα λεγόμενα **διακριτά προβλήματα αποφάσεων πολλαπλών βαθμίδων**, όπου η μετάβαση από ένα σημείο μιας βαθμίδας στην επόμενη βαθμίδα είναι δυνατή μέσω ενός πεπερασμένου αριθμού εναλλακτικών αποφάσεων. Κάθε απόφαση, και μαζί της η αντίστοιχη μετάβαση στην επόμενη βαθμίδα, συνδέεται με ένα αντίστοιχο κόστος. Το πρόβλημα έγκειται στην ελαχιστοποίηση του συνολικού κόστους κατά τη μετάβαση από μια αρχική βαθμίδα 0 στην τελική βαθμίδα K. Η προβληματική αυτή και η μέθοδος επίλυσης μέσω Δυναμικού Προγραμματισμού γίνονται πιο κατανοητά μέσα από ένα παράδειγμα.

Παράδειγμα 2.1: Ένα όχημα ξεκινά από το σημείο A του δικτύου μονοδρόμων που απεικονίζει το Σχήμα 2.1 και έχει προορισμό το σημείο E. Οι αριθμοί στους συνδέσμους του δικτύου ανταποκρίνονται στους αντίστοιχους χρόνους διαδρομής ενώ τα βέλη δείχνουν την αντίστοιχη κατεύθυνση των μονοδρόμων (πάντα από αριστερά προς τα δεξιά). Κάθε κόμβος του δικτύου μπορεί να θεωρηθεί ότι ανήκει σε μια βαθμίδα, και σε κάθε κόμβο υπάρχουν το πολύ δύο εναλλακτικές δυνατότητες (αποφάσεις) για τη μετάβαση στην επόμενη βαθμίδα. Ποιά πορεία πρέπει να ακολουθήσει το όχημα, εκκινώντας από την βαθμίδα 0 (σημείο A) για να φθάσει στη βαθμίδα 4 (σημείο E) σε **ελάχιστο χρόνο**;

Ένας πρώτος τρόπος επίλυσης του προβλήματος είναι η σύγκριση **όλων** των δυνατών διαδρομών και η επιλογή της συντομοτάτης. Στο δίκτυο του Σχ. 2.1 έχουμε πραγματικά μόνον 6 δυνατές διαδρομές (ποιές;), σε μεγαλύτερα δίκτυα όμως ο απαιτούμενος υπολογιστικός φόρτος για τον προσδιορισμό και την εκτίμηση **όλων** των διαδρομών είναι υπέρογκος.



Σχήμα 2.1: Δίκτυο μονοδρόμων και χρόνοι διαδρομής για το Παραδ. 2.1.

Ένας δεύτερος τρόπος επίλυσης είναι η συστηματική εφαρμογή της Αρχής του Βέλτιστου. Προς το σκοπό αυτό αρχίζουμε με το σημείο προορισμού Ε και απαντούμε, μεταβαίνοντας βαθμιαία προς τα πίσω, για κάθε κόμβο τις ακόλουθες ερωτήσεις:

– Ποιός είναι ο ελάχιστος χρόνος διαδρομής από αυτόν τον κόμβο μέχρι το σημείο Ε;
– Ποιά είναι η αντίστοιχη ακολουθητέα κατεύθυνση σ' αυτόν τον κόμβο;

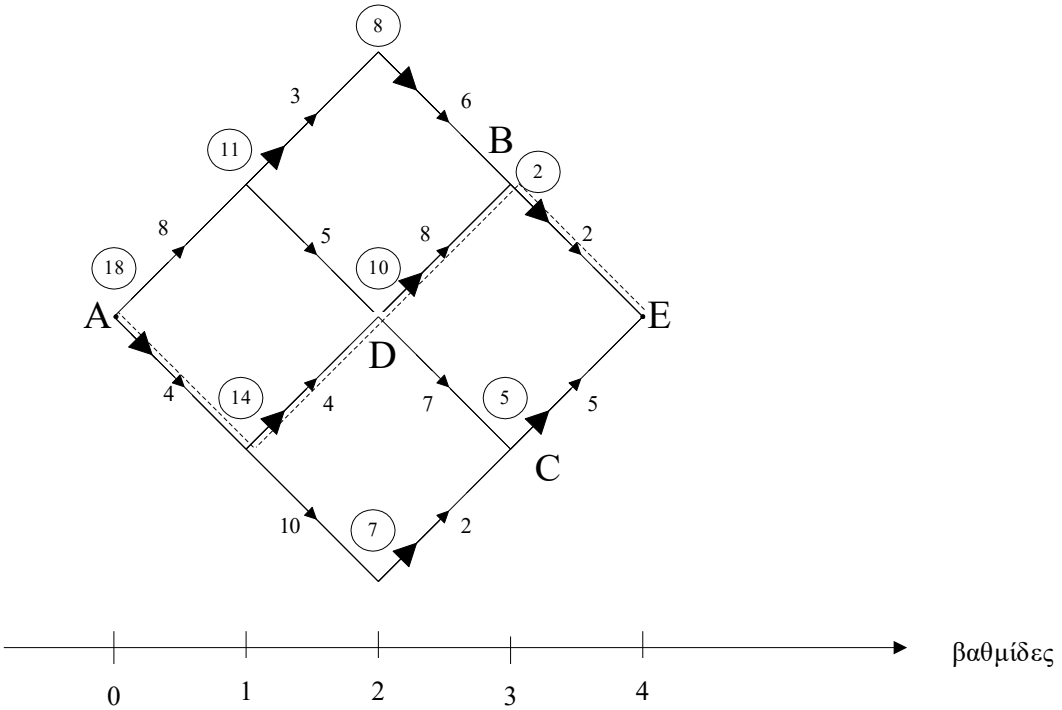
Για την ακριβέστερη περιγραφή της διαδικασίας αυτής θεωρούμε και πάλι το οδικό δίκτυο του προβλήματος στο Σχ. 2.2. Η απάντηση των παραπάνω ερωτήσεων για τους κόμβους Β, C της βαθμίδος 3 είναι προφανής. Σημειώνουμε τους αντίστοιχους ελάχιστους χρόνους **2** και **5** σε κύκλους πάνω από τους δύο κόμβους και υποδεικνύουμε τις αντίστοιχες ακολουθητέες κατευθύνσεις με μεγάλα βέλη.

Αφού, με τον τρόπο αυτό, έχουμε επεξεργαστεί όλους τους κόμβους της βαθμίδας 3, μεταφερόμαστε πίσω κατά μία βαθμίδα, στη βαθμίδα 2. Ας θεωρήσουμε εδώ, παραδειγματικά, τον κόμβο D. Προσβλέποντας στο Ε, το όχημα έχει την δυνατότητα στον κόμβο D να πάρει την αριστερή ή την δεξιά κατεύθυνση. Αν κατευθυνθεί προς τα αριστερά, η διάρκεια διαδρομής μέχρι το Ε είναι ίση προς το άθροισμα δύο χρόνων:

– πρώτον, του χρόνου διαδρομής του συνδέσμου DB, δηλαδή 8
– δεύτερον, του ελάχιστου χρόνου διαδρομής από το Β ως το Ε, δηλαδή **2**.

Συμπερασματικά έχουμε διάρκεια διαδρομής από το D ως το Ε από την αριστερή κατεύθυνση 8 + **2** = 10. Από την δεξιά κατεύθυνση έχουμε παρόμοια 7 + **5** = 12. Άρα, η ελάχιστη διάρκεια διαδρομής από το σημείο D επιτυγχάνεται προς τα αριστερά και η διάρκειά της είναι 10. Υποδεικνύουμε λοιπόν με ένα μεγάλο βέλος την βέλτιστη κατεύθυνση στο D και σημειώνουμε την ελάχιστη διάρκεια **10** πάνω από τον κόμβο.

Κατά πανομοιότυπο τρόπο μπορούμε να προσδιορίσουμε τα βέλη βέλτιστης κατεύθυνσης και τους αντίστοιχους ελάχιστους χρόνους για όλους τους κόμβους της βαθμίδας 2 και εν συνεχεία και των βαθμίδων 1 και 0, οπότε και έχουμε επιλύσει το πρόβλημα. Η βέλτιστη διαδρομή, που



Σχήμα 2.2: Εφαρμογή της Αρχής του Βέλτιστου.

Αριθμός συνδέσμων ανά πλευρά τετραγώνου	n	2	3	5	10
Αριθμός διαδρομών	$(2n)!/(n!)^2$	6	20	252	$1.85\ 10^5$
Αριθμός κόμβων αποφάσεων	$(n + 1)^2–1$	8	15	35	120

Πίνακας 2.1: Υπολογιστικός φόρτος για τετράγωνο δίκτυο.

σημειώνεται με διακεκομμένη γραμμή στο Σχ. 2.2, μπορεί να προσδιορισθεί αν εκκινήσει κανείς στο Α και, ακολουθώντας τα βέλη, καταλήξει στο Ε.

Για την επίλυση του προβλήματος απαιτήθηκε η επεξεργασία συνολικά 8 κόμβων αποφάσεων. Ο Πίνακας 2.1 δείχνει την αύξηση του υπολογιστικού φόρτου σε συνάρτηση του αριθμού συνδέσμων ανά πλευρά του τετράγωνου δικτύου για τους δύο αναφερθέντες τρόπους επίλυσης. Βλέπουμε πως η χρήση της Αρχής του Βέλτιστου μειώνει σημαντικά τον υπολογιστικό φόρτο για την επίλυση προβλημάτων αποφάσεων πολλαπλών βαθμίδων σε μεγάλα δίκτυα συγκρινόμενη με την εκτίμηση όλων των δυνατών διαδρομών.

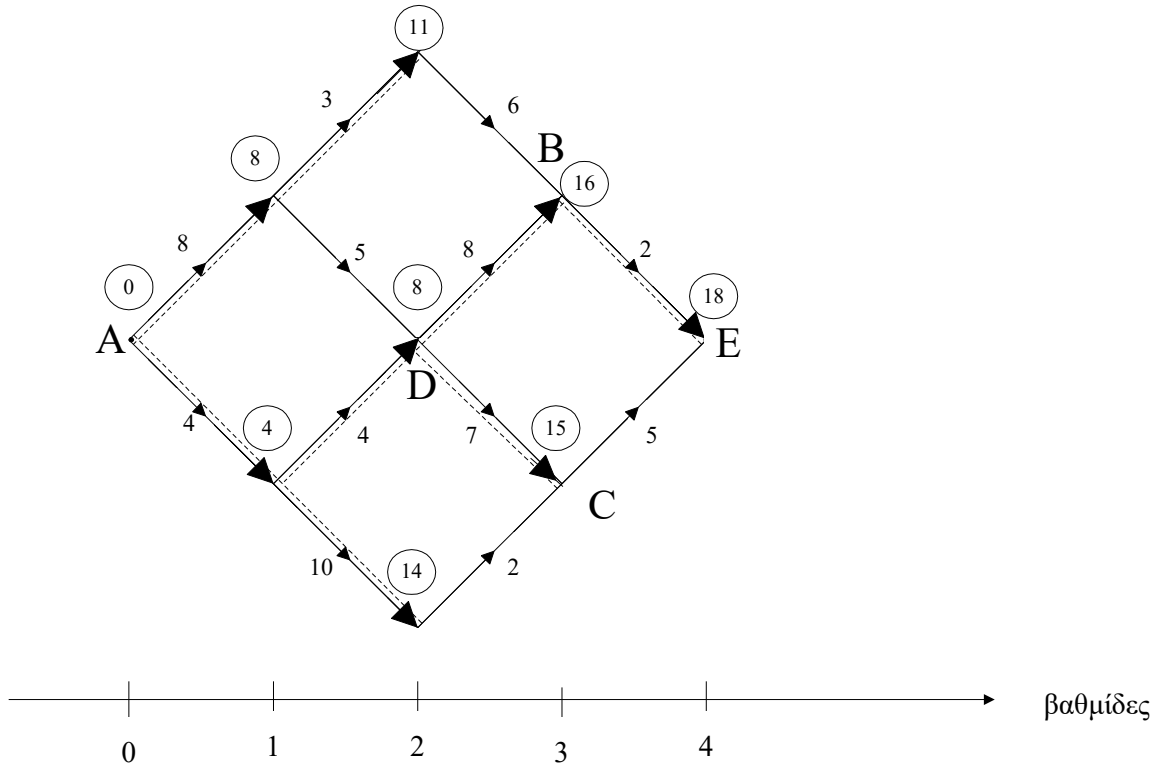
■

Το πρόβλημα του Παραδ. 2.1 ανήκει στη γενικότερη κατηγορία προβλημάτων **βραχύτατων μονοπατιών** σε γράφους με οποιαδήποτε δομή, όπου ζητείται η βραχύτατη σύνδεση μεταξύ δύο κόμβων. Όπως είδαμε, ο Δυναμικός Προγραμματισμός βρίσκει εφαρμογή σε προβλήματα βραχύτατων μονοπατιών, υπό την προϋπόθεση ότι ο αντίστοιχος γράφος είναι ή μπορεί να μετασχηματισθεί σε σύστημα αποφάσεων πολλαπλών βαθμίδων. Σε μερικές περιπτώσεις αυτό επιτυγχάνεται με την εισαγωγή νέων, εικονικών κόμβων και συνδέσμων με μηδενική τιμή. Όμως η Αρχή του Βέλτιστου μπορεί να εφαρμοσθεί και άμεσα σε γράφους χωρίς βαθμίδες, όταν η διέλευση επί όλων των συνδέσμων είναι δυνατή μόνο προς την ίδια κατεύθυνση (προς τα εμπρός), χωρίς παρουσία κύκλων εντός του δικτύου (βλ. Προβλ. 2.2 b).

Ας σημειωθεί ότι η βαθμιαία μέθοδος που εφαρμόστηκε στο Παραδ. 2.1 οδήγησε σε ένα **απόλυτο ελάχιστο** του προβλήματος, καθότι συγκρίθηκαν, εμμέσως, όλες οι δυνατές διαδρομές και επελέγει η συντομοτάτη.

Σημαντική είναι επίσης η παρατήρηση ότι η επιτευχθείσα λύση του Παραδ. 2.1 εμπεριέχει όχι μόνο την βραχύτατη διαδρομή (τροχιά) από το Α στο Ε, αλλά και την βέλτιστη ακολουθητέα κατεύθυνση (απόφαση) σε **όλους τους κόμβους** του δικτύου. Η λύση αυτή είναι επομένως ένας **βέλτιστος έλεγχος κλειστού βρόχου** με την έννοια του Παραρτ. Β.2. Πράγματι, η εφαρμογή της Αρχής του Βέλτιστου οδηγεί, σε όλους τους τύπους προβλημάτων όπου εφαρμόζεται, κατά κανόνα σε λύση βέλτιστου ελέγχου κλειστού βρόχου, όπως θα διαπιστώσουμε και στα επόμενα κεφάλαια.

Η περιγραφείσα διαδικασία του Δυναμικού Προγραμματισμού μπορεί βέβαια να εφαρμοσθεί και αντίστροφα, αρχίζοντας δηλαδή από την βαθμίδα 0 και προχωρώντας **έμπροσθεν**, βαθμίδα προς βαθμίδα, μέχρι την τελική βαθμίδα Κ, όπως φαίνεται στο Σχ. 2.3. Τότε όμως, το αναγραφόμενο (σε κύκλο) βέλτιστο κόστος κάθε κόμβου έχει άλλη σημασία: Υποδηλώνει το ελάχιστο κόστος μετάβασης από τον αρχικό κόμβο (Α) σε κάθε κόμβο του δικτύου. Επίσης, το παχύ βέλος στην περίπτωση αυτή υποδηλώνει την βέλτιστη κατεύθυνση **άφιξης** στον αντίστοιχο κόμβο. Επομένως η αντίστροφη αυτή διαδικασία δεν οδηγεί σε κανόνα ελέγχου



Σχήμα 2.3: Έμπροσθεν εφαρμογή του Δυναμικού Προγραμματισμού στο Παραδ. 2.1.

κλειστού βρόχου, αλλά στον προσδιορισμό **βέλτιστων τροχιών** από τον αρχικό κόμβο Α **προς κάθε κόμβο του δικτύου**, συμπεριλαμβανομένου βεβαίως και του τελικού κόμβου (Ε), βλ. διακεκομμένες γραμμές στο Σχ. 2.3. Για να προσδιορισθεί η βέλτιστη διαδρομή που οδηγεί σε κάποιο κόμβο του δικτύου (π.χ. στον Ε), εκκινούμε από τον κόμβο αυτό και προχωρούμε όπισθεν επί του συνδέσμου που φέρει το παχύ βέλος, προς την προηγούμενη βαθμίδα (π.χ. από τον Ε προς τον Β) κ.ο.κ. μέχρι το αρχικό σημείο Α.

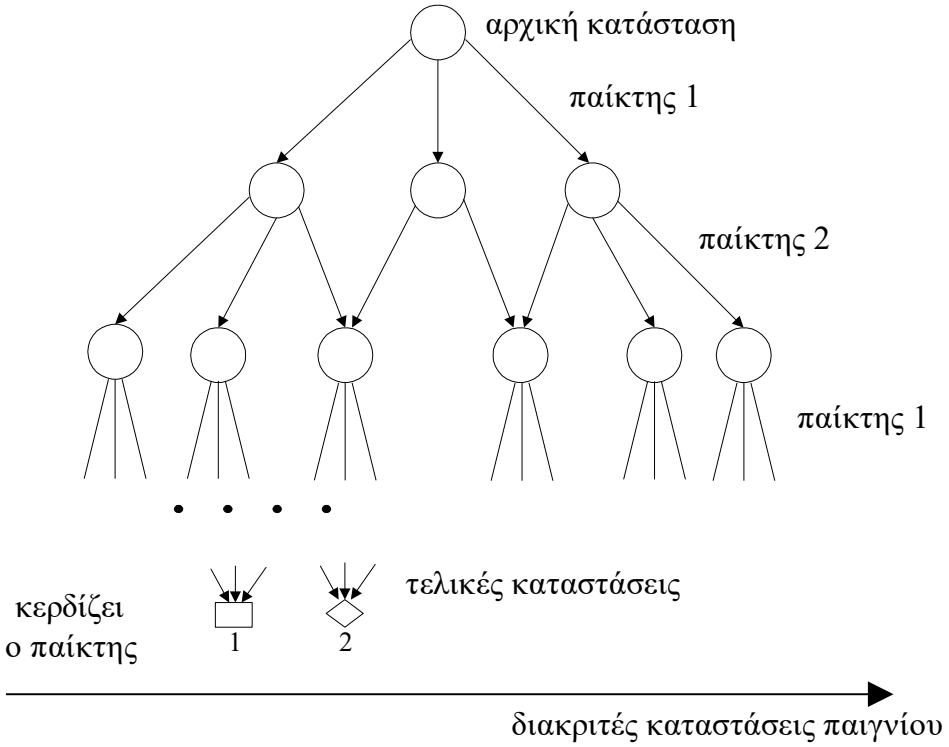
2.2 Παίγνια Δύο Προσώπων

Μια μορφή Δυναμικού Προγραμματισμού είναι εφαρμόσιμη σε παίγνια δύο προσώπων με τα ακόλουθα χαρακτηριστικά:

- Πλήρης πληροφορία αμφότερων των παικτών, δηλαδή κάθε παίκτης γνωρίζει την εξέλιξη και την παρούσα κατάσταση του παιχνιδιού.
- Οι παίκτες παίζουν εναλλάξ και έχουν την δυνατότητα να επιλέξουν μεταξύ ενός πεπερασμένου αριθμού γνωστών διακριτών αποφάσεων (κινήσεων).
- Η τελική έκβαση του παιγνίου είναι μια διακριτή κατάσταση (π.χ. κερδίζει ο παίκτης 1) από ένα πεπερασμένο σύνολο διακριτών τελικών καταστάσεων.

Στην κατηγορία αυτή ανήκει επί παραδείγματι το σκάκι, η ντάμα και άλλα αιτιοκρατικά παίγνια. Εννοείται ότι υπό τον όρο παίγνιο μπορούν να μοντελοποιηθούν διεργασίες ανταγωνισμού στο οικονομικό, πολιτικό, στρατιωτικό και άλλα πεδία. Η μεθοδολογία που θα παρουσιασθεί σ' αυτό το κεφάλαιο μπορεί εύκολα να επεκταθεί σε στοχαστικά παίγνια (π.χ. τάβλι) κάνοντας χρήση των κατάλληλων εννοιών και αλγορίθμων του Κεφαλαίου 7.

Για την κατανόηση της μεθοδολογίας θεωρούμε την ανάπτυξη κάποιου παιγνίου στο Σχ. 2.4.



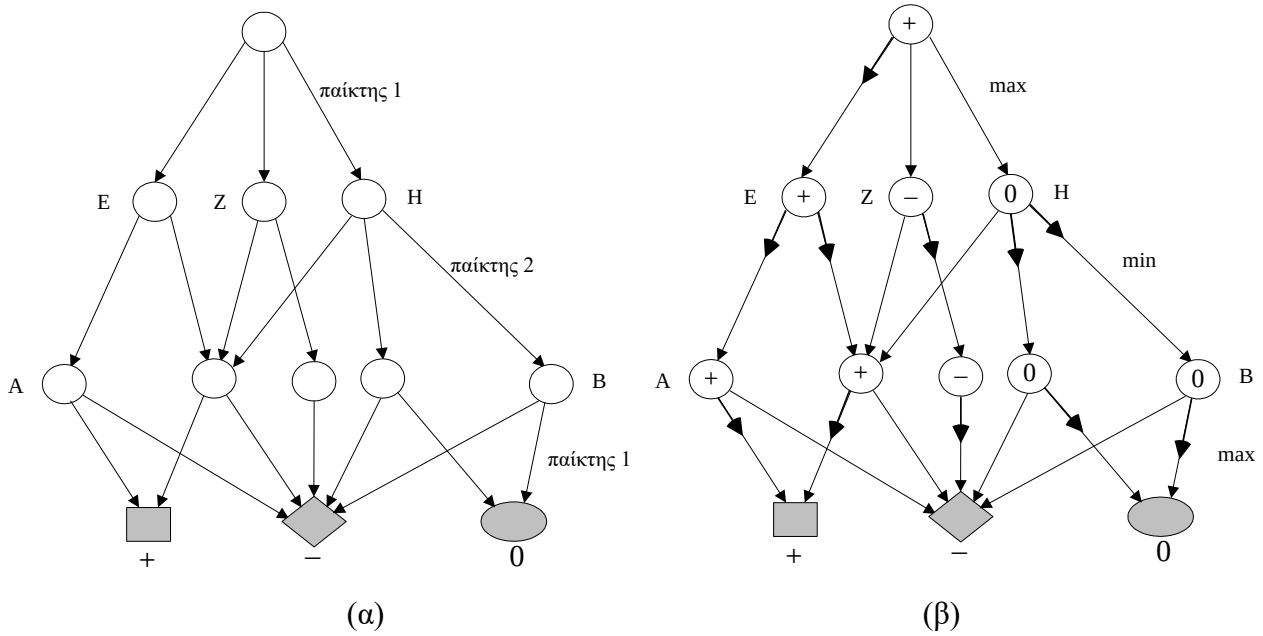
Σχήμα 2.4: Ανάπτυξη παιγνίου δύο προσώπων.

Εκκινώντας από μια αρχική κατάσταση θεωρούμε ότι ο παίκτης 1 έχει να επιλέξει μεταξύ τριών δυνατών κινήσεων που οδηγούν σε αντίστοιχες νέες καταστάσεις του παιγνίου. Εν συνεχεία καλείται να επιλέξει ο παίκτης 2, κατόπιν και πάλι ο παίκτης 1 κ.ο.κ. μέχρι την τελική έκβαση του παιγνίου, που μπορεί να αναδείξει νικητή τον παίκτη 1 ή 2 ανάλογα με τις επιλεχθείσες κινήσεις. Προφανής στόχος κάθε παίκτη είναι να επηρεάσει την εξέλιξη του παιγνίου έτσι ώστε να είναι ο τελικός νικητής.

Η πλήρης ανάπτυξη κατά το Σχ. 2.4 δεν είναι δυνατή παρά μόνο σε πολύ απλά παίγνια. Για το σκάκι θα απαιτούνταν περίπου 10^{120} ενδιάμεσοι κόμβοι. (Ας σημειωθεί ότι 10^{80} είναι το σύνολο των ατόμων στο γνωστό σύμπαν.) Έτσι ένας υπολογιστής με χρόνο επεξεργασίας 0.3 ns για κάθε κόμβο, θα απαιτούσε 10^{21} αιώνες υπολογισμών για τον προσδιορισμό της βέλτιστης στρατηγικής. Μια εναλλακτική, υποβέλτιστη δυνατότητα είναι να αναπτύσσεται το παίγνιο σε κάθε κίνηση ενός παίκτου, εκκινώντας από την παρούσα (αντί της αρχικής) κατάσταση και μέχρι ένα συγκεκριμένο βάθος. Για παράδειγμα, το πρόγραμμα σκακιού Deep Thought (έκδοση 1991) ανέπτυσσε την παρούσα κατάσταση πριν από κάθε κίνηση σε βάθος 15 ως 30 κινήσεων με ταχύτητα επεξεργασίας 10^9 κόμβων ανά δευτερόλεπτο.

Το ακόλουθο παράδειγμα δείχνει τον τρόπο υπολογισμού της βέλτιστης στρατηγικής στην περίπτωση πλήρους ανάπτυξης του παιγνίου, ενώ στη συνέχεια θα ασχοληθούμε και με τον υπολογισμό υποβέλτιστων στρατηγικών βάσει μερικής ανάπτυξης του παιγνίου.

Παράδειγμα 2.2: Θεωρούμε το αναπτυγμένο παίγνιο του Σχ. 2.5α από την σκοπιά του παίκτη 1 και επιθυμούμε να υπολογίσουμε μια βέλτιστη στρατηγική κλειστού βρόχου για τον παίκτη αυτό. Θεωρούμε (πράγμα όχι απαραίτητο) ότι ο παίκτης 1 ανοίγει το παιγνίδι και ότι υπάρχουν τρεις δυνατές τελικές καταστάσεις: νίκη, ήττα (δηλαδή νίκη του παίκτη 2) και ισοπαλία.



Σχήμα 2.5: Παίγνιο του Παραδ. 2.2 (α) και η λύση του (β).

Για τον υπολογισμό της βέλτιστης στρατηγικής για τον παίκτη 1, τοποθετούμε καταρχήν τα +, – και 0 στους τελικούς κόμβους όπου ο παίκτης 1 κερδίζει, χάνει ή έρχεται ισόπαλος, αντίστοιχα, και εκκινούμε (ως συνήθως στον Δυναμικό Προγραμματισμό) από την τελική βαθμίδα. Στον κόμβο A ο παίκτης 1 έχει δύο επιλογές και προφανώς θα επιλέξει εκείνη που οδηγεί στο +. Άρα εγγράφουμε + στον κόμβο αυτό (βλ. Σχ. 2.5β) και χαρακτηρίζουμε την βέλτιστη κίνηση με ένα παχύ βέλος. Στο σημείο B ο παίκτης 1 έχει να επιλέξει μεταξύ ήττας και ισοπαλίας και προφανώς θα επιλέξει την ισοπαλία. Κατά τον ίδιο τρόπο μπορούμε να επεξεργαστούμε όλους τους κόμβους της βαθμίδας αυτής και έχουμε τα αντίστοιχα αποτελέσματα στο Σχ. 2.5β. Προχωρώντας στην αμέσως προηγούμενη βαθμίδα θεωρούμε τις πιθανές κινήσεις του παίκτη 2 στο σημείο Z. Έχοντας τις επιλογές της επακόλουθης νίκης του (δηλαδή –) ή ήττας του (δηλαδή +), ο παίκτης 2 θα επιλέξει προφανώς –, πράγμα που εγγράφεται, μαζί με το αντίστοιχο βέλος, στον κόμβο Z. Κατά παρόμοιο τρόπο υπολογίζονται + και 0 για τους κόμβους E και H αντίστοιχα. Ας σημειωθεί ότι η βέλτιστη κίνηση για τον παίκτη 2 στους κόμβους E και H δεν είναι μοναδική. Προχωρώντας μια ακόμη βαθμίδα προς τα πίσω, υπολογίζουμε κατά παρόμοιο τρόπο την βέλτιστη κίνηση του παίκτη 1 στην αρχική κατάσταση (Σχήμα 2.5β) και ευρίσκουμε +. Αφού η αρχική κατάσταση χαρακτηρίζεται από +, ο παίκτης 1 μπορεί να είναι σίγουρος ότι θα κερδίσει το παίγνιο αν ακολουθήσει την υπολογισθείσα βέλτιστη στρατηγική, όποια και αν είναι η αντίδραση του παίκτη 2.

■

Το παράδειγμα κάνει φανερό ότι η διαδικασία προσδιορισμού βέλτιστης πολιτικής για τα εξεταζόμενα παίγνια αντιστοιχεί σε βασικές γραμμές με εκείνη του Δυναμικού Προγραμματισμού με δύο επί μέρους διαφορές:

– Το κόστος μετάβασης από κατάσταση σε κατάσταση του παιγνίου είναι συνήθως μηδενικό.
– Σε κάθε βαθμίδα του παιγνίου εφαρμόζεται εναλλάξ μεγιστοποίηση ή ελαχιστοποίηση, ανάλογα με τον παίκτη που αποφασίζει.

Παράδειγμα 2.3: Θεωρούμε το παίγνιο των επτά σπίρτων που έχει ως εξής: Επτά σπίρτα τοποθετούνται πάνω στο τραπέζι και δύο παίκτες πρέπει να σηκώνουν, παίζοντας εναλλάξ,

από ένα ως τρία σπίρτα κάθε φορά. Χάνει αυτός που είναι υποχρεωμένος να σηκώσει το τελευταίο σπίρτο.

Το Σχ. 2.6 αναπαριστά την πλήρη ανάπτυξη, εκτίμηση και λύση (βέλτιστη στρατηγική) του παιγνίου, ειδωμένου από την σκοπιά του παίκτη που παίζει πρώτος. Οι αριθμοί μέσα στα τετράγωνα-κόμβους είναι οι αριθμοί σπίρτων των αντίστοιχων καταστάσεων. Οι τελικοί κόμβοι με ένα σπίρτο εκτιμώνται για προφανείς λόγους με +, αν πρόκειται για βαθμίδα ελαχιστοποίησης, και με –, αν πρόκειται για βαθμίδα μεγιστοποίησης.

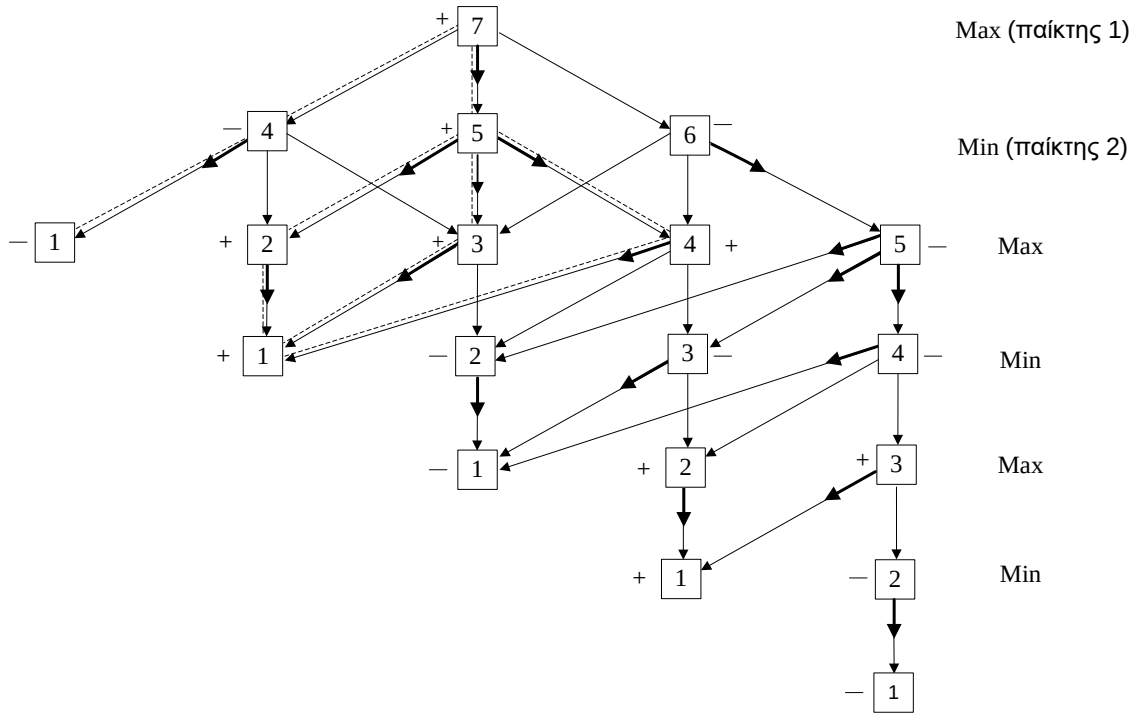
■

Από τα παραπάνω παραδείγματα γεννάται το εύλογο ερώτημα της έκβασης του παιγνίου αν **αμφότεροι** οι παίκτες εφαρμόζουν βέλτιστη στρατηγική. Στην περίπτωση αυτή, η εξέλιξη και τελική έκβαση είναι σαφώς προκαθορισμένες και εξαρτώνται από το συγκεκριμένο παίγνιο. Αν οι τελικές καταστάσεις κάποιου συγκεκριμένου παιγνίου είναι νίκη του παίκτη 1 ή 2, τότε το ποιος από τους δύο θα κερδίσει ή θα χάσει, εξαρτάται από το ποιός παίζει πρώτος. Ανάλογα με το παίγνιο, μπορεί αυτός που ανοίγει το παίγνιο να είναι "καταδικασμένος" να κερδίσει ή να χάσει, αν αμφότεροι οι παίκτες εφαρμόζουν βέλτιστη στρατηγική. Επί παραδείγματι στο παίγνιο των επτά σπίρτων, αν ο παίκτης που παίζει πρώτος εφαρμόζει απαρέγκλιτα την βέλτιστη στρατηγική, δεν υπάρχει τρόπος να χάσει.

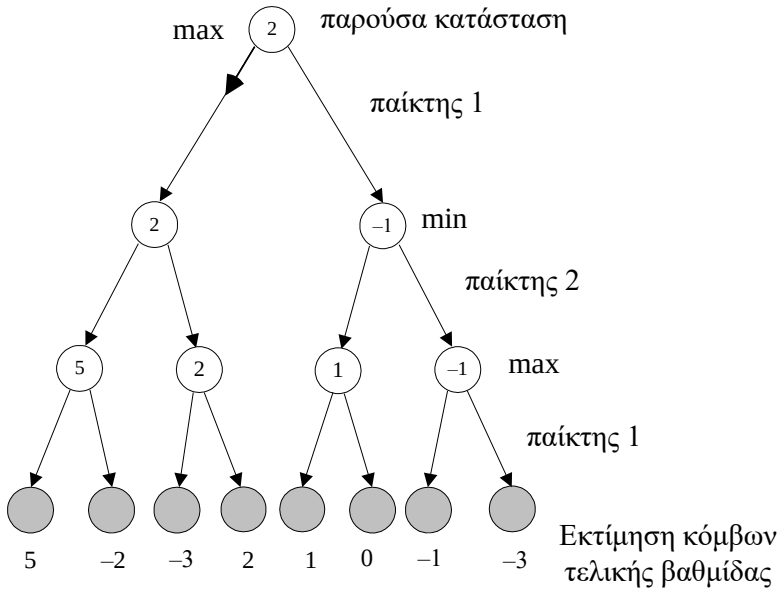
Θα εξετάσουμε τώρα την ακολουθητέα υποβέλτιστη διαδικασία στην περίπτωση παιγνίων που δεν είναι δυνατόν να αναπτυχθούν πλήρως λόγω υψηλής πολυπλοκότητας. **Σε κάθε κίνηση** του παίκτη 1 που μας ενδιαφέρει εκτελούνται τα ακόλουθα βήματα:

(i) Ανάπτυξη του παιγνίου, εκκινώντας από την παρούσα κατάσταση, μέχρι ένα προκαθορισμένο βάθος (π.χ. μέχρι βάθος 3 στο Σχ. 2.7).

(ii) Εκτίμηση (βαθμολογία) των κόμβων της τελευταίας βαθμίδας βάσει εμπειρικών κανόνων, π.χ. κατά τρόπο ώστε θετικές τιμές να αντανακλούν αντίστοιχες ευνοϊκές καταστάσεις για τον παίκτη 1, ενώ αρνητικές τιμές να αντιστοιχούν σε δυσμενείς καταστάσεις του παιγνίου για τον ίδιο παίκτη. Στο Σχ. 2.7 δίδονται, σαν παράδειγμα, στους κόμβους της τελευταίας βαθμίδας τιμές από 5 ως –3.



Σχήμα 2.6: Επίλυση του παιγνίου επτά σπίρτων.



Σχήμα 2.7: Μερική ανάπτυξη παιγνίου.

(iii) Εφαρμογή της διαδικασίας ελαχιστο-μεγιστοποίησης σε κάθε κόμβο κάθε βαθμίδας, εκκινώντας από την προτελευταία βαθμίδα και λαμβάνοντας υπόψη την εκτίμηση της τελευταίας βαθμίδας που προσδιορίσθηκε στο βήμα (ii). Η διαδικασία αυτή οδηγεί στις τιμές των κόμβων που αναγράφονται στο παράδειγμα του Σχ. 2.7.

(iv) Εκτέλεση της προσδιορισθείσας, πλέον ευνοϊκής κίνησης στην παρούσα κατάσταση (βλ. παχύ βέλος στο Σχ. 2.7) από τον παίκτη 1.

Η περιγραφείσα διαδικασία είναι προφανώς υποβέλτιστη, αφού δεν μπορεί – χωρίς την πλήρη ανάπτυξη του παιγνίου – να εγγυηθεί τη νίκη του παίκτη 1. Η αποτελεσματικότητα της αντίστοιχης στρατηγικής εξαρτάται σε μεγάλο βαθμό από το βάθος ανάπτυξης του παιγνίου σε κάθε κίνηση και από την ευστοχία της βαθμολογίας των κόμβων της αντίστοιχης τελικής βαθμίδας.

Παράδειγμα 2.4: Θεωρούμε το παίγνιο τρίλιζα ή tic-tac-toe ή παίγνιο ΧΟ. Δύο παίκτες τοποθετούν εναλλάξ το αντίστοιχο σήμα τους (Χ για τον παίκτη 1 και Ο για τον παίκτη 2) σε ένα ελεύθερο από τα εννέα μικρά τετράγωνα ενός μεγαλύτερου τετραγώνου (Σχ. 2.8). Κερδίζει εκείνος που έχει καλύψει πρώτος με το σήμα του μια ολόκληρη σειρά (τριάδα) που μπορεί να είναι οριζόντια, κάθετη ή διαγώνια.

Για το παίγνιο αυτό θα θεωρήσουμε ότι σε κάθε κίνηση του παίκτη 1 που μας ενδιαφέρει υπάρχει δυνατότητα ανάπτυξης μέχρι βάθους 2, οπότε απαιτείται η εφαρμογή της παραπάνω περιγραφείσας υποβέλτιστης διαδικασίας. Για την εκτίμηση των κόμβων της δεύτερης (τελευταίας) βαθμίδας κάθε ανάπτυξης, ορίζουμε μια συνάρτηση e ως εξής: e = (αριθμός σειρών χωρίς το σύμβολο Ο του αντιπάλου) μείον (αριθμός σειρών χωρίς το σύμβολο Χ του παίκτου 1). Είναι προφανές ότι θετικές τιμές της e αντιστοιχούν σε ευνοϊκές καταστάσεις για τον παίκτη 1 που μας ενδιαφέρει, ενώ αρνητικές τιμές της e σε δυσμενείς καταστάσεις για τον ίδιο παίκτη. Θα θεωρήσουμε επίσης για προφανείς λόγους ότι e = ∞ σε περίπτωση τελικού κόμβου του παιγνίου που σηματοδοτεί νίκη του παίκτη 1, ενώ e = –∞ στην αντίθετη περίπτωση. Για παράδειγμα, η απεικονιζόμενη στο Σχ. 2.8 κατάσταση του παιγνίου βαθμολογείται με e = 6–4 = 2.

Για να μειώσουμε τον αναγκαίο υπολογιστικό φόρτο, θα θεωρήσουμε κατά την ανάπτυξη όλες τις συμμετρικές καταστάσεις του παιγνίου (βλ. π.χ. Σχ. 2.9) ως ταυτόσημες αφού η σημασία τους για την έκβαση του παιγνίου ουδόλως διαφέρει.

Με τα δεδομένα αυτά, και θεωρώντας ότι ο παίκτης 1 παίζει πρώτος, έχουμε την πρώτη ανάπτυξη του παιγνίου μέχρι βάθους 2 κατά το Σχ. 2.10, όπου και η εκτίμηση των κόμβων της δεύτερης (τελευταίας) βαθμίδας κατά τον ορισμό της συνάρτησης e. Βάσει της εκτίμησης αυτής, η διαδικασία ελαχιστο-μεγιστοποίησης οδηγεί τελικά στην κίνηση που σημειώνεται στην αρχική κατάσταση του Σχ. 2.10, η οποία και εκτελείται.

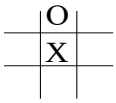
Θεωρούμε τώρα ότι ο παίκτης 2 αντιδρά στην κίνηση αυτή όπως σημειώνεται στην αρχική κατάσταση του Σχ. 2.11, οπότε η ανάπτυξη του παιγνίου μέχρι βάθους 2 και η ελαχιστο-μεγιστοποίηση οδηγούν σε δύο ισότιμες δυνατές κινήσεις του παίκτου 1 που σημειώνονται με παχιά βέλη στην αρχική κατάσταση του Σχ. 2.11. Υποθέτουμε ότι από τις δύο ισότιμες κινήσεις επιλέγεται η δεξιά, οπότε ο παίκτης 2 απαντά με την αρχική κατάσταση του Σχ. 2.12, για να αποφύγει προφανώς την άμεση ήττα του. Μια τελευταία ανάπτυξη του παιγνίου κατά το Σχ. 2.12 μας οδηγεί ήδη σε τελικούς κόμβους: ο παίκτης 1 έχει τώρα πέντε συνολικά επιλογές, εκ των οποίων οι τέσσερις οδηγούν στην άμεση ήττα του, ενώ η πέμπτη του εξασφαλίζει την νίκη.

■

Ας σημειωθεί ότι η πλήρης ανάπτυξη του παιγνίου της τρίλιζας (που είναι εφικτή λόγω περιορισμένης πολυπλοκότητας) οδηγεί στο συμπέρασμα πως, αν αμφότεροι οι παίκτες παίξουν βέλτιστα, το αποτέλεσμα θα είναι ισοπαλία, αφού κανείς από τους δύο παίκτες δεν θα μπορούσε να συμπληρώσει μια τριάδα.

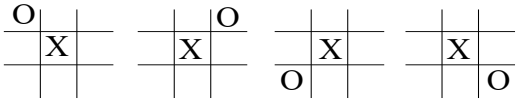
Ας σημειωθεί ότι ο υπολογιστικός φόρτος κατά την εφαρμογή της περιγραφείσας υποβέλτιστης διαδικασίας μπορεί να μειωθεί αισθητά αν τα ως άνω βήματα (i), (ii), (iii) εκτελούνται όχι διαδοχικά, αλλά παράλληλα. Τότε μπορεί να εφαρμοσθεί η λεγόμενη **διαδικασία άλφα-βήτα** που επιτρέπει τον προσδιορισμό της ίδιας υποβέλτιστης λύσης, βάσει μειωμένης όμως ανάπτυξης του παιγνίου.

Για παράδειγμα, στο Σχ. 2.12, αν η ανάπτυξη του παιγνίου και η εκτίμηση των κόμβων της τελευταίας βαθμίδας γίνονται παράλληλα, τότε το γεγονός ότι $e(A) = -\infty$ σημαίνει ότι οι κόμβοι B, C, D δεν χρειάζεται ούτε να αναπτυχθούν ούτε να εκτιμηθούν, αφού δεν πρόκειται προφανώς να επηρεάσουν το αποτέλεσμα της ελαχιστοποίησης.

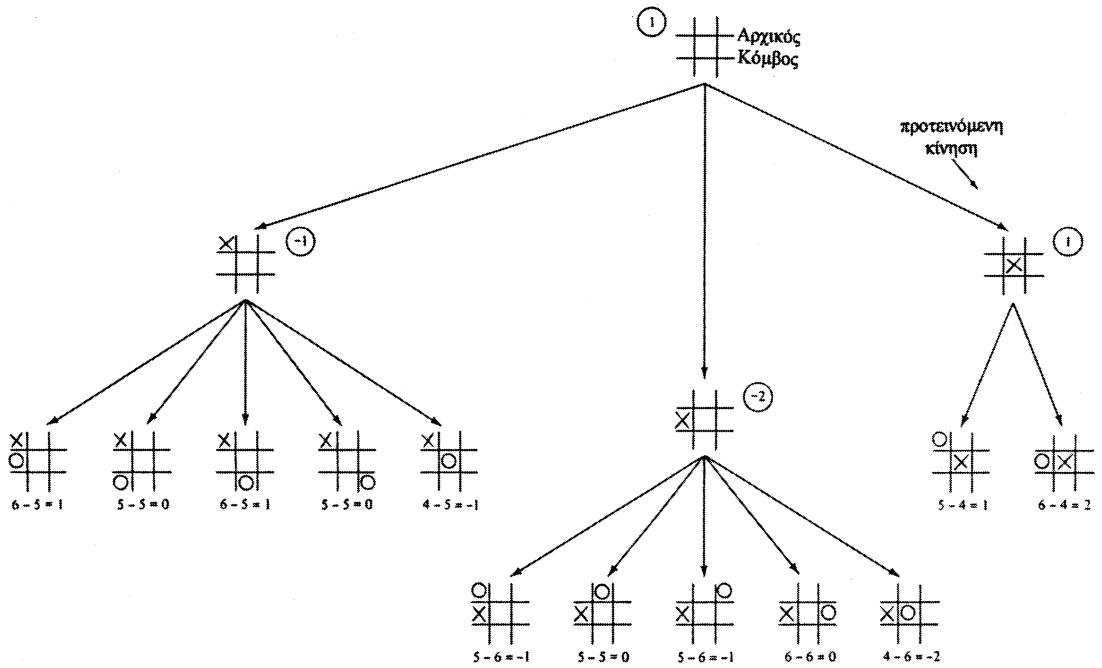


$e = 6 - 4 = 2$

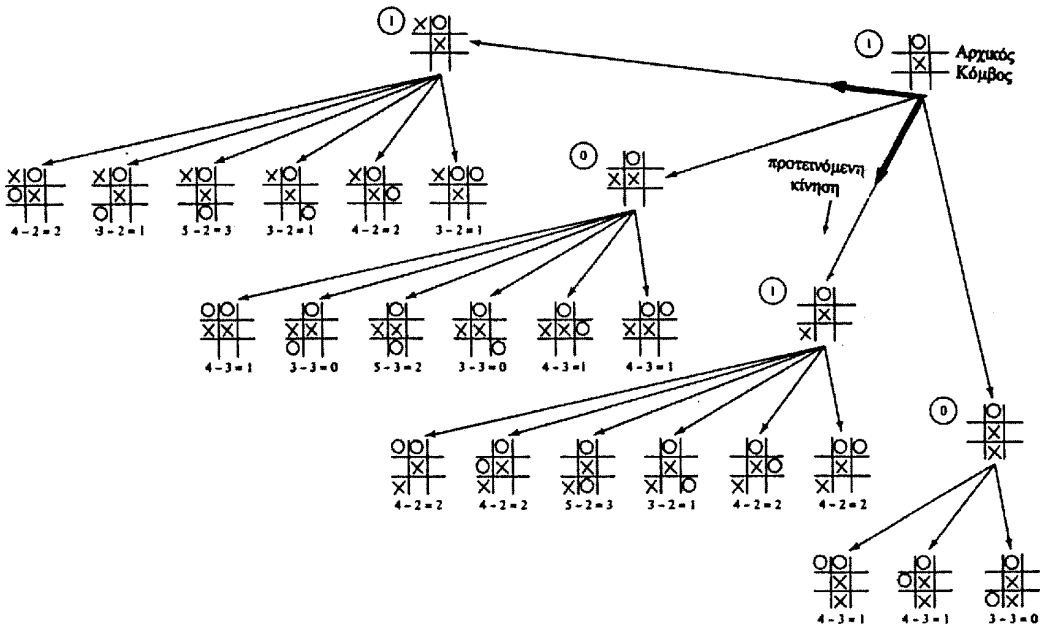
Σχήμα 2.8: Παράδειγμα εκτίμησης κόμβου.



Σχήμα 2.9: Παράδειγμα συμμετρικών, ταυτόσημων καταστάσεων του παιγνίου του Παραδ. 2.4.



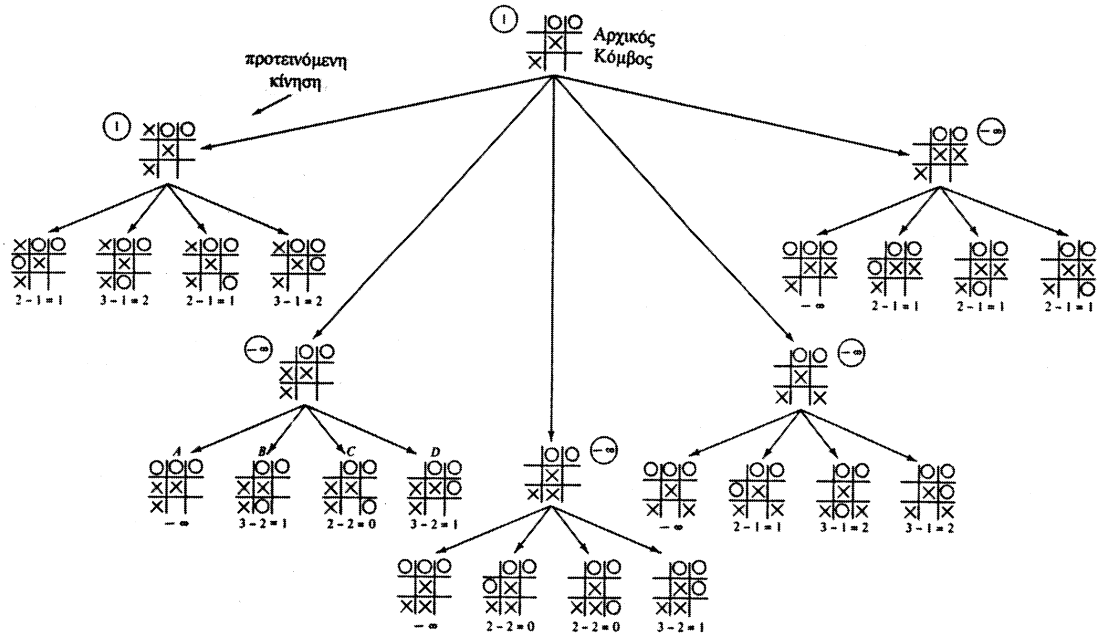
Σχήμα 2.10: Πρώτη ανάπτυξη παιγνίου στο Παραδ. 2.4.



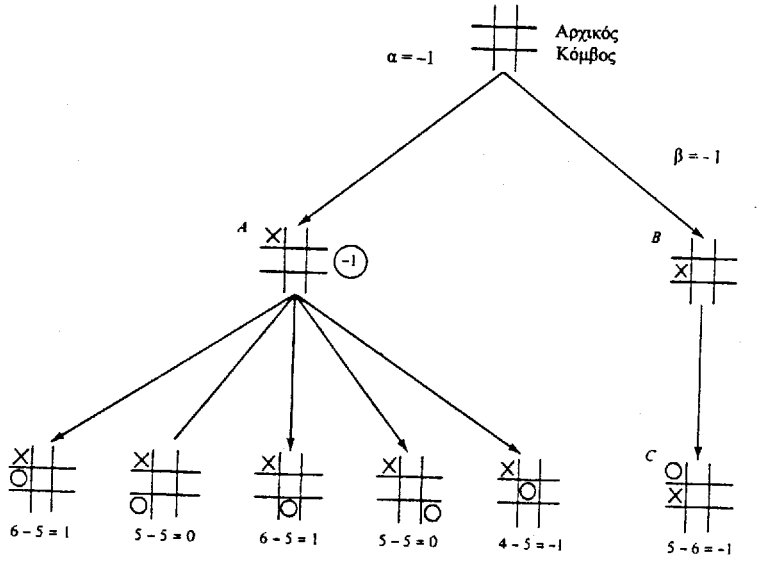
Σχήμα 2.11: Δεύτερη ανάπτυξη παιγνίου στο Παραδ. 2.4.

Σαν δεύτερο παράδειγμα θεωρούμε και πάλι στο Σχ. 2.13 την διαδικασία του Σχ. 2.10. Αφού υπολογισθεί e(A) = –1, όπως ακριβώς και στο Σχ. 2.10, μπορούμε να θέσουμε, μια βαθμίδα υψηλότερα, μια τιμή α = e(A) = –1 γνωρίζοντας ότι στον κόμβο εκείνο (δηλαδή στον αρχικό κόμβο) αναγκαστικά (λόγω μεγιστοποίησης) e ≥ α = –1. Συνεχίζουμε τώρα την ανάπτυξη-

εκτίμηση με τον κόμβο B και εν συνεχεία με τον C. Αφού e(C) = –1, τότε σίγουρα (λόγω ελαχιστοποίησης) e(B) ≤ –1 και γι' αυτό θέτουμε μια τιμή β(B) = –1. Αφού όμως α ≥ β(B), δεν



Σχήμα 2.12: Τρίτη ανάπτυξη παιγνίου στο Παραδ. 2.4.



Σχήμα 2.13: Εφαρμογή της διαδικασίας άλφα-βήτα.

χρειάζεται προφανώς να εξετάσουμε περαιτέρω ανάπτυξη του κόμβου B, πράγμα που εξοικονομεί υπολογιστικό φόρτο.

Συγκεντρωτικά, οι γενικοί κανόνες αποκοπής κλάδων κατά την ανάπτυξη-εκτίμηση του παιγνίου μέσω εφαρμογής της διαδικασίας άλφα-βήτα είναι οι εξής:

(i) Ένας κόμβος ελαχιστοποίησης B για τον οποίο ισχύει $\beta(B) \leq \alpha(E)$, όπου E οποιοσδήποτε προγενέστερός του κόμβος μεγιστοποίησης, μπορεί να αποκοπεί (δηλαδή δεν χρειάζεται να αναπτυχθεί περαιτέρω) θέτοντας $e(B) = \beta(B)$.
(ii) Ένας κόμβος μεγιστοποίησης E για τον οποίο ισχύει $\alpha(E) \geq \beta(F)$, όπου F οποιοσδήποτε προγενέστερός του κόμβος ελαχιστοποίησης, μπορεί να αποκοπεί (δηλαδή δεν χρειάζεται να αναπτυχθεί περαιτέρω) θέτοντας $e(E) = \alpha(E)$.

Ο γενικός τρόπος προσδιορισμού των τιμών α και β είναι ο εξής:
- Η τιμή α ενός κόμβου μεγιστοποίησης ενημερώνεται κατά την διαδικασία ανάπτυξης-εκτίμησης έτσι ώστε να αντιστοιχεί πάντοτε στη μέγιστη τιμή μεταξύ των **τελικώς** εκτιμηθέντων διαδόχων του κόμβων.
- Η τιμή β ενός κόμβου ελαχιστοποίησης ενημερώνεται κατά την διαδικασία ανάπτυξης-εκτίμησης έτσι ώστε να αντιστοιχεί πάντοτε στην ελάχιστη τιμή μεταξύ των **τελικώς** εκτιμηθέντων διαδόχων του κόμβων.

Παράδειγμα 2.5: Η εφαρμογή της διεργασίας άλφα-βήτα στο παίγνιο του Παράδ. 2.3 απλουστεύεται επιπλέον λόγω πλήρους ανάπτυξης του παιγνίου. Οι αντίστοιχοι κανόνες έχουν ως εξής:
- Ένας κόμβος βαθμίδας min μπορεί να αποκοπεί, αν $\alpha = +$ σε προγενέστερό του κόμβο βαθμίδος max.
- Ένας κόμβος βαθμίδας max μπορεί να αποκοπεί, αν $\beta = -$ σε προγενέστερό του κόμβο βαθμίδας min.

Η εφαρμογή των κανόνων αυτών στο παίγνιο επτά σπίρτων επιτρέπει τον προσδιορισμό της βέλτιστης λύσης με ανάπτυξη των διακεκομμένων μόνον συνδέσμων στο Σχ. 2.6, δηλαδή με αισθητά λιγότερο υπολογιστικό φόρτο. Ας σημειωθεί ότι αν η διεργασία ανάπτυξης-εκτίμησης άρχιζε από τα δεξιά (και όχι από τα αριστερά), η μείωση του υπολογιστικού φόρτου θα ήταν πολύ μικρότερη, όπως μπορεί εύκολα να διαπιστωθεί.

■

2.3 Άλλα Συνδυαστικά Προβλήματα

Εκτός των βραχύτατων μονοπατιών και των παιγνίων δύο προσώπων, ο Δυναμικός Προγραμματισμός μπορεί να χρησιμοποιηθεί για την επίλυση και άλλων συνδυαστικών προβλημάτων βελτιστοποίησης. Βασική προϋπόθεση για την εφαρμογή του Δυναμικού Προγραμματισμού είναι η διαμόρφωση του συνδυαστικού προβλήματος ως **προβλήματος αποφάσεων πολλαπλών βαθμίδων** (βλ. αρχή του Κεφαλαίου 2.1). Η διαμόρφωση αυτή ήταν άμεση σε προβλήματα βραχύτατων μονοπατιών οργανωμένων σε βαθμίδες, λόγω της φύσεως του συγκεκριμένου προβλήματος. Αντιθέτως, η διαμόρφωση άλλων συνδυαστικών προβλημάτων σε αντίστοιχα προβλήματα αποφάσεων πολλαπλών βαθμίδων μπορεί να μην είναι προφανής χωρίς κατάλληλη επεξεργασία (ή και να μην είναι δυνατή). Ένα σημαντικό συνδυαστικό πρόβλημα που μπορεί, μετά από κατάλληλη μορφοποίηση, να παρουσιασθεί ως πρόβλημα αποφάσεων πολλαπλών βαθμίδων και να επιλυθεί με εφαρμογή Δυναμικού Προγραμματισμού είναι το **πρόβλημα του περιπλανώμενου πωλητού**.

Στην βασική εκδοχή του προβλήματος του περιπλανώμενου πωλητού έχουμε ένα δίκτυο που αποτελείται από N κόμβους (π.χ. πόλεις) όπου το κόστος μετάβασης (π.χ. η απόσταση ή ο χρόνος διαδρομής) από την πόλη i στην πόλη j είναι d_{ij}. Σημειώνουμε ότι:

- Ισχύει $d_{ii} = 0 \ \forall i$ (γιατί;).
- Είναι δυνατόν να έχουμε $d_{ij} \neq d_{ji}$ (π.χ. λόγω μονοδρόμων ή άλλων αιτίων).
- Αν κάποια πόλη i δεν συνδέεται άμεσα με κάποια πόλη j, τότε $d_{ij} = \infty$.

Ένας πωλητής ξεκινά από την πόλη 1, πρέπει να επισκεφθεί κάθε μια από τις υπόλοιπες $N-1$ πόλεις **ακριβώς μία φορά** και να επιστρέψει στην πόλη 1. Ζητείται η πορεία του πωλητή που ελαχιστοποιεί το συνολικό κόστος διαδρομής του.

Το συγκεκριμένο πρόβλημα βελτιστοποίησης εμφανίζεται για παράδειγμα σε πολλές εταιρείες κατά τον σχεδιασμό διαδρομών που πρέπει να ακολουθήσουν τα οχήματα παράδοσης ή αποκομιδής προϊόντων ή υλικών κλπ.

Η μορφοποίηση του προβλήματος αυτού σε πρόβλημα αποφάσεων πολλαπλών βαθμίδων μπορεί να επιτευχθεί ως εξής. Θεωρούμε ότι ο αριθμός των βαθμίδων K ισούται με τον αριθμό των πόλεων, δηλαδή K=N. Στις βαθμίδες 0 και K υπάρχει ένας μόνον κόμβος απόφασης που αντιστοιχεί στην πόλη αφετηρίας και τελικού προορισμού, δηλαδή στην πόλη 1 (βλ. Σχ. 2.14). Στις βαθμίδες $k = 1,\ldots, K-1$, κάθε κόμβος απόφασης χαρακτηρίζεται από ένα ζεύγος (S, j), όπου j είναι η πόλη προορισμού για τον κόμβο αυτό και S είναι ένα σύνολο ενδιάμεσων πόλεων για την μετάβαση από την πόλη 1 στην πόλη j, άρα $1 \notin S$ και $j \notin S$. Ο αριθμός |S| των ενδιάμεσων πόλεων για κάθε κόμβο της βαθμίδας k ορίζεται ίσος με $k-1$, άρα στην βαθμίδα 1 έχουμε για κάθε κόμβο S=Ø, στην βαθμίδα 2 έχουμε $|S| = 1$ κ.ο.κ., μέχρι την βαθμίδα $K-1$ όπου το σύνολο των ενδιάμεσων πόλεων κάθε κόμβου έχει $N-2$ στοιχεία, δηλαδή $|S| = N-2$ (βλ. Σχ. 2.14 για $N = 5$). Έτσι, ο αριθμός κόμβων της βαθμίδας k καθορίζεται από όλους τους δυνατούς συνδυασμούς ζευγών (S, j), για j=2, …, N, και για όλα τα δυνατά μη διατεταγμένα σύνολα S που ικανοποιούν τις ως άνω προδιαγραφές |S| = k – 1, $1 \notin S,\ j \notin S$.

Σε κάθε κόμβο απόφασης (S, j) μιας βαθμίδας $k = 1,\ldots, K-1$ υπάρχουν τόσες εναλλακτικές αποφάσεις για την μετάβαση στην επόμενη βαθμίδα k + 1, όσος και ο αριθμός των πόλεων που μπορεί να επισκεφθεί ο πωλητής μετά την πόλη j δεδομένου ότι έχει ήδη επισκεφθεί τις πόλεις του συνόλου S (και απαγορεύεται να τις ξαναεπισκεφθεί). Άρα ο αριθμός μεταβάσεων κάθε κόμβου της βαθμίδας k είναι $N-1-|S|-1$ (δηλαδή $N-1$ πόλεις προς μετάβαση μείον αυτές του συνόλου S μείον η πόλη j), και αφού $|S| = k-1$, ο αριθμός μεταβάσεων είναι $N-k-1$, βλ. Σχ. 2.14. Το κόστος μετάβασης από τον κόμβο (S, j) της βαθμίδος k σε έναν κόμβο (S', i) της βαθμίδος k+1 είναι d_{ji}, δηλαδή όσο και το κόστος μετακίνησης από την πόλη j στην πόλη i, και ισχύει $S' = S \cup \{j\}$.

Για παράδειγμα (Σχ. 2.14), αν $N = 5$, υπάρχουν για τον κόμβο (Ø, 2) της βαθμίδας 1 τρεις μεταβάσεις στην βαθμίδα 2, αφού μετά την πόλη 2 (και χωρίς καμία ενδιάμεση πόλη αφού S=Ø) μπορεί ο πωλητής να επισκεφθεί τις πόλεις 3 ή 4 ή 5. Έτσι η μετάβαση από τον κόμβο (Ø, 2) της βαθμίδας 1 στον κόμβο ({2}, 3) της βαθμίδας 2 έχει κόστος ίσο με το κόστος μετακίνησης από την πόλη 2 στην πόλη 3, δηλαδή d_{23}, ενώ τα κόστη των άλλων δύο δυνατών μεταβάσεων, δηλαδή από τον (Ø, 2) προς τους ({2}, 4) ή ({2}, 5) είναι d_{24} ή d_{25}, αντίστοιχα. Στο ίδιο παράδειγμα, ο κόμβος ({2}, 3} της βαθμίδας 2 έχει δύο εναλλακτικές αποφάσεις για

την μετάβαση στην βαθμίδα 3, αφού μετά την πόλη 3 μπορεί ο πωλητής να επισκεφθεί μόνον τις πόλεις 4 ή 5 (την 2 την έχει ήδη επισκεφθεί), δηλαδή τις μεταβάσεις προς τους κόμβους ({2, 3}, 4) ή ({2, 3}, 5) με κόστη d_{34} ή d_{35}, αντίστοιχα. Σημειώνουμε ότι μετάβαση προς τον ως άνω κόμβο ({2, 3}, 4} της βαθμίδας 3 γίνεται, π.χ., και από τον κόμβο ({3}, 2) της βαθμίδας 2 (με κόστος d_{24}) αφού τα σύνολα S δεν είναι διατεταγμένα.

j i	1	2	3	4	5
1	0	3	1	5	4
2	1	0	5	4	3
3	5	4	0	2	1
4	3	1	3	0	3
5	5	2	4	1	0

Πίνακας 2.2: Αποστάσεις d_{ij} από την πόλη i στην πόλη j για ένα δίκτυο 5 πόλεων.

Οι μεταβάσεις από τον μοναδικό κόμβο της βαθμίδας 0 στους $N-1$ κόμβους της βαθμίδος 1 (δηλαδή τους κόμβους (Ø, j), j = 2,…, N) έχουν κόστη d_{1j}, j = 2,…, N. Τέλος, στην βαθμίδα $K-1$ έχουμε ακριβώς $N-1$ κόμβους (S, j), j = 2,…, N, όπου $|S| = N-2$, και κάθε κόμβος έχει μία μετάβαση πρός τον μοναδικό κόμβο της βαθμίδας K (τελικός προορισμός) με κόστος d_{j1}.

Με τον τρόπο αυτό, το πρόβλημα του περιπλανώμενου πωλητού έχει μετασχηματισθεί σε ένα πρόβλημα αποφάσεων πολλαπλών βαθμίδων σε επίπεδο γράφο που μπορεί να επιλυθεί με τον ίδιο τρόπο όπως και το πρόβλημα βραχύτατου μονοπατιού του Παραδ. 2.1 με χρήση Δυναμικού Προγραμματισμού, είτε στην όπισθεν (Σχ. 2.2) είτε στην έμπροσθεν (Σχ. 2.3) εκδοχή του. Αν χρησιμοποιηθεί η έμπροσθεν εκδοχή του Δυναμικού Προγραμματισμού, το βέλτιστο κόστος κάθε κόμβου (S, j) (που μπορεί να αναγράφεται κατά την γραφική επίλυση σε κύκλο πάνω από τον κόμβο όπως στο Παράδ. 2.1) δηλώνει το κόστος του βραχύτατου μονοπατιού από την πόλη 1 στην πόλη j μέσω των πόλεων του συνόλου S. Σημειώνουμε ότι το βραχύτατο αυτό μονοπάτι αναφέρεται στην βέλτιστη διάταξη (μεταξύ όλων των δυνατών διατάξεων) των ενδιάμεσων πόλεων του συνόλου S.

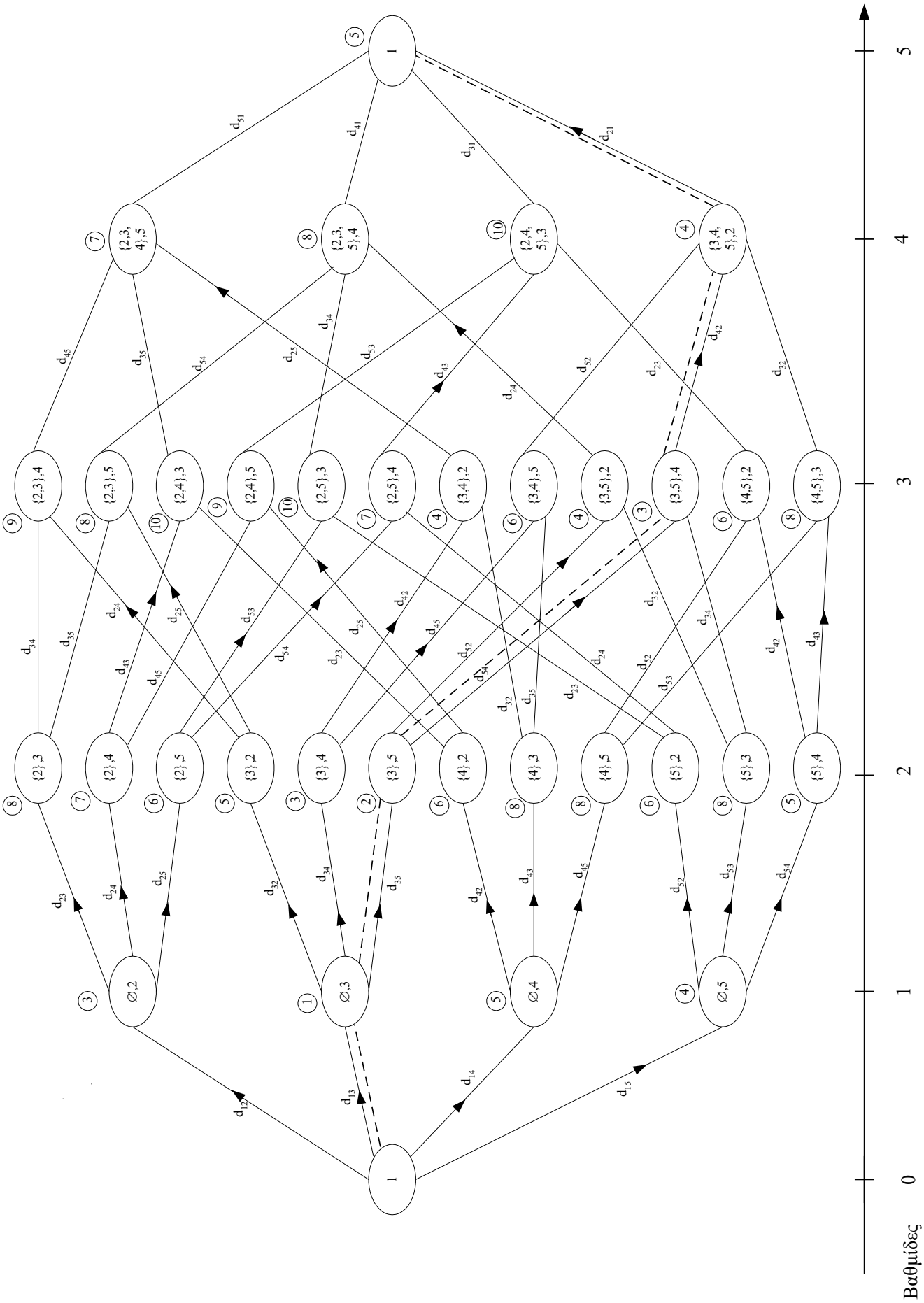
Παράδειγμα 2.6: Θεωρούμε το πρόβλημα του περιπλανώμενου πωλητού για $N=5$ με τις αποστάσεις μεταξύ των 5 πόλεων που αναγράφονται στον Πίνακα 2.2.

Το Σχ. 2.14 απεικονίζει τον επίπεδο γράφο που αντιστοιχεί στο αντίστοιχο πρόβλημα αποφάσεων πολλαπλών βαθμίδων καθώς και την διαδικασία επίλυσης μέσω (έμπροσθεν) Δυναμικού Προγραμματισμού. Οι αριθμοί σε κύκλους πάνω από τους κόμβους υποδηλώνουν τις αντίστοιχες ελάχιστες αποστάσεις ενώ τα βέλη επί μεταβάσεων υποδηλώνουν την βέλτιστη μετάβαση πρός τον αντίστοιχο κόμβο. Έτσι το μήκος του βραχύτατου μονοπατιού του πωλητή είναι 5 και η αντίστοιχη βέλτιστη διαδρομή του είναι $1-3-5-4-2-1$ (βλ. διακεκομμένη γραμμή στον γράφο επίλυσης του Σχ. 2.14).

■

Ο απαιτούμενος αριθμός κόμβων κάθε βαθμίδος k για την επίλυση του προβλήματος του περιπλανώμενου πωλητού είναι

$$(N-1)\binom{N-2}{k-1}$$



Σχήμα 2.14: Έμπροσθεν εφαρμογή Δυναμικού Προγραμματισμού στο Παράδ. 2.6.

Αριθμός πόλεων	5	6	7	8	9	10	15
Αριθμός κόμβων ΔΠ	34	82	194	450	1026	2306	114690
Αριθμός διαδρομών	24	120	720	$5\cdot 10^3$	$4\cdot 10^4$	$3.6\cdot 10^5$	$8.7\cdot 10^{10}$

Πίνακας 2.3: Υπολογιστικός φόρτος για το πρόβλημα του περιπλανώμενου πωλητή.

όπου για α, β ακέραιους ίσχύει

$$\binom{\alpha}{\beta} = \frac{\alpha!}{\beta!(\alpha-\beta)!}.$$

Έτσι έχουμε συνολικό αριθμό κόμβων ίσο προς

$$(N-1)(N-2)!\sum_{k=0}^{N}[(k-1)!(N-k-1)!]^{-1}.$$

Εξάλλου, ο αριθμός των δυνατών εναλλακτικών διαδρομών του περιπλανώμενου πωλητή είναι $(N-1)!$. Η επίλυση του προβλήματος του περιπλανώμενου πωλητή μέσω σύγκρισης όλων των δυνατών διαδρομών θα απαιτούσε υπέρογκο υπολογιστικό φόρτο όπως γίνεται φανερό από τον Πίνακα 2.3.

Υπάρχουν διάφορες παραλλαγές του προβλήματος του περιπλανώμενου πωλητή, όπως για παράδειγμα:

- Ο πωλητής ξεκινά από την πόλη 1, επισκέπτεται όλες τις άλλες πόλεις ακριβώς μια φορά, αλλά δεν επιστρέφει στην πόλη 1.
- Ο πωλητής ξεκινά από την πόλη 1, επισκέπτεται κάθε μία από τις άλλες πόλεις **τουλάχιστον μια φορά** και επιστρέφει στην πόλη 1.

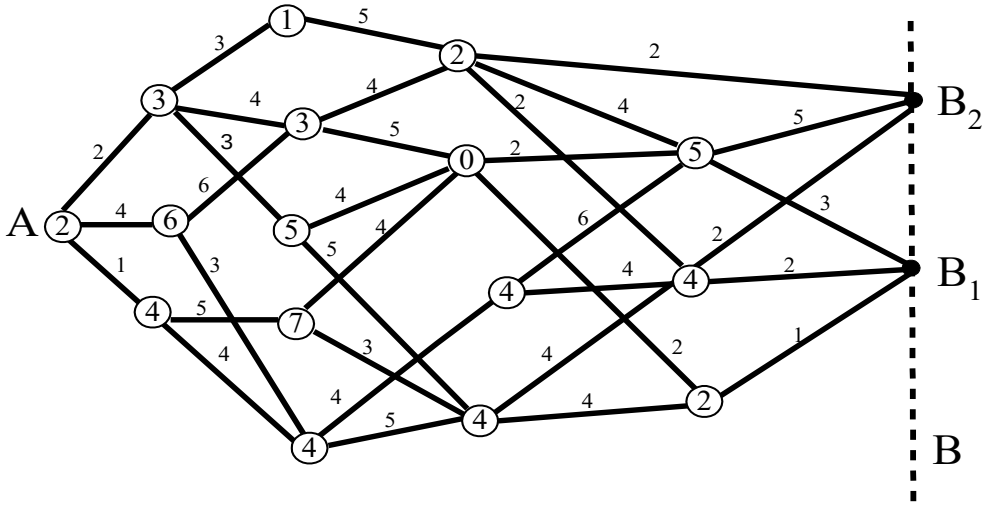
Προβλήματα

2.1 Ένα λεωφορείο με μεταβαλλόμενη διαδρομή διασχίζει το οδικό δίκτυο του παρακάτω σχήματος από το σημείο Α προς Β={B_1, B_2} κινούμενο πάντα προς τα εμπρός. Ο αριθμός των ατόμων που αναμένουν σε κάθε στάση-κόμβο αναγράφεται μέσα στον αντίστοιχο κύκλο, ενώ ο χρόνος διαδρομής μεταξύ δύο στάσεων αναγράφεται πάνω από τον αντίστοιχο σύνδεσμο. Προορισμός όλων των ατόμων είναι το Β. Ο χρόνος παραμονής του λεωφορείου σε κάθε στάση θεωρείται αμελητέος.

(a) Υπολογίστε την βραχύτατη διαδρομή
  (a1) από το Α στο B_1.
  (a2) από το Α στο Β.

(b) Υπολογίστε την διαδρομή που μεγιστοποιεί τον αριθμό των μεταφερομένων ατόμων.

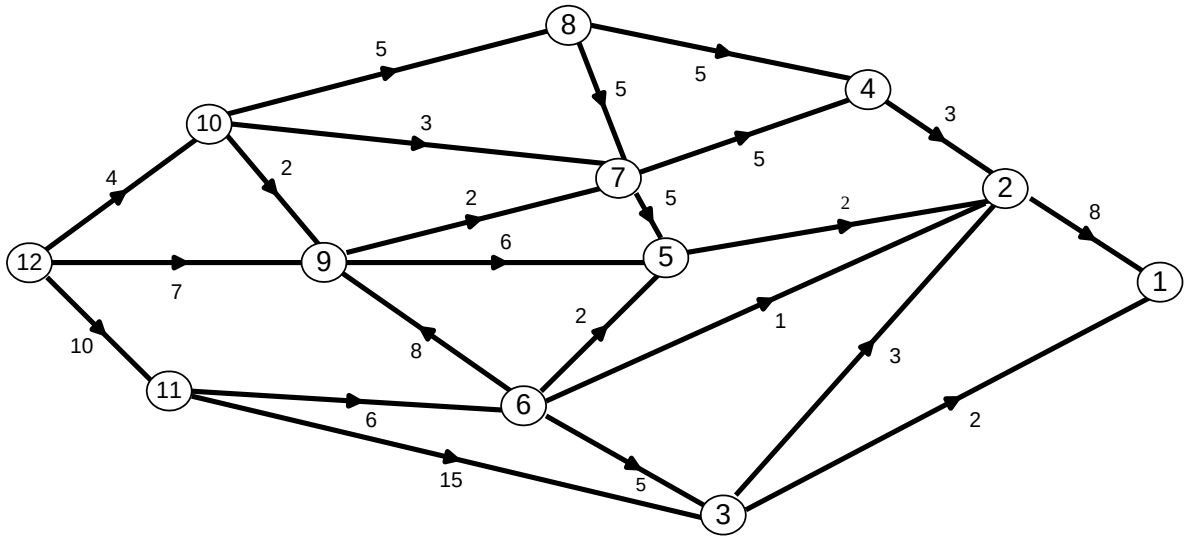
(c) Συγκρίνετε τις παραπάνω διαδρομές σχετικά με τους χρόνους και τον αριθμό μεταφερομένων ατόμων.



2.2 Για το δίκτυο του παρακάτω σχήματος ζητείται η συντομότατη διαδρομή από τον κόμβο 12 στον κόμβο 1.

(a) Μετασχηματίστε το δίκτυο εισάγοντας, όπου χρειάζεται, εικονικούς κόμβους και συνδέσμους, σε σύστημα αποφάσεων πολλαπλών βαθμίδων και επιλύστε το πρόβλημα με την μέθοδο του Δυναμικού Προγραμματισμού.

(b) Υπολογίστε την συντομότατη διαδρομή απ' ευθείας στο απεικονιζόμενο δίκτυο (χωρίς οργάνωση σε βαθμίδες) εφαρμόζοντας την Αρχή του Βέλτιστου.



2.3 Ο παρακάτω πίνακας δείχνει τους χρόνους διαδρομής ενός εργοστασιακού οχήματος μεταξύ αντίστοιχων θέσεων ενός εργοστασιακού χώρου. Οι χρόνοι διαδρομής συμπεριλαμβάνουν τον μέσο όρο παραμονής του οχήματος στις αντίστοιχες θέσεις. Ζητείται η συντομότατη διαδρομή μεταξύ των θέσεων Α και Ε.

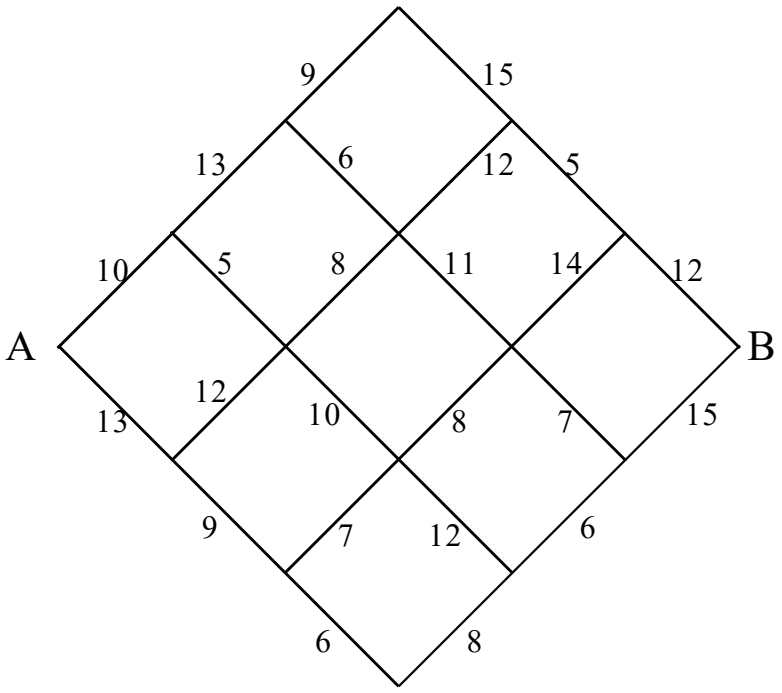
(a) Αναπτύξτε τον γράφο διαδρομών.

(b) Υπολογίστε την συντομότατη διαδρομή και την διάρκειά της.

από / προς	A	B	C	D	E	F	G	H	I	J	K
A	–	2	–	–	–	–	–	6	7	4	–
B	–	–	6	–	–	–	–	–	5	3	–
C	–	–	–	8	–	–	–	–	–	–	3
D	–	–	–	–	5	–	–	–	–	–	–
E	–	–	–	–	–	–	–	–	–	–	–
F	–	–	–	–	6	–	–	–	–	–	–
G	–	–	–	–	–	4	–	–	–	–	2
H	–	–	–	–	–	–	5	–	2	–	5
I	–	–	–	–	–	–	3	–	–	1	5
J	–	–	2	10	–	–	–	–	–	–	6
K	–	–	–	4	10	7	–	–	–	–	–

2.4 Θεωρούμε το παρακάτω δίκτυο οδών του οποίου όλες οι μεταβάσεις νοούνται από αριστερά προς τα δεξιά. Κάθε μονοπάτι από A σε B περιλαμβάνει ακριβώς 6 οδούς.

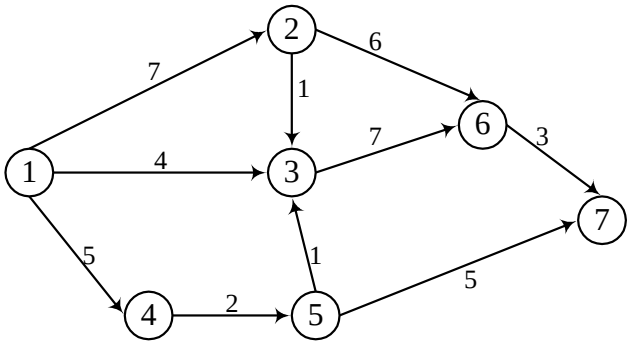
(a) Θεωρούμε ότι τα αναγραφόμενα κόστη μεταβάσεων αντιπροσωπεύουν μήκος διαδρομής. Προσδιορίστε μέσω Δυναμικού Προγραμματισμού το βραχύτατο μονοπάτι από A σε B.



(b) Θεωρούμε τώρα ότι τα αναγραφόμενα κόστη μεταβάσεων αντιπροσωπεύουν το εύρος των αντίστοιχων οδών. Επειδή θέλουμε να αποστείλουμε ένα φορτηγό με υπέρμετρο φάρδος από A σε B, μας ενδιαφέρει το μονοπάτι του οποίου η πιο στενή οδός είναι η ευρύτατη αν συγκριθεί με την πιο στενή οδό άλλων μονοπατιών. Το κριτήριο κόστους

του προβλήματος δεν έχει επομένως την τυπική αθροιστική μορφή άλλων προβλημάτων. Το πρόβλημα μπορεί όμως να επιλυθεί εφαρμόζοντας την γενική λογική του Δυναμικού Προγραμματισμού. Επιλύστε το πρόβλημα και προσδιορίστε το βέλτιστο μονοπάτι και την μέγιστη τιμή του κριτηρίου κόστους.

2.5 Για το δίκτυο του παρακάτω σχήματος ζητείται η συντομότατη διαδρομή από τον κόμβο 1 στον κόμβο 7 θεωρώντας ότι τα αναγραφόμενα κόστη μεταβάσεων αντιπροσωπεύουν τα μήκη των αντίστοιχων διαδρομών.



(a) Υπολογίστε τη συντομότατη διαδρομή απ' ευθείας στο απεικονιζόμενο δίκτυο εφαρμόζοντας την Αρχή του Βέλτιστου.

(b) Ποια είναι η προϋπόθεση που πρέπει να ισχύει έτσι ώστε το παραπάνω πρόβλημα να μπορεί να επιλυθεί με εφαρμογή της μεθόδου του Δυναμικού Προγραμματισμού;

(c) Λαμβάνοντας υπόψη την παραπάνω προϋπόθεση, μετασχηματίστε κατάλληλα το δίκτυο και επιλύστε το πρόβλημα με τη μέθοδο του Δυναμικού Προγραμματισμού.

2.6 Η παραγωγική διαδικασία ενός προϊόντος απαιτεί την εκτέλεση τεσσάρων εργασιών 1, 2, 3, 4 από τις οποίες η εργασία 1 πρέπει να εκτελείται πάντα πρώτη. Το κόστος παραγωγής του προϊόντος εξαρτάται από τη σειρά εκτέλεσης των υπολοίπων εργασιών (2, 3, 4). Στον παρακάτω πίνακα δίνεται το κόστος εκτέλεσης κάθε εργασίας σε συνάρτηση της αμέσως προηγούμενης εργασίας. Για παράδειγμα αν εκτελεστεί η εργασία 2 μετά την εργασία 1 το κόστος είναι 5, αν εκτελεστεί η εργασία 3 μετά την εργασία 2 το κόστος είναι 6, και ούτω καθεξής. Η βιομηχανία ενδιαφέρεται να προσδιορίσει τη σειρά με την οποία θα πραγματοποιηθούν οι εργασίες έτσι ώστε να έχει το ελάχιστο δυνατό κόστος.

	1	2	3	4
1	0	5	4	7
2		0	6	4
3		2	0	4
4		3	5	0

(a) Σε ποιό χαρακτηριστικό συνδυαστικό πρόβλημα μπορεί να αναχθεί το συγκεκριμένο πρόβλημα και γιατί;

(b) Λαμβάνοντας υπόψη την απάντηση στο ερώτημα (a) επιλύστε γραφικά το πρόβλημα με έμπροσθεν δυναμικό προγραμματισμό.

2.7 Θεωρούμε το παιδικό παίγνιο δύο προσώπων τα οποία στέκονται κατ' αρχήν αντικριστά σε απόσταση Α. Οι παίκτες προσεγγίζουν εναλλάξ ο ένας τον άλλον τοποθετώντας το εκάστοτε πίσω πέλμα τους ακριβώς μπροστά από το άλλο. Κατά την προσέγγισή τους οι παίκτες έχουν την επιλογή να τοποθετήσουν το πέλμα τους είτε κατά μήκος (οπότε προσεγγίζουν τον αντίπαλό τους κατά Δ), είτε κάθετα (οπότε προσεγγίζουν τον αντίπαλό τους κατά δ, όπου δ < Δ). Θεωρούμε ότι τα μεγέθη δ, Δ είναι ίδια για τους δύο παίκτες. Κερδίζει εκείνος που θα πατήσει (ή θα αγγίξει) πρώτος το πέλμα του αντιπάλου.

(a) Θεωρώντας ότι η τρέχουσα κατάσταση του παιγνίου χαρακτηρίζεται από την τρέχουσα απόσταση μεταξύ των δύο παικτών, αναπτύξτε το παίγνιο από την κατάσταση εκκίνησης (απόσταση Α) μέχρι βάθος 4 (δύο κινήσεις για κάθε αντίπαλο). Υπόδειξη: Κάθε κατάσταση του παιγνίου αρκεί να εμφανίζεται μία μόνο φορά σε κάθε βαθμίδα.

(b) Προσδιορίστε την βέλτιστη στρατηγική κάθε παίκτου για Α = 21, Δ = 5, δ = 2. Ποιός παίκτης είναι "καταδικασμένος" να κερδίσει και ποιός να χάσει εφ' όσον παίζουν βέλτιστα αμφότεροι;

2.8 Θεωρούμε ένα παίγνιο δύο προσώπων το οποίο έχει ως εξής: Το παίγνιο εκκινά με ένα σκορ ίσο με 0. Οι δύο παίκτες, παίζοντας εναλλάξ, έχουν την επιλογή να προσθέτουν στο σκορ αυτό 1 ή 2 κάθε φορά. Μόλις το σκορ φτάσει στο 11, το παίγνιο τερματίζει και νικητής είναι ο παίκτης ο οποίος έφτασε το σκορ στο 11.

(a) Θεωρώντας ότι η τρέχουσα κατάσταση του παιγνίου χαρακτηρίζεται από την τρέχουσα τιμή του σκορ, αναπτύξτε πλήρως το παίγνιο από την κατάσταση εκκίνησης (σκορ 0) μέχρι την κατάσταση τερματισμού (σκορ 11).

(b) Προσδιορίστε τη βέλτιστη στρατηγική για κάθε παίκτη. Ποιός παίκτης είναι "καταδικασμένος" να κερδίσει και ποιός να χάσει εφ' όσον παίζουν βέλτιστα αμφότεροι;

(c) Θέλουμε να επιλύσουμε το ίδιο πρόβλημα χρησιμοποιώντας τη μέθοδο άλφα–βήτα. Υποδείξτε από ποιά πλευρά του δέντρου που αναπτύξατε μας συμφέρει να ξεκινήσουμε την εφαρμογή της μεθόδου και γιατί.

2.9 Θεωρούμε ένα παίγνιο με επτά σπίρτα που έχει ως εξής: Επτά σπίρτα τοποθετούνται πάνω στο τραπέζι (αρχικός σωρός σπίρτων) και δύο παίκτες Α και Β πρέπει, παίζοντας εναλλάξ, να χωρίσουν κάθε φορά ένα σωρό από σπίρτα του τραπεζιού σε δύο σωρούς που έχουν διαφορετικό αριθμό σπίρτων. Για παράδειγμα ένας σωρός με 6 σπίρτα μπορεί να χωριστεί σε δύο σωρούς των 4 και 2 ή 5 και 1, αλλά όχι 3 και 3. Ο πρώτος παίκτης που δεν μπορεί να κάνει νέα κίνηση χάνει το παιχνίδι.

(a) Αναπτύξτε πλήρως το παίγνιο από την σκοπιά του παίκτη Α. Η τρέχουσα κατάσταση του παιγνίου χαρακτηρίζεται από ένα σύνολο που εμπεριέχει τον αριθμό σπίρτων κάθε σωρού, π.χ. {7}, {4, 3}, {2, 2, 2, 1} κ.λπ. **Υπόδειξη**: Κάθε κατάσταση του παιγνίου αρκεί να εμφανίζεται μία μόνο φορά σε κάθε βαθμίδα.

(b) Προσδιορίστε τη βέλτιστη στρατηγική για κάθε παίκτη. Ποιος παίκτης είναι «καταδικασμένος» να κερδίσει και ποιός να χάσει εφ' όσον παίζουν βέλτιστα αμφότεροι;

(c) Επιλύστε το ίδιο πρόβλημα χρησιμοποιώντας την διαδικασία άλφα-βήτα.

3. ΒΕΛΤΙΣΤΟΣ ΕΛΕΓΧΟΣ ΔΙΑΚΡΙΤΟΥ ΧΡΟΝΟΥ

Το κεφάλαιο αυτό αναφέρεται στην εφαρμογή της Αρχής του Βέλτιστου στο πρόβλημα βέλτιστου ελέγχου διακριτού χρόνου (βλ. Παράρτ. Β.1).

3.1 Διατύπωση του Προβλήματος

Θεωρούμε την ελαχιστοποίηση του κριτηρίου κόστους διακριτού χρόνου

$$J = \vartheta\left[\mathbf{x}(K)\right] + \sum_{k=0}^{K-1} \varphi\left[\mathbf{x}(k), \mathbf{u}(k), k\right] \tag{3.1}$$

με καθορισμένο χρονικό ορίζοντα K, λαμβάνοντας υπόψη τους καταστατικούς περιορισμούς

$$\mathbf{x}(k+1) = \mathbf{f}\left[\mathbf{x}(k), \mathbf{u}(k), k\right] \tag{3.2}$$

με αρχική κατάσταση $\mathbf{x}(0) = \mathbf{x}_0$ και τελικό στόχο που προσδιορίζεται μέσω

$$\mathbf{g}\left[\mathbf{x}(K)\right] = \mathbf{0} \tag{3.3}$$

ή με οποιοδήποτε άλλο τρόπο προσδιορισμού ενός συνόλου σημείων $\mathbf{x}(K)$ στα οποία και μόνον επιτρέπεται να καταλήξει η κατάσταση $\mathbf{x}$ του προβλήματος. Εννοείται ότι όλα τα αποτελέσματα του κεφαλαίου αυτού είναι εφαρμόσιμα στην ειδική περίπτωση απουσίας ενός τελικού στόχου. Σε σύγκριση με το Παραρτ. Β.1 περιοριζόμαστε εδώ, χάριν ευκολίας, στην περίπτωση δεδομένου τελικού χρόνου K. Η γενίκευση των αποτελεσμάτων σε περίπτωση ελεύθερου τελικού χρόνου θα εξετασθεί στο Κεφάλαιο 3.4 β).

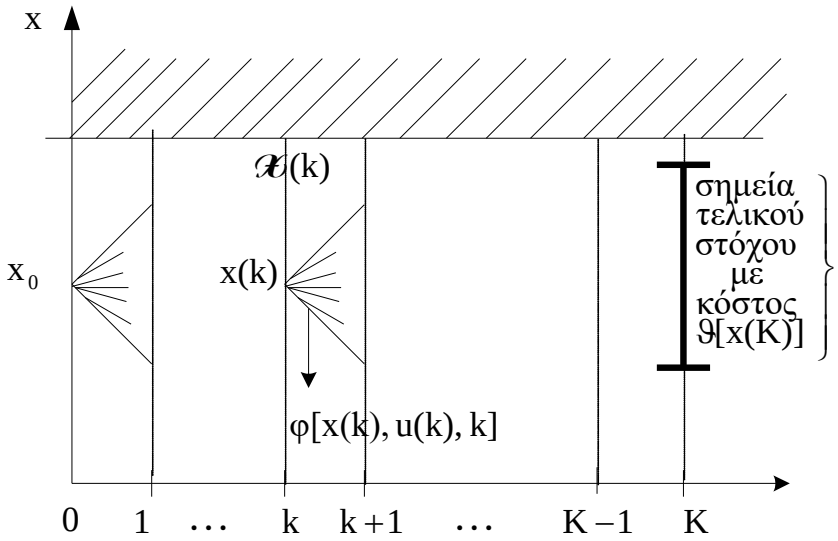
Η επιτρεπτή περιοχή ελέγχου καθορίζεται ως εξής

$$\mathbf{u}(k) \in \mathscr{U}[\mathbf{x}(k), k] = \{\mathbf{u}(k) \,\big|\, \mathbf{h}[\mathbf{x}(k), \mathbf{u}(k), k] \le \mathbf{0}\} \tag{3.4}$$

όπου υποθέτουμε ότι ο Ιακωβιανός πίνακας $\partial\mathbf{h}/\partial\mathbf{u}$ είναι πλήρους βαθμού. Η παραδοχή αυτή δεν εκπληρούται αν έχουμε περιορισμούς ανισότητας κατάστασης. Στην περίπτωση αυτή θεωρούμε επιπλέον τους ακόλουθους περιορισμούς ανισότητας

$$\mathbf{x}(k) \in \mathscr{X}(k) = \{\mathbf{x}(k) \,\big|\, \mathbf{h}^{x}[\mathbf{x}(k), k] \le \mathbf{0}\} \tag{3.5}$$

όπου $\mathscr{X}(k)$ είναι η επιτρεπτή περιοχή κατάστασης. Επί πλέον, όπου οι συγκεκριμένες εφαρμογές το απαιτούν, μπορούμε να θεωρήσουμε και περιορισμούς ισότητας (Β.22), διακριτά σημεία ελέγχου (Β.23) ή κατάστασης και άλλους περιορισμούς.



Σχήμα 3.1: Παραστατική επεξήγηση του ΠΒΕ διακριτού χρόνου.

Για διακριτούς χρόνους k=0,1,...,K μπορούμε να θεωρήσουμε το πρόβλημα αυτό σαν πρόβλημα αποφάσεων πολλαπλών βαθμίδων και να επιχειρήσουμε τη λύση του μέσω της Αρχής του Βέλτιστου, όπως κάναμε για το Παραδ. 2.1. Προς το σκοπό αυτό προϋποθέτουμε την ύπαρξη ενός απόλυτου ελαχίστου του προβλήματος.

Στο σημείο αυτό είναι χρήσιμο να σκιαγραφηθούν η αντιστοιχία, αλλά και οι διαφορές του διατυπωθέντος ΠΒΕ διακριτού χρόνου με το πρόβλημα του Παραδ. 2.1 για την περίπτωση μονοδιάστατης κατάστασης x και ελέγχου u. Είναι καταρχήν προφανές (Σχ. 3.1) ότι ο δείκτης k διακριτού χρόνου αντιστοιχεί στις βαθμίδες του Παραδ. 2.1, ενώ η επιτρεπτή περιοχή κατάστασης $\mathcal{X}$(k) προσδιορίζει τις επιτρεπτές τιμές $x(k) \in \mathcal{X}(k)$ που αντιστοιχούν στους κόμβους του Παραδ. 2.1. Σε αντίθεση όμως με το Παράδ. 2.1, έχουμε εδώ εν γένει άπειρους κόμβους x(k) ανά βαθμίδα k και ενδεχομένως άπειρα σημεία x(K) τελικού στόχου, το κόστος των οποίων σύμφωνα με το αντικειμενικό κριτήριο (3.1) προσδιορίζεται από την συνάρτηση ϑ[x(K)]. Θεωρώντας ένα τυχαίο επιτρεπτό σημείο x(k), κάθε επιτρεπτός έλεγχος $u(k) \in \mathcal{U}(k)$ οδηγεί, μέσω της καταστατικής εξίσωσης (3.2), σε αντίστοιχο σημείο x(k+1). Λόγω των περιορισμών (3.4) και (3.5) δεν είναι επιτρεπτή κάθε μετάβαση στην επομένη βαθμίδα, όμως ο αριθμός των επιτρεπτών μεταβάσεων είναι εν γένει (σε αντίθεση με το Παράδ. 2.1) άπειρος. Το κόστος μετάβασης από ένα σημείο x(k) στην επόμενη βαθμίδα k+1 μέσω ενός ελέγχου u(k) προσδιορίζεται από την (3.1) σε φ[x(k), u(k), k]. Επομένως το αντικειμενικό κριτήριο (3.1) αντανακλά το συνολικό κόστος της διαδρομής (τροχιάς) από την αρχική κατάσταση x_0 έως κάποιο σημείο x(K) του τελικού στόχου.

3.2 Ο Αναδρομικός Τύπος του Bellman

Για την εφαρμογή της Αρχής του Βέλτιστου στο ΠΒΕ διακριτού χρόνου, ορίζουμε καταρχήν το **υπολειπόμενο κόστος** ή **κόστος μεταφοράς** J_k για τη μεταφορά μιας κατάστασης **x**(k) στον τελικό στόχο (3.3) ως εξής (προσοχή στο κάτω όριο του αθροίσματος)

$$J_k = \vartheta[\mathbf{x}(K)] + \sum_{\kappa=k}^{K-1} \varphi[\mathbf{x}(\kappa), \mathbf{u}(\kappa), \kappa]. \qquad (3.6)$$

Για ένα συγκεκριμένο πρόβλημα, το ελάχιστο κόστος μεταφοράς $J_k^* = \min J_k$ (ικανοποιώντας και όλους τους περιορισμούς) εξαρτάται αποκλειστικά από την υπό μεταφορά κατάσταση $\mathbf{x}(k)$ και από το χρόνο k. Ονομάζουμε αυτό το ελάχιστο κόστος $V[\mathbf{x}(k), k]$ και έχουμε

$$V[\mathbf{x}(k),k] = \min J_k = \min\{\varphi[\mathbf{x}(k), \mathbf{u}(k), k] + J_{k+1}\} \qquad (3.7)$$

όπου το ελάχιστο νοείται μεταξύ όλων των τροχιών $\mathbf{u}(\kappa)$, $\kappa = k,...,K-1$, που ικανοποιούν τις (3.2)-(3.5). Εφαρμόζοντας την Αρχή του Βέλτιστου στην (3.7) έχουμε

$$V[\mathbf{x}(k), k] = \min\{\varphi[\mathbf{x}(k), \mathbf{u}(k),k] + V[\mathbf{x}(k+1), k+1]\}. \qquad (3.8)$$

Σημειώνουμε ότι το ελάχιστο κόστος $V[\mathbf{x}(k), k]$ αντιστοιχεί στο βέλτιστο κόστος που αναγράφονταν σε κύκλο πάνω από κάθε κόμβο του Παραδ. 2.1, οπότε η αντιστοιχία της (3.8) με την διαδικασία ελαχιστοποίησης σε κάθε κόμβο του Παραδ. 2.1 καθίσταται προφανής. Αντικαθιστώντας την $\mathbf{x}(k+1)$ από την (3.2) στην (3.8) έχουμε

$$V[\mathbf{x}(k), k] = \min\{\varphi[\mathbf{x}(k), \mathbf{u}(k), k] + V[\mathbf{f}(\mathbf{x}(k), \mathbf{u}(k), k), k+1]\}. \qquad (3.9)$$

Η δεξιά πλευρά της ισότητας αυτής, που φέρει το όνομα **αναδρομικός τύπος του Bellman**, εξαρτάται από τον $\mathbf{u}(k)$, όχι όμως από τους μεταγενέστερους $\mathbf{u}(\kappa)$, $\kappa = k+1,...,K-1$. Επομένως, η ελαχιστοποίηση στον αναδρομικό τύπο του Bellman νοείται μόνο για τις μεταβλητές ελέγχου του χρόνου k, δηλαδή για $\mathbf{u}(k)$, πάντοτε βέβαια υπό τους περιορισμούς (3.4), (3.5). Αυτή η μονοβαθμιαία ελαχιστοποίηση μπορεί να πραγματοποιηθεί ανεξάρτητα για κάθε βαθμίδα, αρχίζοντας από τον τελικό χρόνο, δηλαδή για $k=K-1, K-2,...,0$, οπότε το αρχικό πρόβλημα αποφάσεων πολλαπλών βαθμίδων κατανέμεται σε K προβλήματα αποφάσεων μιας βαθμίδος. Ας παρουσιάσουμε λεπτομερώς αυτή την βαθμιαία διαδικασία (**Δυναμικός Προγραμματισμός**):

Βαθμίδα K–1: Για κάθε $\mathbf{x}(K-1) \in \mathscr{X}(K-1)$ καθορίζεται ο αντίστοιχος $\mathbf{u}(K-1)$ που ελαχιστοποιεί

$$J_{K-1} = \vartheta[\mathbf{x}(K)] + \varphi[\mathbf{x}(K-1), \mathbf{u}(K-1), K-1]$$

λαμβάνοντας υπόψη

$$\begin{aligned} &\mathbf{x}(K) = \mathbf{f}[\mathbf{x}(K-1), \mathbf{u}(K-1), K-1] \\ &\mathbf{g}[\mathbf{x}(K)] = \mathbf{0} \\ &\mathbf{u}(K-1) \in \mathscr{U}[\mathbf{x}(K-1), K-1] \\ &\mathbf{x}(K) \in \mathscr{X}(K). \end{aligned}$$

Το αποτέλεσμα αυτής της μονοβαθμιαίας ελαχιστοποίησης για όλες τις $\mathbf{x}(K-1) \in \mathscr{X}(K-1)$ εκφράζεται μέσω $\mathbf{u}(K-1) = \mathbf{R}[\mathbf{x}(K-1), K-1]$. Οι αντίστοιχες ελάχιστες τιμές του J_{K-1} συμβολίζονται, σύμφωνα με ένα προηγούμενο ορισμό, με $V[\mathbf{x}(K-1), K-1]$.

Βαθμίδα K–2: Για κάθε $\mathbf{x}(K-2) \in \mathscr{X}(K-2)$ καθορίζεται ο αντίστοιχος $\mathbf{u}(K-2)$ που ελαχιστοποιεί

$$\tilde{J}_{K-2} = V[\mathbf{x}(K-1), K-1] + \varphi[\mathbf{x}(K-2), \mathbf{u}(K-2), K-2]$$

λαμβάνοντας υπόψη

$$\mathbf{x}(K-1) = \mathbf{f}[\mathbf{x}(K-2), \mathbf{u}(K-2), K-2]$$
$$\mathbf{u}(K-2) \in \mathcal{U}[\mathbf{x}(K-2), K-2]$$
$$\mathbf{x}(K-1) \in \mathcal{X}(K-1).$$

(Ο τελικός στόχος (3.3) δεν χρειάζεται πλέον να ληφθεί υπόψη.) Τα αποτελέσματα αυτής της μονοβαθμιαίας βελτιστοποίησης εκφράζονται με τις συναρτήσεις $\mathbf{u}(K-2) = \mathbf{R}[\mathbf{x}(K-2), K-2]$ και $V[\mathbf{x}(K-2), K-2]$.

-
-
-

Βαθμίδα 0: Για $\mathbf{x}(0) = \mathbf{x}_0$ καθορίζεται ο αντίστοιχος $\mathbf{u}(0)$ που ελαχιστοποιεί

$$\tilde{J}_0 = V[\mathbf{x}(1), 1] + \varphi[\mathbf{x}(0), \mathbf{u}(0), 0]$$

λαμβάνοντας υπόψη

$$\mathbf{x}(1) = \mathbf{f}[\mathbf{x}(0), \mathbf{u}(0), 0]$$
$$\mathbf{u}(0) \in \mathcal{U}[\mathbf{x}(0), 0]$$
$$\mathbf{x}(1) \in \mathcal{X}(1).$$

Τα αποτελέσματα αυτής της τελευταίας μονοβαθμιαίας ελαχιστοποίησης εκφράζονται με τις συναρτήσεις $\mathbf{u}(0) = \mathbf{R}[\mathbf{x}(0), 0]$ και $V[\mathbf{x}(0), 0]$.

Σε μερικές βαθμίδες μπορεί να υπάρχουν καταστάσεις $\mathbf{x}(k) \in \mathcal{X}(k)$ για τις οποίες το αντίστοιχο πρόβλημα μονοβαθμιαίας ελαχιστοποίησης δεν έχει λύση επειδή η επιτρεπτή περιοχή που απορρέει από τους περιορισμούς είναι κενή. Τέτοια σημεία $\mathbf{x}(k)$ προφανώς δεν μπορούν να μεταφερθούν στον τελικό στόχο ικανοποιώντας όλους τους περιορισμούς. Για τη διαδικασία του Δυναμικού Προγραμματισμού, αυτά τα σημεία μπορούν να θεωρηθούν ότι έχουν βέλτιστο κόστος μεταφοράς $V[\mathbf{x}(k), k] = \infty$.

Στο τέλος της περιγραφείσας διαδικασίας K βαθμίδων του Δυναμικού Προγραμματισμού έχουν υπολογισθεί, όπως και στο Παραδ. 2.1, όχι απλώς οι βέλτιστες τροχιές για την μεταφορά της αρχικής κατάστασης $\mathbf{x}(0) = \mathbf{x}_0$ στον τελικό στόχο (3.3), αλλά ένας **βέλτιστος κανόνας κλειστού βρόχου** που εκφράζεται από τα αποτελέσματα των προβλημάτων μονοβαθμιαίας ελαχιστοποίησης

$$\mathbf{u}(k) = \mathbf{R}[\mathbf{x}(k), k], \; k = 0, \ldots, K-1. \qquad (3.10)$$

Για να διευρύνουμε το πεδίο εφαρμογής του βέλτιστου κανόνα ελέγχου, μπορούμε να επιλύσουμε το παραπάνω πρόβλημα της βαθμίδας 0 **για κάθε** $\mathbf{x}(0) \in \mathcal{X}(0)$. Η (3.10) περιλαμβάνει τότε επαρκή στοιχεία για την βέλτιστη μεταφορά όχι μόνο του $\mathbf{x}_0$ αλλά οποιουδήποτε σημείου $\mathbf{x}(k) \in \mathcal{X}(k)$, $k=0,\ldots,K-1$, στον τελικό στόχο (3.3), υπό την προϋπόθεση ότι η μεταφορά του είναι εφικτή.

Η μονοβαθμιαία ελαχιστοποίηση σε κάθε βαθμίδα μπορεί να επιχειρηθεί αναλυτικά ή αριθμητικά (με την βοήθεια υπολογιστή). Αναλυτικές λύσεις είναι όμως εφαρμόσιμες μόνο για σχετικά απλά προβλήματα, όπως αυτό του επόμενου παραδείγματος. Η αριθμητική επεξεργασία της μονοβαθμιαίας ελαχιστοποίησης θα περιγραφεί στο επόμενο Κεφάλαιο 3.3. Τέλος, θα πρέπει να τονισθεί ότι ο Δυναμικός Προγραμματισμός καταλήγει σε ένα **απόλυτο ελάχιστο** του ΠΒΕ διακριτού χρόνου, εφ' όσον τα υπολογισθέντα μονοβαθμιαία ελάχιστα είναι και αυτά απόλυτα.

Παράδειγμα 3.1a: Θεωρούμε το ΠΒΕ διακριτού χρόνου που αφορά την ελαχιστοποίηση του

$$J = \frac{1}{2}\sum_{k=0}^{3}\left[x(k)^2 + u(k)^2\right]$$

λαμβάνοντας υπόψη τους περιορισμούς

$$x(k+1) = x(k) + u(k), \quad x(0) = 1, \quad x(4) = 0. \qquad (3.11)$$

Για την επίλυση του προβλήματος εφαρμόζουμε τον Δυναμικό Προγραμματισμό ως εξής:

Βαθμίδα 3: Για κάθε $x(3)\in\mathcal{R}$ πρέπει να υπολογισθεί ο αντίστοιχος u(3) που ελαχιστοποιεί

$$J_3 = \frac{1}{2}\left[x(3)^2 + u(3)^2\right]$$

λαμβάνοντας υπόψη x(4) = x(3) + u(3) και x(4) = 0. Διαπιστώνουμε αμέσως ότι υπάρχει μια μοναδική επιτρεπτή λύση που οδηγεί στον τελικό στόχο, δηλαδή

$$u(3) = -x(3). \qquad (3.13a)$$

Με τη λύση αυτή έχουμε

$$V[x(3), 3] = \min J_3 = x(3)^2.$$

Βαθμίδα 2: Για κάθε $x(2)\in\mathcal{R}$ πρέπει να υπολογισθεί ο αντίστοιχος u(2) που ελαχιστοποιεί

$$\tilde{J}_2 = x(3)^2 + \frac{1}{2}\left[x(2)^2 + u(2)^2\right]$$

λαμβάνοντας υπόψη x(3) = x(2)+u(2). Δι' αντικαταστάσεως έχουμε

$$\tilde{J}_2 = \left[x(2) + u(2)\right]^2 + \frac{1}{2}\left[x(2)^2 + u(2)^2\right]$$

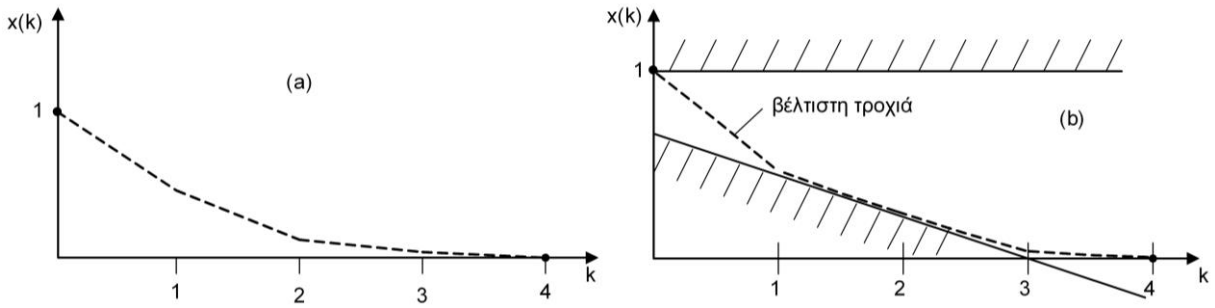
και εξ αυτής το ελάχιστο

$$u(2) = -\frac{2}{3}x(2) \qquad (3.14a)$$

$$V\left[x(2), 2\right] = \min J_2 = \frac{5}{6}x(2)^2.$$

Βαθμίδα 1: Για κάθε $x(1)\in\mathcal{R}$ πρέπει να υπολογισθεί ο αντίστοιχος u(1) που ελαχιστοποιεί

$$\tilde{J}_1 = \frac{5}{6}x(2)^2 + \frac{1}{2}\left[x(1)^2 + u(1)^2\right]$$



Σχήμα 3.2: Επιτρεπτή περιοχή και βέλτιστη τροχιά κατάστασης για τα Παραδ. 3.1a και b.

λαμβάνοντας υπόψη x(2)=x(1)+u(1). Δι' αντικαταστάσεως έχουμε

$$\tilde{J}_1 = \frac{4}{3}\left[x(1)^2 + u(1)^2\right] + \frac{5}{3}x(1)\cdot u(1)$$

και εξ αυτής το ελάχιστο

$$u(1) = -\frac{5}{8}x(1) \qquad (3.15a)$$

$$V[x(1),1] = \frac{13}{16}x(1)^2 \ .$$

Βαθμίδα 0: Για κάθε $x(0)\in\mathcal{R}$ πρέπει να υπολογισθεί ο u(0) που ελαχιστοποιεί

$$\tilde{J}_0 = \frac{13}{16}x(1)^2 + \frac{1}{2}\left[x(0)^2 + u(0)^2\right]$$

λαμβάνοντας υπόψη x(1) = x(0)+u(0). Δι' αντικαταστάσεως έχουμε

$$\tilde{J}_0 = \frac{21}{16}\left[x(0)^2 + u(0)^2\right] + \frac{13}{8}x(0)\cdot u(0)$$

και εξ αυτής το ελάχιστο

$$u(0) = -\frac{13}{21}x(0) \qquad (3.16a)$$

$$V[x(0), 0] = \frac{17}{21}x(0)^2.$$

Οι ζητούμενες βέλτιστες τροχιές για την μεταφορά της x(0) = 1 στην x(4) = 0 είναι λοιπόν (βλ. Σχ. 3.2a)

$$u^*(0) = -\frac{13}{21},\quad u^*(1) = -\frac{5}{21},\quad u^*(2) = -\frac{2}{21},\quad u^*(3) = -\frac{1}{21}$$

$$x^*(1) = \frac{8}{21},\ x^*(2) = \frac{3}{21},\ x^*(3) = \frac{1}{21},\ x^*(4) = 0.$$

Πέραν τούτου έχουμε με τις (3.13a)-(3.16a) έναν βέλτιστο κανόνα ελέγχου που μεταφέρει κάθε κατάσταση $x(k)\in\mathcal{R}$, k=0, 1, 2, 3, κατά βέλτιστο τρόπο στον τελικό στόχο x(4)=0.

Παράδειγμα 3.1b: Θεωρούμε το ΠΒΕ του Παραδ. 3.1a, τώρα όμως με την παρουσία του περιορισμού ανισότητας

$$0.6-0.2k \le x(k) \le 1. \qquad (3.12)$$

Η επιτρεπτή περιοχή κατάστασης (3.12) του προβλήματος αυτού απεικονίζεται στο Σχ. 3.2b. Για την επίλυση του προβλήματος εφαρμόζουμε τον Δυναμικό Προγραμματισμό, η επίλυση των μονοβαθμιαίων προβλημάτων βελτιστοποίησης είναι όμως τώρα πιο δύσκολη εξ αιτίας του περιορισμού (3.12):

Βαθμίδα 3: Για κάθε x(3), $0 \le x(3) \le 1$, πρέπει να υπολογισθεί ο αντίστοιχος u(3) που ελαχιστοποιεί

$$J_3 = \frac{1}{2}\left[x(3)^2 + u(3)^2\right]$$

λαμβάνοντας υπόψη x(4) = x(3) + u(3), x(4) = 0 και $-0.2 \le x(4) \le 1$. Διαπιστώνουμε αμέσως (όπως και στο Παράδ. 3.1a) ότι υπάρχει μία μοναδική επιτρεπτή λύση που οδηγεί στον τελικό στόχο

$$u(3) = -x(3), \quad 0 \le x(3) \le 1. \qquad (3.13b)$$

Με τη λύση αυτή έχουμε

$$V[x(3), 3] = \min J_3 = x(3)^2, \quad 0 \le x(3) \le 1.$$

Βαθμίδα 2: Για κάθε x(2), $0.2 \le x(2) \le 1$, πρέπει να υπολογισθεί ο αντίστοιχος u(2) που ελαχιστοποιεί

$$\tilde{J}_2 = x(3)^2 + \frac{1}{2}\left[x(2)^2 + u(2)^2\right]$$

λαμβάνοντας υπόψη x(3) = x(2)+u(2) και $0 \le x(3) \le 1$. Δι' αντικαταστάσεως έχουμε

$$\tilde{J}_2 = \left[x(2) + u(2)\right]^2 + \frac{1}{2}\left[x(2)^2 + u(2)^2\right] = \frac{3}{2}\left[x(2)^2 + u(2)^2\right] + 2x(2)u(2).$$

Το απεριόριστο ελάχιστο της συνάρτησης αυτής είναι (όπως και στο Παράδ. 3.1a)

$$u(2) = -\frac{2}{3}x(2), \qquad 0.2 \le x(2) \le 1 \qquad (3.14b)$$

το οποίο δεν παραβιάζει τους παραπάνω περιορισμούς του x(3). Το αντίστοιχο ελάχιστο κόστος μεταφοράς είναι

$$V\left[x(2), 2\right] = \min J_2 = \frac{5}{6}x(2)^2, \quad 0.2 \le x(2) \le 1.$$

Βαθμίδα 1: Για κάθε x(1), $0.4 \le x(1) \le 1$, πρέπει να υπολογισθεί ο αντίστοιχος u(1) που ελαχιστοποιεί

$$\tilde{J}_1 = \frac{5}{6}x(2)^2 + \frac{1}{2}\left[x(1)^2 + u(1)^2\right]$$

λαμβάνοντας υπόψη x(2)=x(1)+u(1) και $0.2 \le x(2) \le 1$. Δι' αντικαταστάσεως του περιορισμού ισότητας έχουμε

$$\underset{u(1)}{\text{Min}}\left\{\frac{4}{3}\left[x(1)^2+u(1)^2\right]+\frac{5}{3}x(1)u(1)\right\}$$

λαμβάνοντας υπόψη

$$0.2-x(1)\le u(1)\le 1-x(1).$$

Η επίλυση αυτού του προβλήματος μη γραμμικού προγραμματισμού προσδιορίζεται ως εξής: Η ελαχιστοποίηση του ως άνω κυρτού αντικειμενικού κριτηρίου ως προς u(1) (μέσω μηδενισμού της αντίστοιχης πρώτης παραγώγου) δίδει καταρχήν $u(1)=(-5/8)\cdot x(1)$ (όπως και στο Παράδ. 3.1a). Η λύση αυτή είναι βέλτιστη εφόσον ικανοποιούνται οι υφιστάμενοι περιορισμοί ανισότητας. Ο άνω περιορισμός $u(1)\le 1-x(1)$ απαιτεί, μετά από αντικατάσταση του απεριόριστου βέλτιστου u(1), $x(1)\le 8/3$, και άρα ο περιορισμός αυτός δεν παραβιάζεται, αφού η επιτρεπτή περιοχή του $x(1)$ επιβάλει $x(1)\le 1$. Ο κάτω περιορισμός $u(1)\ge 0.2-x(1)$ απαιτεί, μετά από αντικατάσταση του απεριόριστου βέλτιστου u(1), $x(1)\ge 8/15$, και άρα ο περιορισμός αυτός παραβιάζεται από το απεριόριστο βέλτιστο u(1), αφού, βάσει της επιτρεπτής του περιοχής, το x(1) μπορεί να πάρει τιμές $x(1)\ge 0.4$. Άρα ο κάτω περιορισμός ενεργοποιείται για $0.4\le x(1)\le 8/15$, οπότε έχουμε το περιορισμένο βέλτιστο $u(1)=0.2-x(1)$. Είναι εύκολο να πιστοποιηθεί ότι η τιμή αυτή ικανοποιεί όντως τις συνθήκες Kuhn-Tucker του κυρτού περιορισμένου προβλήματος και έτσι καταλήγουμε τελικά στην λύση

$$u(1)=\begin{cases}-\frac{5}{8}x(1) & \text{για} \quad \frac{8}{15}\le x(1)\le 1\\ \frac{1}{5}-x(1) & \text{για} \quad \frac{2}{5}\le x(1)\le \frac{8}{15}.\end{cases} \tag{3.15b}$$

Εξ αυτής εξάγεται το ελάχιστο κόστος μεταφοράς

$$V\left[x(1),1\right]=\begin{cases}\frac{13}{16}x(1)^2 & \text{για} \quad \frac{8}{15}\le x(1)\le 1\\ x(1)^2-\frac{1}{5}x(1)+\frac{4}{75} & \text{για} \quad \frac{2}{5}\le x(1)\le \frac{8}{15}.\end{cases}$$

Βαθμίδα 0: Για κάθε x(0), $0.6\le x(0)\le 1$, πρέπει να υπολογισθεί ο u(0) που ελαχιστοποιεί

$$\tilde{J}_0=V\left[x(1),1\right]+\frac{1}{2}\left[x(0)^2+u(0)^2\right]$$

λαμβάνοντας υπόψη x(1)=x(0)+u(0) και $0.4\le x(1)\le 1$. Αντικαθιστώντας τον περιορισμό ισότητας και την συνάρτηση $V\left[x(1),1\right]$ από την Βαθμίδα 1 και σημειώνοντας χάριν συντομίας x(0)=x, u(0)=u, έχουμε το πρόβλημα

$$\underset{u}{\text{Min}}\begin{cases}\frac{13}{16}(x+u)^2+\frac{1}{2}(x^2+u^2) & \text{για} \quad \frac{8}{15}-x\le u\le 1-x\\ (x+u)^2-\frac{1}{5}(x+u)+\frac{4}{75}+\frac{1}{2}(x^2+u^2) & \text{για} \quad \frac{2}{5}-x\le u\le \frac{8}{15}-x\end{cases}$$

λαμβάνοντας υπόψη

$$0.4-x\le u\le 1-x.$$

Η πρώτη παράγωγος ως προς u της ως άνω αντικειμενικής συνάρτησης $\tilde{J}_0$ είναι

$$\tilde{J}_0' = \begin{cases} \frac{13}{8}x + \frac{21}{8}u & \text{για} \quad \frac{8}{15} - x \le u \le 1 - x \\ 2x + 3u - 0.2 & \text{για} \quad \frac{2}{5} - x \le u \le \frac{8}{15} - x. \end{cases}$$

Είναι εύκολο να αποδειχθεί ότι, για τις τιμές x = x(0) ≤ 1 που ενδιαφέρουν εδώ, ισχύει $\tilde{J}_0' \ge 0$ εντός της επιτρεπτής περιοχής $0.4 - x \le u \le 1 - x$ και άρα έχουμε το ελάχιστο στο κάτω όριο του περιορισμού αυτού, δηλαδή

$$u(0) = 0.4 - x(0), \qquad 0.6 \le x(0) \le 1 \tag{3.16b}$$

$$V[x(0), 0] = x(0)^2 - 0.4\,x(0) + \frac{16}{75}.$$

Οι ζητούμενες βέλτιστες τροχιές για τη μεταφορά της x(0) = 1 στην x(4) = 0 λαμβάνοντας υπόψη τον περιορισμό (3.12) είναι λοιπόν (βλ. Σχ. 3.2b)

$$u^*(0) = -\frac{3}{5},\ u^*(1) = -\frac{1}{5},\ u^*(2) = -\frac{2}{15},\ u^*(3) = -\frac{1}{15}$$

$$x^*(1) = \frac{2}{5},\ x^*(2) = \frac{1}{5},\ x^*(3) = \frac{1}{15},\ x^*(4) = 0.$$

Πέραν τούτου έχουμε με τις (3.13)-(3.16) έναν βέλτιστο κανόνα ελέγχου που μεταφέρει κάθε επιτρεπτή κατάσταση

$$x(k) \in \mathscr{X}(k) = \{x(k) \mid 0.6 - k \le x(k) \le 1\}, \quad k = 0, 1, 2, 3$$

κατά βέλτιστο τρόπο στον τελικό στόχο x(4) = 0.

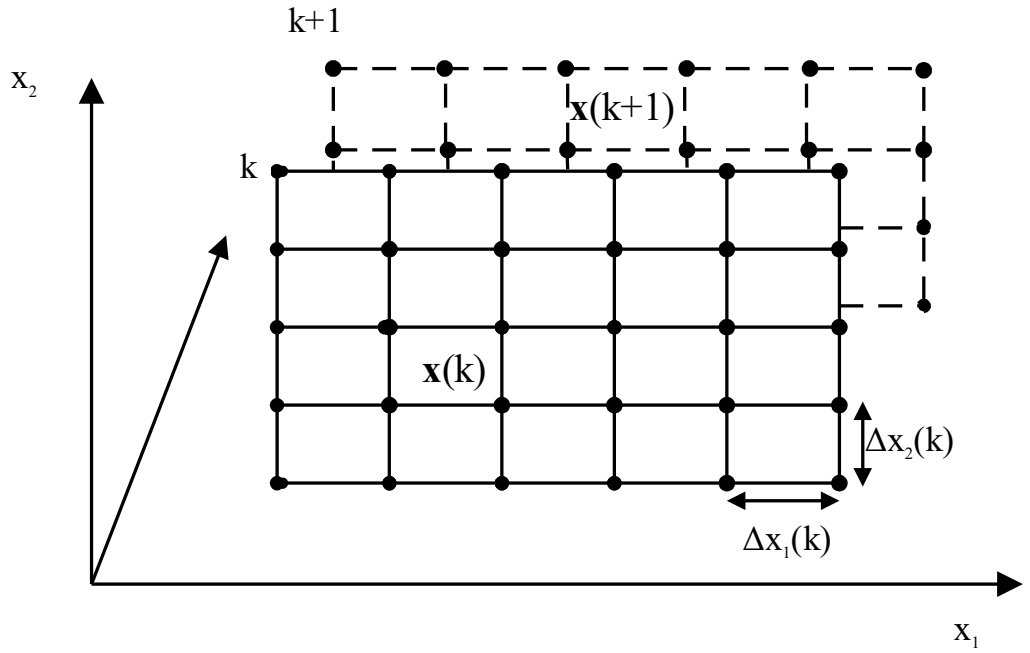
■

3.3 Διακριτός Δυναμικός Προγραμματισμός

α) Διακριτοποίηση

Η γενική αριθμητική λύση ΠΒΕ διακριτού χρόνου είναι δυνατή αν εισάγουμε ένα διακριτό πλέγμα σημείων στην επιτρεπτή περιοχή κατάστασης $\mathscr{X}(k)$ και επιτρεπτή περιοχή ελέγχου $\mathscr{U}[\mathbf{x}(k),k]$ (βλ. Σχ. 3.3). Τα διακριτά διαστήματα $\Delta\mathbf{x}(k)$ και $\Delta\mathbf{u}(k)$ μπορούν να επιλεγούν ανάλογα με το συγκεκριμένο πρόβλημα και την επιθυμητή ακρίβεια της λύσης. Σε περίπτωση απεριόριστης επιτρεπτής περιοχής κατάστασης ή ελέγχου, είναι απαραίτητο να εισαχθούν κατάλληλα όρια έτσι ώστε να υπάρχει πεπερασμένος αριθμός διακριτών σημείων.

Αν εφαρμόσουμε σε μια διακριτή κατάσταση $\mathbf{x}^i(k)$ όλους τους διακριτούς ελέγχους $\mathbf{u}^j(k)$, έχουμε ένα πεπερασμένο αριθμό μεταβάσεων στην επόμενη βαθμίδα k+1 με αντίστοιχα κόστη $\varphi[\mathbf{x}^i(k),\mathbf{u}^j(k),k]$. Εφαρμόζοντας την διαδικασία αυτή σε όλα τα διακριτά σημεία κατάστασης όλων των διακριτών χρόνων, καταλήγουμε συνολικά σε ένα διακριτό σύστημα αποφάσεων πολλαπλών βαθμίδων με την έννοια του Κεφαλαίου 2.



Σχήμα 3.3: Διακριτοποίηση της επιτρεπτής περιοχής κατάστασης για σύστημα δύο διαστάσεων.

Η εφαρμογή ενός διακριτού ελέγχου $\mathbf{u}^j(k)$ σε μια διακριτή κατάσταση $\mathbf{x}^i(k)$ οδηγεί στην κατάσταση

$$\mathbf{x}(k+1) = \mathbf{f}\left[\mathbf{x}^i(k), \mathbf{u}^j(k), k\right] \qquad (3.17)$$

της βαθμίδας k+1. Υπάρχουν δύο ιδιαίτερες περιπτώσεις που πρέπει να αναφερθούν στο σημείο αυτό:

- Η κατάσταση $\mathbf{x}(k+1)$ βρίσκεται εκτός της επιτρεπτής περιοχής $\mathscr{X}(k+1)$. Στην περίπτωση αυτή η συγκεκριμένη μετάβαση δεν λαμβάνεται υπόψη.
- Η κατάσταση $\mathbf{x}(k+1)$ δεν συμπίπτει με ένα διακριτό σημείο της βαθμίδας k+1. Στην περίπτωση αυτή μπορούμε κατά προσέγγιση να θεωρήσουμε ότι η κατάσταση $\mathbf{x}(k+1)$ συμπίπτει με το πλησιέστερο διακριτό σημείο. Αν όμως επιθυμούμε ακριβέστερη λύση, θα πρέπει να προχωρήσουμε σε κατάλληλη **παρεμβολή** όπως θα δούμε παρακάτω (Κεφ. 3.3δ). Σημειώνουμε εδώ ότι, αν η καταστατική εξίσωση (3.2) του προβλήματος είναι αντιστρεπτή ως προς $\mathbf{u}(k)$ (πράγμα που προϋποθέτει $\dim(\mathbf{x}) = \dim(\mathbf{u})$), δηλαδή αν

$$\mathbf{u}(k) = \mathbf{F}[\mathbf{x}(k), \mathbf{x}(k+1), k] \qquad (3.18)$$

μπορεί να εξαχθεί αναλυτικά από την (3.2), τότε δεν χρειάζεται εκ των προτέρων διακριτοποίηση της επιτρεπτής περιοχής ελέγχου. Αντ' αυτού, για κάθε μετάβαση από ένα διακριτό σημείο $\mathbf{x}^i(k)$ σε ένα διακριτό σημείο $\mathbf{x}^j(k+1)$ της επόμενης βαθμίδας, μπορεί να υπολογίζεται μέσω της (3.18) ο απαιτούμενος έλεγχος $\mathbf{u}^{ij}(k) = \mathbf{F}[\mathbf{x}^i(k), \mathbf{x}^j(k+1), k]$, και αν $\mathbf{u}^{ij}(k) \in \mathscr{U}[\mathbf{x}^i(k),k]$ τότε η μετάβαση είναι επιτρεπτή, άλλως η μετάβαση παραβιάζει τους περιορισμούς ελέγχου και δεν λαμβάνεται υπόψη. Με τον τρόπο αυτό, κάθε μετάβαση οδηγεί σε διακριτό σημείο της επόμενης βαθμίδας παρακάμπτοντας την ανάγκη παρεμβολής.

Το Σχ. 3.3 δείχνει το τρισδιάστατο δίκτυο επίλυσης στην περίπτωση 2 μεταβλητών κατάστασης x_1, x_2. Αντίστοιχα δίκτυα (n+1) διαστάσεων μπορούν να θεωρηθούν για την επίλυση προβλημάτων με n μεταβλητές κατάστασης x_1, x_2, …, x_n. Θα πρέπει όμως να τονισθεί πως οποιοδήποτε διακριτό (ή διακριτοποιημένο) πρόβλημα μπορεί να επιλυθεί και σε επίπεδο γράφο, όπου η μία διάσταση χαρακτηρίζει, ως συνήθως, τις βαθμίδες k, ενώ η άλλη διάσταση απεικονίζει όλους τους δυνατούς συνδυασμούς επιτρεπτών τιμών των μεταβλητών κατάστασης, π.χ., για n = 2, όλους τους συνδυασμούς $\{x_1^i, x_2^j\}$ $\forall i$ $\forall j$. Ο συνολικός αριθμός αυτών των συνδυασμών πρέπει προφανώς να είναι πεπερασμένος (βλ. Παράδ. 5.3).

Ο τελικός στόχος $\mathbf{g}[\mathbf{x}(K)] = \mathbf{0}$ μπορεί επίσης, αν απαιτείται, να προσαρμοσθεί στο διακριτό περιβάλλον του προβλήματος. Εισάγοντας μια ζώνη ανεκτικότητας $\pm\delta$ γύρω από τον στόχο αυτό, καταλήγουμε στον εξής διακριτό τελικό στόχο

$$\mathcal{G} = \left\{\mathbf{x}^i(K) \,\middle|\, \exists\boldsymbol{\beta}: \left|\mathbf{x}^i(K) - \boldsymbol{\beta}\right| \le \delta, \mathbf{g}(\boldsymbol{\beta}) = \mathbf{0}\right\}. \tag{3.19}$$

Κατά τον τρόπο αυτό μπορούμε να θεωρήσουμε ότι ο τελικός στόχος έχει επιτευχθεί (κατά προσέγγιση) εφ' όσον $\mathbf{x}(K) \in \mathcal{G}$.

β) Αλγόριθμος Διακριτού Δυναμικού Προγραμματισμού

Συμπερασματικά έχουμε κατασκευάσει ένα διακριτό περιβάλλον για την πραγματοποίηση της πολυβαθμιαίας διαδικασίας του Δυναμικού Προγραμματισμού. Αυτή η μέθοδος λύσης ΠΒΕ ονομάζεται **Διακριτός Δυναμικός Προγραμματισμός**. Σαν συνέπεια της διακριτοποίησης υπάρχει για κάθε σημείο $\mathbf{x}^i(k)$ ένας πεπερασμένος αριθμός μεταβάσεων $\mathbf{u}^j(k)$ που οδηγούν στην επόμενη βαθμίδα k+1. Η μονοβαθμιαία βελτιστοποίηση μπορεί επομένως να πραγματοποιηθεί με την άμεση σύγκριση αυτών των μεταβάσεων, όπως ακριβώς έγινε και στο Κεφάλαιο 2.1. Υπό τις συνθήκες αυτές έχουμε την ακόλουθη διαδικασία πολυβαθμιαίας βελτιστοποίησης, η οποία μπορεί εύκολα να προγραμματισθεί σε ηλεκτρονικό υπολογιστή:

Βαθμίδα K–1: Για κάθε διακριτό σημείο $\mathbf{x}^i(K-1) \in \mathcal{X}(K-1)$ πρέπει να καθορισθεί η αντίστοιχη διακριτή τιμή $\mathbf{u}^j(K-1) \in \mathcal{U}[\mathbf{x}^i(K-1), K-1]$ που ελαχιστοποιεί

$$J_{K-1} = \vartheta[\mathbf{x}(K)] + \varphi[\mathbf{x}^i(K-1), \mathbf{u}^j(K-1), K-1]$$

όπου η τελική κατάσταση

$$\mathbf{x}(K) = \mathbf{f}[\mathbf{x}^i(K-1), \mathbf{u}^j(K-1), K-1]$$

ικανοποιεί τον τελικό στόχο $\mathbf{g}[\mathbf{x}(K)] = \mathbf{0}$ ή $\mathbf{x}(K) \in \mathcal{G}$ καθώς και $\mathbf{x}(K) \in \mathcal{X}(K)$. Έστω $\mathbf{u}^{l(i)}(K-1)$ η τιμή του διακριτού ελέγχου που ανταποκρίνεται στη λύση του μονοβαθμιαίου αυτού προβλήματος ελαχιστοποίησης για κάθε $\mathbf{x}^i(K-1)$ και έστω $V[\mathbf{x}^i(K-1), K-1]$ η αντίστοιχη ελάχιστη τιμή του J_{K-1}.

Βαθμίδα K–2: Για κάθε διακριτό σημείο $\mathbf{x}^i(K-2) \in \mathcal{X}(K-2)$ πρέπει να καθορισθεί η αντίστοιχη διακριτή τιμή $\mathbf{u}^j(K-2) \in \mathcal{U}[\mathbf{x}^i(K-2), K-2]$ που ελαχιστοποιεί

$$\tilde{J}_{K-2} = V[\mathbf{x}(K-1), K-1] + \varphi[\mathbf{x}^i(K-2), \mathbf{u}^j(K-2), K-2]$$

όπου

$$\mathbf{x}(K-1) = \mathbf{f}[\mathbf{x}^i(K-2), \mathbf{u}^j(K-2), K-2].$$

Αν όμως η ως άνω τιμή $\mathbf{x}(K-1)$ δεν συμπίπτει με ένα διακριτό σημείο, δεν έχουμε και αντίστοιχη τιμή $V[\mathbf{x}(K-1), K-1]$. Στην περίπτωση αυτή μπορεί να θεωρηθεί ότι η κατάσταση $\mathbf{x}(k+1)$ συμπίπτει με το πλησιέστερο επιτρεπτό διακριτό σημείο. Εναλλακτικά μπορεί να υπολογισθεί μια τιμή $V[\mathbf{x}(K-1), K-1]$ μέσω γραμμικής παρεμβολής από τις γνωστές τιμές V των γειτονικών επιτρεπτών διακριτών σημείων (βλ. Κεφάλαιο 3.3δ). Η ως άνω ελαχιστοποίηση πραγματοποιείται μόνο για μεταβάσεις που οδηγούν σε $x(K-1) \in \mathscr{X}(K-1)$. Έστω $\mathbf{u}^{l(i)}(K-2)$ και $V[\mathbf{x}^i(K-2), K-2]$ τα αποτελέσματα της μονοβαθμιαίας ελαχιστοποίησης.

•

•

•

Βαθμίδα 0: Για $\mathbf{x}^i(0) = \mathbf{x}_0$ πρέπει να καθορισθεί η αντίστοιχη διακριτή τιμή $\mathbf{u}^j(0) \in \mathscr{U}[\mathbf{x}^i(0), 0]$ που ελαχιστοποιεί

$$\tilde{J}_0 = V[\mathbf{x}(1), 1] + \varphi[\mathbf{x}^i(0), \mathbf{u}^j(0), 0]$$

όπου

$$\mathbf{x}(1) = \mathbf{f}[\mathbf{x}^i(0), \mathbf{u}^j(0), 0].$$

Αν η $\mathbf{x}(1)$ δεν είναι ένα διακριτό σημείο, πρέπει να υπολογισθεί η $V[\mathbf{x}(1),1]$ κατά προσέγγιση όπως αναφέρθηκε και παραπάνω. Η ελαχιστοποίηση πραγματοποιείται μόνο για μεταβάσεις όπου $\mathbf{x}(1) \in \mathscr{X}(1)$. Έστω $\mathbf{u}^{l(i)}(0)$ και $V[\mathbf{x}^i(0), 0]$ τα αποτελέσματα αυτής της τελικής μονοβαθμιαίας ελαχιστοποίησης.

Επειδή κάθε μονοβαθμιαία ελαχιστοποίηση πραγματοποιείται δι' αμέσου συγκρίσεως ενός πεπερασμένου αριθμού μεταβάσεων, μπορούμε να εγγυηθούμε τον υπολογισμό απόλυτων ελάχιστων. Τότε όμως οδηγεί και η ολική πολυβαθμιαία διαδικασία σε **απόλυτο ελάχιστο** του όλου προβλήματος, καθότι, ως γνωστόν, ο Δυναμικός Προγραμματισμός επεξεργάζεται έμμεσα όλους τους συνδυασμούς μεταβάσεων. Βέβαια, αυτό το απόλυτο ελάχιστο αφορά την διακριτή προσέγγιση του προβλήματος και επομένως είναι μια κατά προσέγγιση λύση του αρχικού ΠΒΕ διακριτού χρόνου αλλά συνεχών τιμών. Η προσέγγιση αυτή είναι συνήθως περισσότερο ακριβής για μικρότερα διακριτά διαστήματα $\Delta\mathbf{x}(k)$ και $\Delta\mathbf{u}(k)$.

Το αποτέλεσμα του Διακριτού Δυναμικού Προγραμματισμού είναι ένας **βέλτιστος κανόνας ελέγχου** ο οποίος υφίσταται σε μορφή πίνακος: Για κάθε διακριτή κατάσταση $\mathbf{x}^i(k)$ γνωρίζουμε την αντίστοιχη διακριτή τιμή ελέγχου $\mathbf{u}^{l(i)}(k)$ που πραγματοποιεί μια βέλτιστη μετάβαση στην επόμενη βαθμίδα k+1. Αν θελήσουμε να εξάγουμε από τον κανόνα ελέγχου μια βέλτιστη τροχιά ελέγχου για την συγκεκριμένη αρχική κατάσταση $\mathbf{x}_0$, μπορεί και πάλι να αντιμετωπίσουμε την περίπτωση όπου κάποια απορρέουσα κατάσταση $\mathbf{x}(k+1)$ δεν συμπίπτει με ένα διακριτό σημείο, οπότε και πάλι δεν εμπεριέχεται αντίστοιχη διακριτή τιμή ελέγχου στη λύση του προβλήματος. Και σ' αυτή την περίπτωση είμαστε αναγκασμένοι να προσφύγουμε στην βέλτιστη τιμή ελέγχου του πλησιέστερου προς το $\mathbf{x}(k+1)$ επιτρεπτού διακριτού σημείου. Εναλλακτικά μπορούμε να υπολογίσουμε κατά προσέγγιση την αντίστοιχη τιμή ελέγχου μέσω

γραμμικής παρεμβολής χρησιμοποιώντας τις αντίστοιχες τιμές των γειτονικών διακριτών σημείων όπως θα δούμε στο Κεφάλαιο 3.3δ.

γ) Υπολογιστικός φόρτος

Έστω $\alpha_i(k), i = 1,...,n,$ ο αριθμός διακριτών σημείων της συνιστώσας i του διανύσματος **x**(k) και $\beta_j(k)$, j = 1,...,m, ο αριθμός διακριτών σημείων της συνιστώσας j του διανύσματος **u**(k). Το ολικό πλέγμα κατάστασης περιλαμβάνει τότε

$$\sum_{k=0}^{K} \prod_{i=1}^{n} \alpha_i(k)$$

διακριτά σημεία και ο αριθμός μεταβάσεων για κάθε σημείο είναι

$$\prod_{j=1}^{m} \beta_j(k).$$

Αν υποθέσουμε χάριν ευκολίας $\alpha_i(k) = \alpha$ και $\beta_j(k) = \beta$ για κάθε i, j, k, τότε ο απαιτούμενος χρόνος υπολογισμού τ για την επεξεργασία της πολυβαθμιαίας διαδικασίας του Διακριτού Δυναμικού Προγραμματισμού είναι ανάλογος προς τον συνολικό αριθμό κόμβων, δηλαδή

$$\tau \sim K \cdot \alpha^n \cdot \beta^m. \qquad (3.20)$$

Ο αναγκαίος χώρος αποθήκευσης του πίνακα κανόνα ελέγχου είναι m τιμές για κάθε διακριτό σημείο $\mathbf{x}^i(k)$ (για τις m συνιστώσες του διανύσματος ελέγχου $\mathbf{u}^{l(i)}(k)$), και άρα συνολικά

$$m\sum_{k=0}^{K} \prod_{i=1}^{n} \alpha_i(k) = m \cdot K \cdot \alpha^n \qquad (3.21)$$

τιμές προς αποθήκευση. Οι (3.20), (3.21) δείχνουν ότι ο υπολογιστικός φόρτος κατά την εφαρμογή του Διακριτού Δυναμικού Προγραμματισμού αυξάνει **εκθετικά** με τις διαστάσεις n, m του προβλήματος, κάτι που συνιστά σημαντικό μειονέκτημα της μεθόδου, γνωστό και ως ″κατάρα των διαστάσεων″ (curse of dimensionality).

Παράδειγμα 3.2: Για να πάρουμε μια ιδέα του τι σημαίνει εκθετική αύξηση, ας θεωρήσουμε ένα υποθετικό πρόβλημα με K=10 βαθμίδες και μονοδιάστατο έλεγχο (m = 1). Υποθέτουμε α = β = 100 διακριτά σημεία και τ_s=100 μs τον απαιτούμενο χρόνο για τον υπολογισμό του κόστους μιας μετάβασης. Εκτιμούμε τότε με την (3.20) τον απαραίτητο χρόνο υπολογισμού της ολικής λύσης μέσω Διακριτού Δυναμικού Προγραμματισμού ως εξής:

για	n = 1	τ = 10 s
για	n = 2	τ = 17 min
για	n = 3	τ = 28 h
για	n = 4	τ = 3,9 μήνες
για	n = 5	τ = 32 χρόνια.

■

Εξαιτίας της εκθετικής αύξησης του υπολογιστικού φόρτου υφίστανται στενά όρια στις διαστάσεις των προβλημάτων που μπορούν να επιλυθούν μέσω του Διακριτού Δυναμικού Προγραμματισμού. Μια σχετική διεύρυνση των ορίων αυτών είναι δυνατόν να επιτευχθεί με τους ακόλουθους τρόπους (βλ. και Κεφάλαιο 6.5):

- Για κάθε συγκεκριμένο πρόβλημα πρέπει να ελέγχεται αν η επιτρεπτή περιοχή κατάστασης και ελέγχου μπορεί να περιορισθεί περισσότερο, έτσι ώστε να μειωθεί αντίστοιχα ο αριθμός των σημείων του πλέγματος και ο αριθμός των μεταβάσεων. Βέβαια, ο επί πλέον αυτός περιορισμός πρέπει να γίνεται κατά τρόπο που να μην αποκλείει πιθανές βέλτιστες λύσεις, πράγμα που είναι συχνά δύσκολο να εκτιμηθεί **πριν** τη λύση του προβλήματος.
- Το μήκος των διακριτών διαστημάτων $\Delta\mathbf{x}(k)$, $\Delta\mathbf{u}(k)$ πρέπει να επιλέγεται όσο το δυνατόν μεγαλύτερο, δηλαδή όχι μικρότερο του αναγκαίου. Ως γνωστόν, μεγαλύτερα διαστήματα οδηγούν συνήθως σε λιγότερο ακριβείς λύσεις (βλ. όμως το Προβλ. 3.4). Δυστυχώς, η αξιόπιστη **εκ των προτέρων** εκτίμηση της ακρίβειας της λύσης είναι σε πολλές εφαρμογές δύσκολη.
- Έχουν προταθεί παραλλαγές και απλουστεύσεις της μεθοδολογίας του Διακριτού Δυναμικού Προγραμματισμού με στόχο τη μείωση του υπολογιστικού φόρτου. Μερικοί τέτοιοι αλγόριθμοι εκτίθενται στο Κεφάλαιο 6.
- Οι τελευταίες και συνεχιζόμενες εξελίξεις στην τεχνολογία υπολογιστών χρησιμοποιούνται άμεσα για την διεύρυνση των διαστάσεων των εφαρμογών του Διακριτού Δυναμικού Προγραμματισμού. Η απλή δομή της πολυβαθμιαίας διαδικασίας του Δυναμικού Προγραμματισμού προσφέρεται ιδιαίτερα για παράλληλη επεξεργασία σε συστήματα πολλαπλών υπολογιστών.

Παράδειγμα 3.3: Θεωρούμε το ακόλουθο ΠΒΕ διακριτού χρόνου:

Ελαχιστοποίηση του

$$J = \frac{1}{2}\sum_{k=0}^{3}\left[x(k)^2 + u(k)^2\right]$$

λαμβάνοντας υπόψη

$$x(k+1) = x(k) + u(k), \quad x(0) = 2, \quad x(4) = 0.$$

Παρά την έλλειψη ρητών περιορισμών, θεωρούμε τις ακόλουθες περιοχές κατάστασης και ελέγχου όπου και θα επιχειρήσουμε τη λύση του προβλήματος

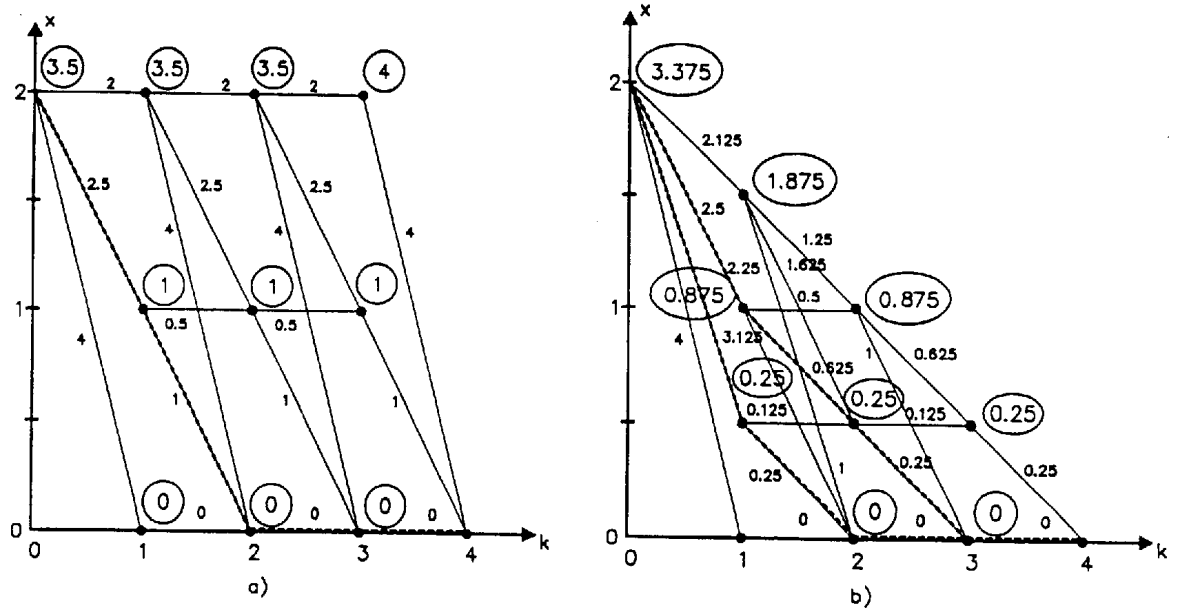
$$\mathscr{X}(k) = \{x(k) \mid 0 \le x(k) \le 2\}, \quad \mathscr{U}(k) = \{u(k) \mid -2 \le u(k) \le 0\}.$$

Τα πλέγματα και τις μεταβάσεις του προβλήματος αυτού για διακριτά διαστήματα $\Delta x = \Delta u = 1$ και $\Delta x = \Delta u = 0.5$ απεικονίζουν τα Σχ. 3.4 a) και b), αντίστοιχα. Λόγω της ειδικής καταστατικής εξίσωσης και της επιλογής των διακριτών διαστημάτων $\Delta x(k)$, $\Delta u(k)$, δεν εμφανίζεται το πρόβλημα παρεμβολής για τον υπολογισμό ενδιάμεσων τιμών. Η λύση του προβλήματος επιτυγχάνεται κατά πανομοιότυπο τρόπο όπως και στο Κεφάλαιο 2. Οι βέλτιστες τροχιές κατάστασης σημειώνονται στα δύο διαγράμματα με διακεκομμένες γραμμές. Στην περίπτωση του Σχ. 3.4b έχουμε δύο ισότιμες λύσεις. (Στο Σχ. 3.4b δεν αναπτύχθηκε πλήρως το δίκτυο δυνατών μεταβάσεων χάριν συντομίας.) Οι αντίστοιχες βέλτιστες τροχιές ελέγχου είναι:

Σχήμα 3.3 a: $u^*(0) = u^*(1) = -1,\ u^*(2) = u^*(3) = 0.$

Σχήμα 3.3 b: $u^*(0) = -1.5,\ u^*(1) = -0.5,\ u^*(2) = u^*(3) = 0$

ή: $u^*(0) = -1,\ u^*(1) = u^*(2) = -0.5,\ u^*(3) = 0.$



Σχήμα 3.4: Διακριτό πλέγμα και μεταβάσεις για το Παραδ. 3.3.

Οι υπολογισθείσες ελάχιστες τιμές του κριτηρίου κόστους J είναι 3.5 για το Σχ. 3.4a και 3.375 για το Σχ. 3.4b. Να σημειωθεί ότι το ακριβές ελάχιστο είναι 3.238, έχουμε δηλαδή μια απόκλιση 8.1% και 4.2%, αντίστοιχα, που οφείλεται στη διακριτοποίηση του προβλήματος. Ο πίνακας βέλτιστου ελέγχου στην περίπτωση του Σχ. 3.4a έχει ως εξής

		k			
		0	1	2	3
x	2	–1	–1	–1	–2
	1		–1	–1	–1
	0		0	0	0

■

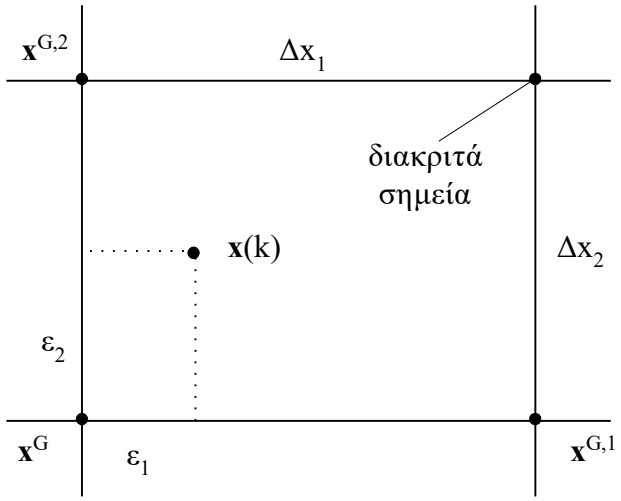
δ) Γραμμική παρεμβολή

Επανερχόμαστε τώρα στο θέμα της γραμμικής παρεμβολής που μπορεί ενδεχομένως να χρησιμοποιηθεί στον αλγόριθμο Διακριτού Δυναμικού Προγραμματισμού. Συγκεκριμένα ενδιαφερόμαστε για την τιμή V[**x**(k),k] στην περίπτωση που η **x**(k) δεν συμπίπτει με ένα σημείο του πλέγματος. Για τον υπολογισμό της τιμής V[**x**(k),k] υποθέτουμε ότι η συνάρτηση αυτή είναι κατά προσέγγιση τοπικά γραμμική και θεωρούμε το διακριτό σημείο $\mathbf{x}^G$, το οποίο ορίζεται έτσι ώστε όλες οι συνιστώσες του να είναι μικρότερες από τις αντίστοιχες συνιστώσες του σημείου **x**(k) (βλ. Σχ. 3.5), δηλαδή

$$x_j^G = l_j \Delta x_j \qquad (3.22)$$

όπου l_j ακέραιος, j = 1,...,n, και

$$x_j = l_j \Delta x_j + \varepsilon_j, \quad 0 < \varepsilon_j < \Delta x_j. \qquad (3.23)$$



Σχήμα 3.5: Ορισμός των διακριτών σημείων $\mathbf{x}^G$, $\mathbf{x}^{G,j}$ στην περίπτωση δύο διαστάσεων (n = 2).

Θεωρούμε επίσης τα γειτονικά διακριτά σημεία $\mathbf{x}^{G,j}$, j=1,...,n, με συνιστώσες $x_i^{G,j} = x_i^G$, $\forall i \neq j$, και $x_j^{G,j} = x_j^G + \Delta x_j$. Βάσει των ορισμών αυτών έχουμε τότε τον ακόλουθο τύπο γραμμικής παρεμβολής

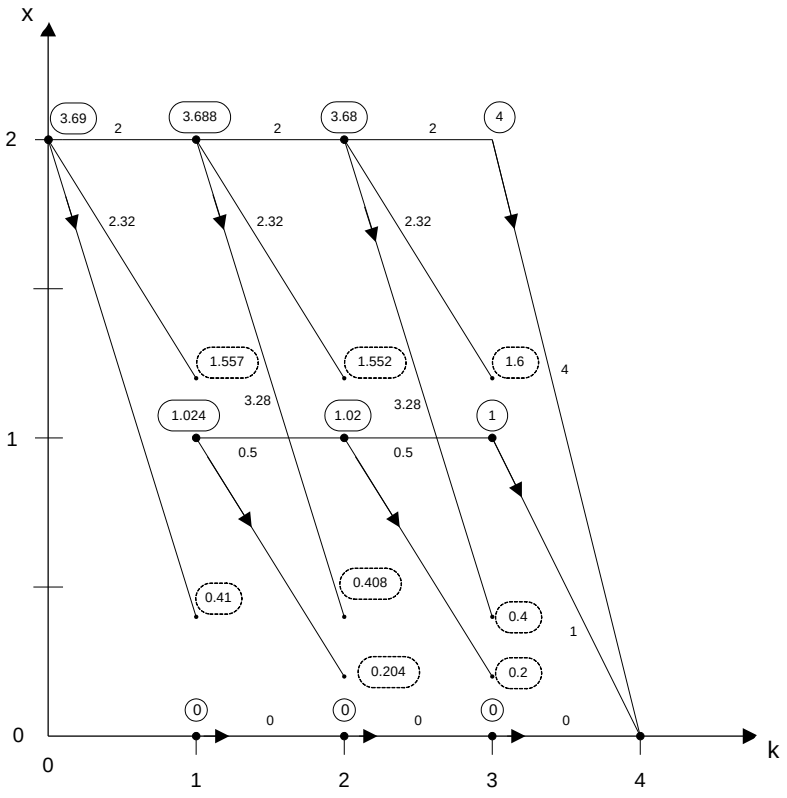
$$V[\mathbf{x}(k),k] = V[\mathbf{x}^G(k),k] + \sum_{j=1}^{n}\{V[\mathbf{x}^{G,j}(k),k] - V[\mathbf{x}^G(k),k]\}\,\varepsilon_j/\Delta x_j\,. \quad (3.24)$$

Μια ενδεχόμενη επιπλοκή παρουσιάζεται αν κάποια από τα γειτονικά διακριτά σημεία $\mathbf{x}^G$, $\mathbf{x}^{G,j}$ δεν είναι επιτρεπτά και άρα δεν υπάρχουν αντίστοιχες τιμές της συνάρτησης V. Στην περίπτωση αυτή μπορεί να θεωρηθεί ως V[**x**(k),k] κατά προσέγγιση η τιμή V του πλησιέστερου επιτρεπτού διακριτού σημείου.

Το ερώτημα της γραμμικής παρεμβολής μπορεί επίσης να εμφανισθεί μετά την επίλυση του προβλήματος, κατά την εξαγωγή μιας βέλτιστης τροχιάς από τον πίνακα κανόνα ελέγχου **R**[**x**(k), k], αν η κατάσταση **x**(k) δεν συμπίπτει με ένα διακριτό σημείο. Και στην περίπτωση αυτή μπορεί να ακολουθηθεί η περιγραφείσα διαδικασία γραμμικής παρεμβολής, αυτή τη φορά για την διανυσματική συνάρτηση **R**[**x**(k),k], δηλαδή για κάθε συνιστώσα της. Η συνάρτηση όμως αυτή μπορεί να είναι για πολλά προβλήματα ασυνεχής, π.χ. αν έχουμε την επιτρεπτή περιοχή ελέγχου $\mathcal{U} = \{0, 1\}$, οπότε η γραμμική παρεμβολή ενδέχεται να οδηγήσει σε μη επιτρεπτά αποτελέσματα, π.χ. $u(k) = 0.5 \notin \mathcal{U}$. Επίσης, ακόμη και αν η συνάρτηση **R**[**x**(k), k] δεν είναι ασυνεχής, μπορεί η γραμμική παρεμβολή να οδηγήσει σε μη επιτρεπτό έλεγχο (π.χ. υπέρβαση ορίων επιτρεπτών τιμών), παρότι ο βέλτιστος έλεγχος των γειτονικών διακριτών σημείων είναι επιτρεπτός. Τέλος, ακόμη και αν η υπολογισθείσα τιμή ελέγχου είναι επιτρεπτή, μπορεί να οδηγεί σε μη επιτρεπτό σημείο $\mathbf{x}(k+1) \notin \mathcal{X}(k+1)$ της επόμενης βαθμίδας, παραβιάζοντας έτσι τους περιορισμούς κατάστασης. Για να αποφευχθούν αυτές οι πιθανές παρενέργειες, μπορεί, εναλλακτικά, κατά τον υπολογισμό της βέλτιστης τροχιάς ελέγχου, να επιλύεται σε κάθε βαθμίδα το ακόλουθο μονοβαθμιαίο πρόβλημα ελαχιστοποίησης:

> Για δεδομένο **x**(k) πρέπει να υπολογισθεί η διακριτή τιμή $\mathbf{u}^j(k) \in \mathcal{U}[\mathbf{x}(k),k]$ που ελαχιστοποιεί

$$\tilde{J}_k = V[\mathbf{x}(k+1), k+1] + \varphi[\mathbf{x}(k), \mathbf{u}^j(k), k] \quad (3.25)$$



Σχήμα 3.6: Διακριτό πλέγμα και μεταβάσεις για το Παραδ. 3.4 (γραμμική παρεμβολή).

όπου

$$\mathbf{x}(k+1) = \mathbf{f}[\mathbf{x}(k), \mathbf{u}^j(k), k]\,, \quad \mathbf{x}(k+1) \in X\,(k+1).$$

Αν η **x**(k+1) δεν συμπίπτει με ένα διακριτό σημείο, χρησιμοποιείται η γραμμική παρεμβολή (3.24) για τον υπολογισμό της τιμής V[**x**(k+1), k+1].

Παράδειγμα 3.4: Θεωρούμε και πάλι το πρόβλημα του Παραδ. 3.3, τώρα όμως με τα διακριτά διαστήματα Δx(k) = 1, k = 1,2,3, και Δu(0) = Δu(1) = Δu(2) = 0.8, Δu(3) = 0.2. Με την επιλογή αυτή οι μεταβάσεις από βαθμίδα σε βαθμίδα δεν οδηγούν πάντα σε διακριτές καταστάσεις (βλ. Σχ. 3.6) οπότε εφαρμόζουμε γραμμική παρεμβολή. Για παράδειγμα, το σημείο x(3) = 1.2 δεν συμπίπτει με ένα διακριτό σημείο και έτσι έχουμε με εφαρμογή της (3.24)

$$V(1.2,3) = V(1,3) + [V(2,3) - V(1,3)]0.2 = 1.6.$$

Οι τιμές της συνάρτησης V που υπολογίζονται με γραμμική παρεμβολή σημειώνονται στο Σχ. 3.6 σε διακεκομμένο πλαίσιο. Ας σημειωθεί ότι η χρήση γραμμικής παρεμβολής οδηγεί σε υπολογισμό των τιμών της συνάρτησης V κατά προσέγγιση, όχι μόνο για τα σημεία όπου εφαρμόζεται η (3.24), αλλά, κατ' ακολουθία, και σε διακριτά σημεία μικρότερων βαθμίδων. Για το λόγο αυτό, το υπολογισθέν ελάχιστο κόστος V(2,0) = 3.69 (Σχ. 3.6) δεν είναι παρά μια προσέγγιση. Για την εξαγωγή της τροχιάς βέλτιστου ελέγχου από τα αποτελέσματα του Σχ. 3.6, έχουμε καταρχήν για x(0) = 2 τον u(0) = –1.6 που μας οδηγεί στην x(1) = 0.4. Η κατάσταση αυτή δεν είναι διακριτό σημείο, άρα πρέπει να επιλύσουμε το πρόβλημα (3.25) που δίδει έναν μόνο επιτρεπτό έλεγχο u(1) = 0. Φθάνουμε έτσι στην x(2) = 0.4 και κατά παρόμοιο τρόπο στον u(2) = 0 και επομένως στην x(3) = 0.4. Από το σημείο αυτό, οδηγεί ο u(3) = –0.4 στον τελικό στόχο x(4) = 0. Αντικαθιστώντας τις βέλτιστες αυτές τροχιές στο

κριτήριο κόστους του προβλήματος έχουμε J* = 3.6 που είναι χειρότερο από τα βέλτιστα αποτελέσματα των Σχ. 3.4a και b.

Αν, αντί της επίλυσης του προβλήματος (3.25), θελήσουμε να εφαρμόσουμε γραμμική παρεμβολή για την εξαγωγή της τροχιάς βέλτιστου ελέγχου, έχουμε και πάλι για x(0) = 2 τον u(0) = –1.6 που οδηγεί στην x(1) = 0.4. Στο σημείο αυτό χρησιμοποιούμε τις βέλτιστες τιμές ελέγχου των γειτονικών διακριτών σημείων και υπολογίζουμε μέσω γραμμικής παρεμβολής την τιμή ελέγχου –0.32. Εδώ έχουμε την δυνατότητα είτε να αποδεχθούμε αυτή τη μη διακριτή τιμή ελέγχου, είτε να την αντικαταστήσουμε με την κοντινότερη διακριτή τιμή, δηλαδή 0. Στην δεύτερη περίπτωση καταλήγουμε στην ίδια βέλτιστη τροχιά ελέγχου όπως και παραπάνω. Στην πρώτη περίπτωση καταλήγουμε στην x(2) = 0.08, οπότε με γραμμική παρεμβολή έχουμε u(2) = –0.064 που οδηγεί στην x(3) = 0.016 εκ της οποίας οδηγούμαστε με u = –0.016 στον τελικό στόχο x(4) = 0. Με την βέλτιστη αυτή τροχιά έχουμε ελάχιστη τιμή του κριτηρίου κόστους J^* = 3.417. Η βελτίωση οφείλεται στο ότι επιτρέψαμε μικρότερα διαστήματα ελέγχου απ' ό, τι στις προηγούμενες περιπτώσεις.

■

ε) Τελικά συμπεράσματα

Παράδειγμα 3.5: Θεωρούμε την ελαχιστοποίηση του κριτηρίου κόστους

$$J = \frac{1}{2}\sum_{k=0}^{K-1}[200x_1(k)^2 + 200x_2(k)^2 + u(k)^2]$$

για το δυναμικό σύστημα

$$x_1(k+1) = x_1(k) + (1-e^{-T})\,x_2(k) + (e^{-T} + T - 1)\,u(k),\ \ x_1(0) = 0$$

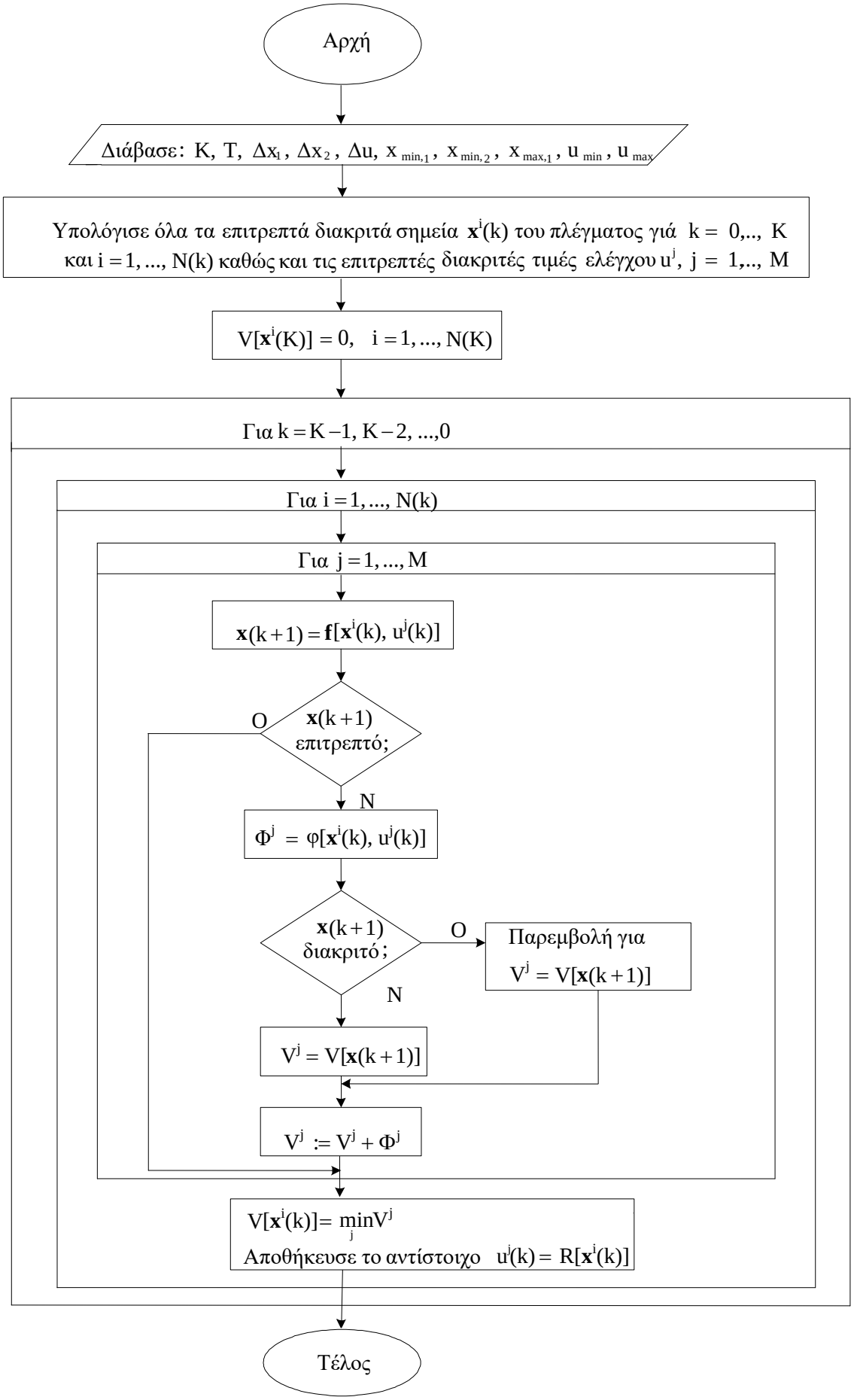
$$x_2(k+1) = e^{-T}x_2(k) + (1-e^{-T})u(k),\ \ x_2(0) = -1$$

με τον περιορισμό ανισότητας κατάστασης

$$x_2(k) \le 8\,(kT - \frac{1}{2})^2 - \frac{1}{2}\,.$$

Το διακριτό διάστημα χρόνου είναι T = 0.01. Επειδή οι μεταβλητές x_1 και u δεν έχουν περιορισμούς και η x_2 έχει μόνον άνω περιορισμό, απαιτείται η εισαγωγή εικονικών περιορισμών $x_{min,1} \le x_1(k) \le x_{max,1}$, $u_{min} \le u(k) \le u_{max}$, $x_{min,2} \le x_2(k)$, ώστε να καταστεί δυνατή η αριθμητική επίλυση του προβλήματος. Η επιλογή των περιορισμών δεν πρέπει βέβαια να επηρεάσει την λύση του προβλήματος. Στο παρόν πρόβλημα επελέγησαν $x_{min,1} = -0.5,\ x_{max,1} = 0,\ u_{min} = -6,\ \ u_{max} = 14$, $x_{min,2}(k) = 8(kT-0.5)^2 - 0.5 - 3.$

Το Σχ. 3.7 απεικονίζει το διάγραμμα ροής του προγράμματος αριθμητικής επίλυσης του προβλήματος, ενώ ο Πίνακας 3.1 δίδει τα αποτελέσματα επίλυσης του προβλήματος σε (σήμερα απαρχαιωμένο) υπολογιστή PENTIUM 4 1.8 GHz για $K = 50$ με διαφορετικές διακριτοποιήσεις καθώς και για την ακριβή λύση (υπολογισθείσα μέσω αλγορίθμου Τετραγωνικού Προγραμματισμού). Κατά την επίλυση δεν χρησιμοποιήθηκε γραμμική παρεμβολή, αλλά επελέγει το εκάστοτε πλησιέστερο επιτρεπτό διακριτό σημείο για την τιμή της συνάρτησης V καθώς και κατά την εξαγωγή της βέλτιστής τροχιάς ελέγχου μετά την



Σχήμα 3.7: Διάγραμμα ροής του προγράμματος αριθμητικής επίλυσης του Παραδ. 3.5.

	Διακριτοποίηση			Ακριβής λύση
	1	2	3	
Δx_1	$2\cdot10^{-3}$	10^{-3}	$5\cdot10^{-4}$	0
Δx_2	10^{-2}	$5\cdot10^{-3}$	$25\cdot10^{-4}$	0
Δu	1	0.5	0.25	0
J^*	5.804	5.685	5.658	5.626
αριθμός μεταβάσεων	$7.875\cdot10^7$	$6.15\cdot10^8$	$4.86\cdot10^9$	
χρόνος CPU	5.3 min	41.3 min	5h 29min	

Πίνακας 3.1: Αποτελέσματα διαφορετικών διακριτοποιήσεων για το Παραδ. 3.5.

επίλυση του προβλήματος. Παρατηρούμε ότι η προσδιορισθείσα βέλτιστη τιμή J^* του κριτηρίου κόστους μειώνεται για μικρότερη διακριτοποίηση, προσεγγίζοντας περισσότερο την βέλτιστη τιμή J^* της ακριβούς λύσης, αυξάνει όμως (με τι ρυθμό;) παράλληλα και ο απαιτούμενος υπολογιστικός φόρτος. Το Σχ. 3.8 απεικονίζει τις προσδιορισθείσες τροχιές ελέγχου για τις τρεις διακριτοποιήσεις του Πίνακα 3.1 καθώς και την ακριβή βέλτιστη τροχιά ελέγχου. Η οδοντωτή μορφή της διακριτοποιημένης τροχιάς ελέγχου οφείλεται στην επιλογή μόνο διακριτών τιμών ελέγχου. Παρατηρούμε την βελτίωση της διακριτοποιημένης προσέγγισης για μειούμενα διακριτά διαστήματα. Το Σχ. 3.9 απεικονίζει τις βέλτιστες τροχιές κατάστασης (ακριβής λύση πάνω, και λύση διακριτοποίησης αρ. 3 κάτω). Παρατηρούμε ότι η διαφορά της διακριτοποιημένης από την ακριβή λύση είναι μηδαμινή.

■

Τα σημαντικά πλεονεκτήματα του Διακριτού Δυναμικού Προγραμματισμού κατά τη λύση ΠΒΕ διακριτού χρόνου μπορούν να συνοψισθούν ως εξής:

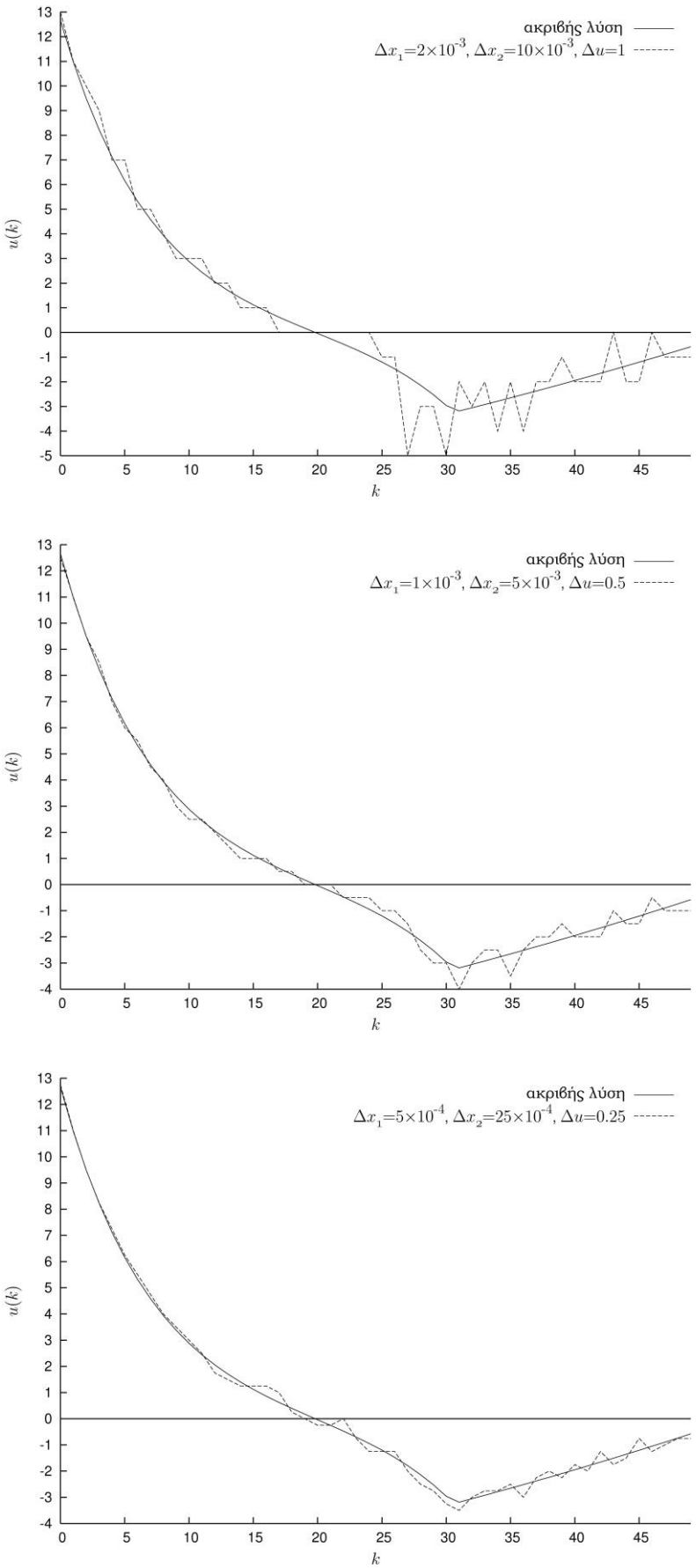
- Η παρουσία **περιορισμών ανισότητας** ευκολύνει παρά δυσκολεύει τη λύση.
- Η λύση αντιπροσωπεύει το **απόλυτο ελάχιστο**, του διακριτοποιημένου βέβαια προβλήματος.
- Η λύση παίρνει την μορφή (πίνακα) **κανόνα ελέγχου**.
- Οι συναρτήσεις ϑ, φ, **f**, **h**, **g** του προβλήματος δεν χρειάζεται να είναι **συνεχείς ή παραγωγίσιμες**, μπορούν μάλιστα και να δοθούν σε μορφή πίνακα.
- Τα υπολογιστικά προγράμματα επίλυσης είναι γενικά εφαρμόσιμα, απλά και ευκρινή.

Τα σημαντικότερα μειονεκτήματα της μεθόδου συνοψίζονται ως εξής:

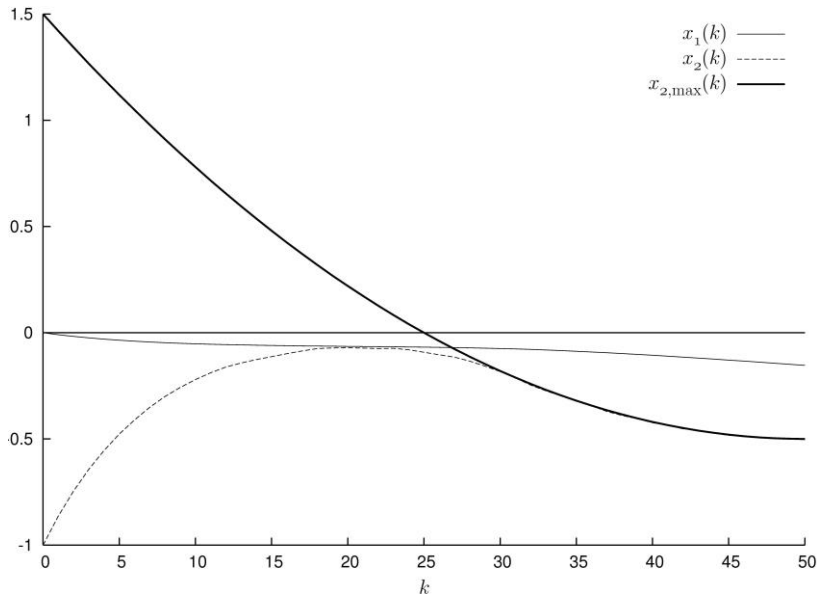
- Η λύση υπολογίζεται κατά **προσέγγιση** (λόγω διακριτοποίησης και γραμμικής παρεμβολής).
- Ο υπολογιστικός φόρτος αυξάνει **εκθετικά** με τις διαστάσεις του προβλήματος.

3.4 Επεκτάσεις του Προβλήματος

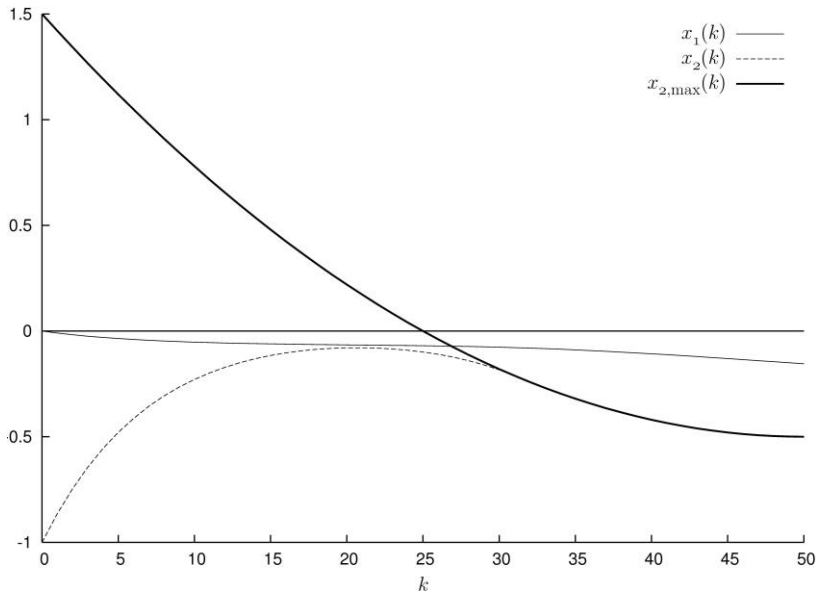
Στο κεφάλαιο αυτό θα θεωρήσουμε μερικές επεκτάσεις του ΠΒΕ διακριτού χρόνου και την αντίστοιχη επίλυσή τους μέσω Δυναμικού Προγραμματισμού.



Σχήμα 3.8: Βέλτιστες τροχιές ελέγχου για το Παράδ. 3.5: ακριβής λύση και λύση βάσει τριών διακριτοποιήσεων.



a) ακριβής λύση



b) λύση με την διακριτο-ποίηση αρ. 3

Σχήμα 3.9: Βέλτιστες τροχιές κατάστασης για το Παράδ. 3.5.

α) Διαχρονικοί περιορισμοί

Θεωρούμε το ΠΒΕ διακριτού χρόνου (3.1)-(3.5) και επιπλέον τον **διαχρονικό περιορισμό ανισότητος**

$$\sum_{k=k_1}^{k_2} G\left[\mathbf{x}(k), \mathbf{u}(k), k\right] \le 0 \qquad (3.26)$$

όπου $0 \le k_i \le K$, $i = 1,2$. **Διαχρονικοί περιορισμοί ισότητος** μπορούν να συμπεριληφθούν στο πρόβλημα κατά παρόμοιο τρόπο.

Το επεκταθέν πρόβλημα μπορεί να αναχθεί στο κανονικό πρόβλημα του Κεφαλαίου 3.1 μέσω της ακόλουθης εικονικής εξίσωσης κατάστασης

$$x_{n+1}(k+1)=\begin{cases}0 & \text{για } k<k_1\\ x_{n+1}(k)+G[\mathbf{x}(k),\mathbf{u}(k),k] & \text{για } k_1\le k\le k_2\\ x_{n+1}(k) & \text{για } k>k_2\end{cases} \tag{3.27}$$

με τον περιορισμό ανισότητας

$$x_{n+1}(k_2)\le 0. \tag{3.28}$$

Ο μετασχηματισμός αυτός επιτρέπει την επίλυση, μεταξύ άλλων, **προβλημάτων βέλτιστης κατανομής** (βλ. Κεφάλαιο 5.3) όπως αυτό του ακόλουθου παραδείγματος.

Παράδειγμα 3.6: Ένα φορτηγό πρόκειται να φορτώσει u_0 αυτοκίνητα, u_1 επαγγελματικά ψυγεία και u_2 ηλεκτρικές κουζίνες. Το βάρος b_i και η αξία w_i ανά τεμάχιο κάθε είδους είναι $b_0 = 4$, $b_1 = 0.4$, $b_2 = 0.1$ και $w_0 = 3$, $w_1 = 0.28$, $w_2 = 0.05$, αντίστοιχα. Το ολικό βάρος του φορτίου δεν πρέπει να υπερβαίνει B = 11. Ζητείται η κατανομή του φορτίου που μεγιστοποιεί την συνολική αξία του φορτίου.

Θεωρούμε την εξίσωση κατάστασης

$$x(k+1)=x(k)+u(k)\cdot b_k,\ x(0)=0$$

με τους περιορισμούς

$$u(k)\in\{0,1,2,...\},\quad 0\le x(k)\le B$$

και το κριτήριο αξίας φορτίου

$$J=\sum_{k=0}^{2}u(k)w_k.$$

Με τον τρόπο αυτό, το πρόβλημα έχει αναχθεί σε κανονικό ΠΒΕ που μπορεί να επιλυθεί με Διακριτό Δυναμικό Προγραμματισμό. Για να επιτύχουμε ακριβή λύση, θεωρούμε τα διακριτά διαστήματα $\Delta x(1)=b_0$, $\Delta x(2)=b_1$, $\Delta x(3)=b_2$ και έχουμε το πλέγμα του Σχ. 3.10. (Ας σημειωθεί ότι αυτή η διακριτοποίηση είναι εδώ δυνατή επειδή το b_0 είναι πολλαπλάσιο των b_1, b_2, και το b_1 είναι πολλαπλάσιο του b_2. Στη γενική περίπτωση θα πρέπει το διακριτό διάστημα $\Delta = \Delta x(1) = ... = \Delta x(K)$ να είναι κοινός παρανομαστής των $b_0, b_1,...,b_{K-1}$ για την επίτευξη ακριβούς λύσης, βλ. Πρόβλημα 3.7.) Καταρχήν είναι προφανές ότι για τη βέλτιστη λύση ισχύει x(3) = 11 και επομένως έχουμε

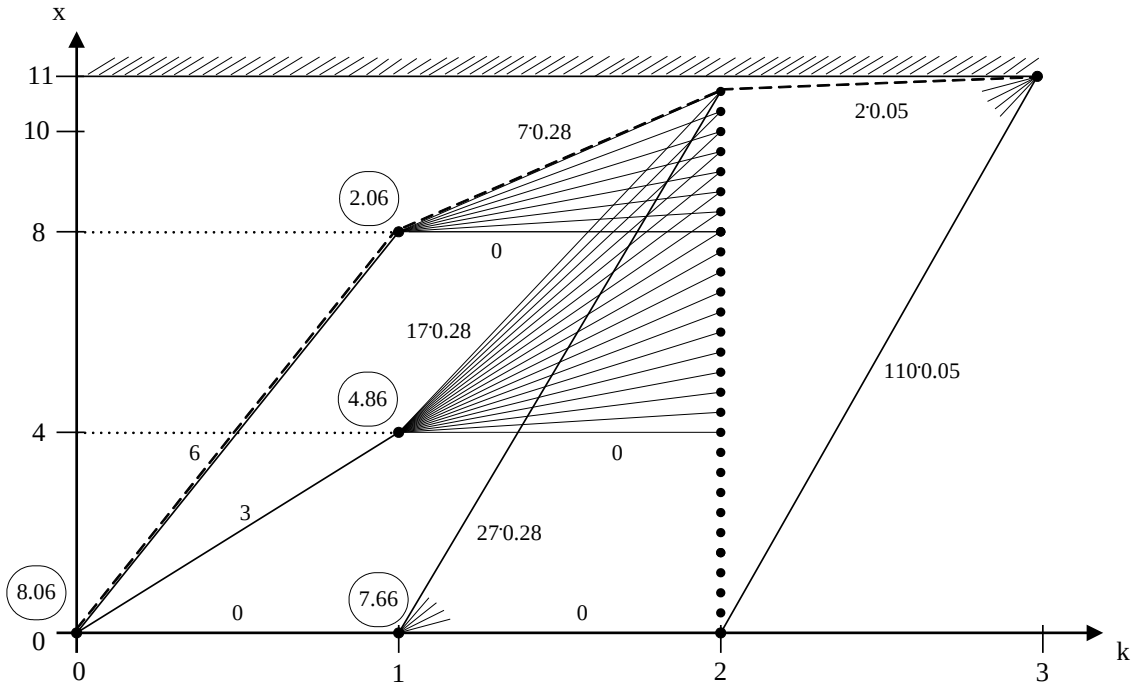
$$V[x(2),2]=[11-x(2)]w_2/b_2.$$

Για τη βαθμίδα k = 1 πρέπει να μεγιστοποιηθεί

$$\tilde{J}_1=u(1)\cdot 0.28+[11-u(1)\cdot b_1-x(1)]w_2/b_2=0.08u(1)+5.5-0.5x(1)$$

λαμβάνοντας υπόψη $x(2) = x(1) + b_1u(1) \le 11$. Προφανώς, η λύση αυτής της μονοβαθμιαίας μεγιστοποίησης δίδεται από το μεγαλύτερο δυνατό u(1), έτσι ώστε να μη παραβιασθεί η ανισότητα, και οδηγεί στις βέλτιστες τιμές V[x(1), 1] που σημειώνονται στο Σχ. 3.10. Η συνέχεια της επίλυσης είναι απλούστερη επί του Σχ. 3.10 και οδηγεί στην κατανομή $u_0 = 2$, $u_1 = 7$, $u_2 = 2$ με μέγιστη αξία φορτίου 8.06.

■



Σχήμα 3.10: Επίλυση του Παραδ. 3.6.

β) Προβλήματα Ελεύθερου Τελικού Χρόνου

Προβλήματα **ελεύθερου τελικού χρόνου** επιχειρούν την βέλτιστη μεταφορά της κατάστασης ενός συστήματος στον τελικό στόχο (αν υπάρχει)

$$\mathbf{g}[\mathbf{x}(K),K] = \mathbf{0} \tag{3.29}$$

χωρίς να προδιαγράφουν τον τελικό χρόνο K, δηλαδή τον αριθμό των βαθμίδων του προβλήματος. Κατά συνέπεια, οι τελικές καταστάσεις του συστήματος μπορούν να επεκτείνονται σε περισσότερες της μιας βαθμίδος. Μια ειδική περίπτωση των προβλημάτων ελεύθερου χρόνου είναι τα **προβλήματα ελάχιστου χρόνου** όπου επιχειρείται η μεταφορά της κατάστασης του συστήματος στον τελικό στόχο (3.29) σε ελάχιστο χρόνο. Το υπό ελαχιστοποίηση κριτήριο κόστους στην περίπτωση αυτή είναι

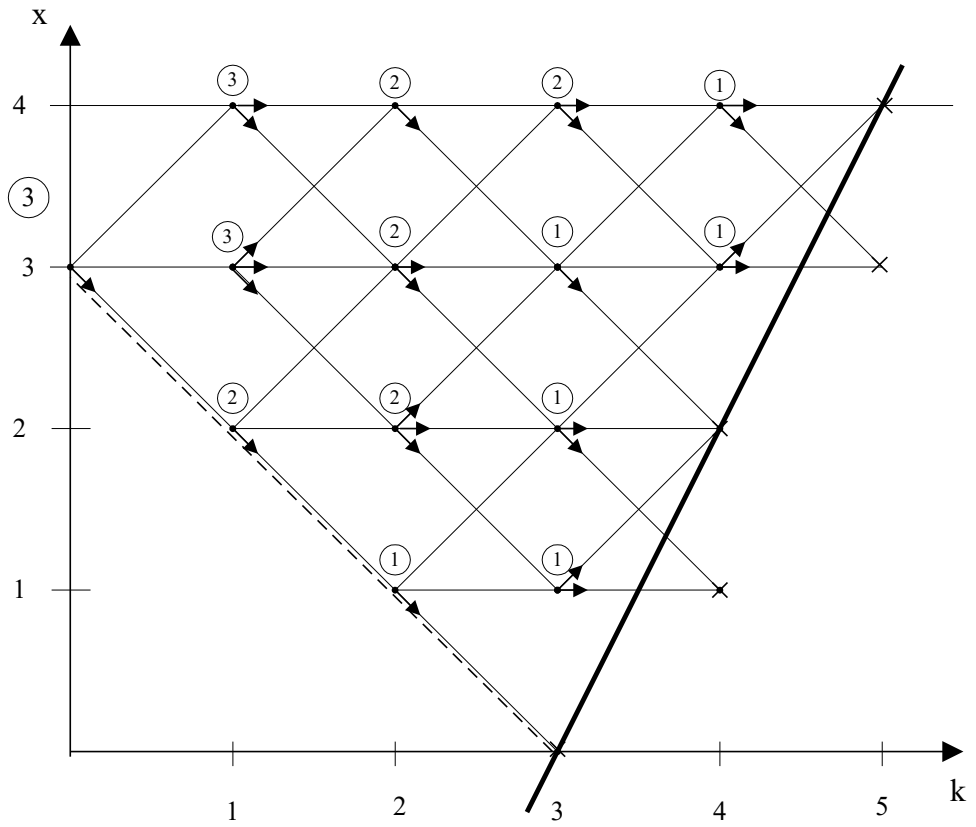
$$J = K = \sum_{k=0}^{K-1} 1 \tag{3.30}$$

έχουμε δηλαδή $\vartheta = 0$ και $\varphi = 1$ ή, εναλλακτικά, $\vartheta = K$ και $\varphi = 0$.

Προβλήματα ελεύθερου τελικού χρόνου μπορούν να επιλυθούν με την μέθοδο Διακριτού Δυναμικού Προγραμματισμού, επιβάλλεται όμως ιδιαίτερη προσοχή στον καθορισμό και την επεξεργασία του διακριτού πλέγματος λόγω της προαναφερθείσας επέκτασης του τελικού στόχου σε περισσότερες της μιας βαθμίδος.

Παράδειγμα 3.7: Θεωρούμε το δυναμικό σύστημα

$$x(k+1) = x(k) + u(k), \quad x(0) = 3$$



Σχήμα 3.11: Επίλυση του Παραδ. 3.7.

με τους περιορισμούς

$$u(k) \in U = \{1, 0, -1\}$$
$$x(k) \in X = [0, 4].$$

Ζητείται η βέλτιστη τροχιά ελέγχου για τη μεταφορά της αρχικής κατάστασης σε ελάχιστο χρόνο στην τελική τροχιά

$$x(K) - 2K + 6 = 0.$$

Για την εφαρμογή Διακριτού Δυναμικού Προγραμματισμού θεωρούμε διακριτοποίηση $\Delta x = 1$. Διακριτοποιώντας την τελική τροχιά κατά προσέγγιση (βλ. Σχ. 3.11), θεωρούμε ότι έχουμε τα διακριτά σημεία τελικού στόχου

$$[x(K), K] \in \{(0,3), (1,4), (2,4), (3,5), (4,5)\}.$$

Η επίλυση του προβλήματος μπορεί έτσι να επιτευχθεί κατά πανομοιότυπο τρόπο όπως στο Κεφάλαιο 2 λαμβάνοντας υπόψη ότι το κόστος κάθε μετάβασης αντιστοιχεί σε 1. Ο ελάχιστος τελικός χρόνος και η βέλτιστη τροχιά σημειώνονται στο Σχ. 3.11.

■

γ) Μη Μεταβλητός Κανόνας Ελέγχου

Τα ΠΒΕ που θεωρήσαμε μέχρι τώρα, επιλύονται μέσω ενός βέλτιστου κανόνα ελέγχου που εξαρτάται από την τιμή κατάστασης **x** και τον χρόνο k. Σε πολλές εφαρμογές, π.χ. στην περιοχή του αυτόματου ελέγχου, είναι επιθυμητοί κανόνες ελέγχου

$$\mathbf{u}(k) = \overline{\mathbf{R}}[\mathbf{x}(k)] \tag{3.31}$$

που εξαρτώνται από την κατάσταση, όχι όμως από τον χρόνο (**μη μεταβλητοί κανόνες ελέγχου**). Το γενικό ΠΒΕ οδηγεί σε μη μεταβλητό βέλτιστο κανόνα ελέγχου αν εκπληρούνται οι ακόλουθες προϋποθέσεις:

- Οι συναρτήσεις του προβλήματος φ, **f**, **h**, **g** δεν εξαρτώνται από τον χρόνο k.
- Η συνάρτηση φ είναι θετικά ημιορισμένη.
- Οι περιορισμοί του προβλήματος δεν εξαρτώνται από τον χρόνο k.
- Ο χρονικός ορίζοντας K του προβλήματος είναι είτε ελεύθερος είτε άπειρος.

Στις περιπτώσεις αυτές, η συνάρτηση ελάχιστου κόστους V είναι και αυτή ανεξάρτητη του χρόνου k, έχουμε δηλαδή V[**x**(k)]. Κατά συνέπεια ο τύπος (3.9) του Bellman δίδει (εφ' όσον K>0)

$$V[\mathbf{x}(k)] = \min\{\varphi[\mathbf{x}(k), \mathbf{u}(k)] + V[\mathbf{f}[\mathbf{x}(k), \mathbf{u}(k)]]\} \tag{3.32}$$

όπου η ελαχιστοποίηση νοείται ως προς **u**(k) λαμβάνοντας υπόψη όλους τους περιορισμούς. Σε αντίθεση με την (3.9), η (3.32) δεν είναι μια αναδρομική εξίσωση, αλλά μια **συναρτησιακή εξίσωση**, καθότι έχουμε την ίδια συνάρτηση V και στις δύο πλευρές της. Η αναλυτική επίλυση συναρτησιακών εξισώσεων είναι δυνατή μόνο για πολύ απλά προβλήματα.

Σε περίπτωση **άπειρου ορίζοντα K** θα πρέπει η τιμή ανά βαθμίδα του ελάχιστου κριτηρίου κόστους να είναι πεπερασμένη, τουλάχιστον για μερικά αρχικά σημεία $\mathbf{x}_0$. Εννοείται ότι, λόγω του άπειρου τελικού χρόνου K, δεν τίθεται θέμα ύπαρξης τελικού στόχου (3.3). Στην περίπτωση άπειρου τελικού χρόνου K (και εκπλήρωσης των προαναφερθέντων προϋποθέσεων), η μέθοδος του Δυναμικού Προγραμματισμού οδηγεί σε μη μεταβλητό βέλτιστο κανόνα ελέγχου, του οποίου ο υπολογισμός επιτυγχάνεται αρχίζοντας από έναν (πεπερασμένο) εικονικό τελικό χρόνο K και έναν εικονικό τελικό στόχο $\mathbf{x}(K) \in X(K)$ με μηδενικό τελικό κόστος, και προχωρώντας όπισθεν, υπολογίζοντας $\mathbf{R}[\mathbf{x}(k-1), K-1]$, $\mathbf{R}[\mathbf{x}(K-2), K-2], \ldots$, για έναν ικανό αριθμό βαθμίδων, μέχρι την (ενδεχομένως ασυμπτωτική) σύγκλιση στο μη μεταβλητό (στάσιμο) κανόνα ελέγχου $\overline{\mathbf{R}}[\mathbf{x}(k)]$. Σε περίπτωση αριθμητικού υπολογισμού αρκεί ένας πεπερασμένος αριθμός βαθμίδων, έως ότου να μην υπάρχει σημαντική μεταβολή του κανόνα ελέγχου από βαθμίδα σε βαθμίδα.

Παράδειγμα 3.8: Θεωρούμε το πρόβλημα του εισοδηματία (βλ. Παραδ. Α.2 και Β.2) με καταστατική εξίσωση

$$x(k+1) = \alpha u(k), \quad x(0) = x_0,$$

περιορισμούς ανισότητας

$$0 \le u(k) \le x(k)$$

τελικό κεφάλαιο $x(K) = x_e$ και κριτήριο προς μεγιστοποίηση

$$J = \sum_{k=0}^{K-1} \sqrt{x(k) - u(k)}.$$

(Τί σημαίνει το κριτήριο αυτό στην παρούσα εφαρμογή;)

Εφαρμόζοντας την μεθοδολογία του Δυναμικού Προγραμματισμού (βλ. Κεφάλαιο 3.2) επιχειρούμε την βαθμιαία επίλυση του προβλήματος:

Βαθμίδα K–1: Υπάρχει μια και μόνη λύση για κάθε x(K–1) που ικανοποιεί τον τελικό στόχο

$$x_e = x(K) = \alpha \cdot u(K-1) \Rightarrow u(K-1) = \frac{x_e}{\alpha} \qquad \forall x(K-1) \geq \frac{x_e}{\alpha}.$$

Το αντίστοιχο μέγιστο κριτήριο είναι

$$V\left[x(K-1), K-1\right] = \sqrt{x(K-1) - \frac{x_e}{\alpha}}.$$

Βαθμίδα K–2: Για κάθε x(K–2) ζητείται το u(K–2) που μεγιστοποιεί

$$\tilde{J}_{K-2} = \sqrt{x(K-2) - u(K-2)} + V\left[\alpha \cdot u(K-2), K-1\right].$$

Αντικαθιστώντας τη συνάρτηση V της προηγούμενης βαθμίδας και επιλύοντας την αναγκαία συνθήκη $\partial \tilde{J}_{K-2}/\partial u(K-2) = 0$, προσδιορίζουμε (μετά από αρκετές αλγεβρικές πράξεις) τη λύση

$$u(K-2) = \frac{\alpha}{1+\alpha} x(K-2) + \frac{x_e}{\alpha^2(1+\alpha)} \qquad \forall x(K-2) \geq \frac{x_e}{\alpha^2}.$$

Αντικαθιστώντας τη λύση αυτή στο $\tilde{J}_{K-2}$ προσδιορίζουμε (μετά από αρκετές αλγεβρικές πράξεις) την μέγιστη τιμή:

$$V[x(K-2), K-2] = \sqrt{(1+\alpha)[x(K-2) - x_e/\alpha^2]}.$$

Βαθμίδα K–3: Κατά παρόμοιο τρόπο (και μετά από πολλές αλγεβρικές πράξεις) προσδιορίζουμε

$$u(K-3) = \frac{\alpha+\alpha^2}{1+\alpha+\alpha^2} x(K-3) + \frac{1}{\alpha^3}\frac{x_e}{1+\alpha+\alpha^2} \qquad \forall x(K-3) \geq \frac{x_e}{\alpha^3}$$

$$V[x(K-3), K-3] = \sqrt{(1+\alpha+\alpha^2)[x(K-3) - x_e/\alpha^3]}.$$

Κατ' αναλογία μπορούμε να γενικεύσουμε την λύση κάθε βαθμίδας ως εξής (η απόδειξη δι' επαγωγής είναι εύκολη)

$$u(k) = R[x(k),k] = \frac{\alpha+\alpha^2+\ldots+\alpha^{K-k-1}}{1+\alpha+\ldots+\alpha^{K-k-1}} x(k) + \frac{1}{\alpha^{K-k}}\frac{x_e}{1+\alpha+\ldots+\alpha^{K-k-1}} \quad \forall x(k) \geq \frac{x_e}{\alpha^{K-k}}$$

$$V[x(k),k] = \sqrt{(1+\alpha+\ldots+\alpha^{K-k-1})[x(k) - x_e/\alpha^{K-k}]}.$$

Ο Πίνακας 3.2 δίδει τις βέλτιστες τροχιές των μεταβλητών του προβλήματος για K = 10, x_0 = 100, α = 1.1, x_e = 0. Από τον Πίνακα 3.2 και τη γενική αναλυτική λύση του προβλήματος διαπιστώνουμε τα εξής:

- Ο βέλτιστος σχεδιασμός επιχειρεί την αύξηση του κεφαλαίου x(k) με σχετικά χαμηλή κατανάλωση x(k)–u(k) στις πρώτες βαθμίδες του χρονικού ορίζοντα. Η αύξηση αυτή του κεφαλαίου οδηγεί βέβαια σε υψηλότερη κατανάλωση στις μετέπειτα βαθμίδες.
- Η επιρροή του τελικού κεφαλαίου στο βέλτιστο έλεγχο u(k) μειώνεται με την απομάκρυνση από τον τελικό χρόνο K και μηδενίζεται για K→∞.
- Αν ο εισοδηματίας προγραμματίζει για την αιωνιότητα (δηλαδή K→∞), ο βέλτιστος κανόνας δίδει

k	x(k)	u(k)	x(k)−u(k)	$\frac{x(k)-u(k)}{x(k)}$
0	100.00	93.73	6.27	0.063
1	103.10	95.51	7.59	0.074
2	105.06	95.87	9.19	0.087
3	105.46	94.34	11.12	0.105
4	103.78	90.33	13.45	0.130
5	99.36	83.08	16.27	0.164
6	91.39	71.70	19.69	0.215
7	78.87	55.04	23.83	0.302
8	60.55	31.71	28.83	0.476
9	34.89	0.0	34.89	1.000

Πίνακας 3.2: Βέλτιστες τροχιές για το Παραδ. 3.8.

$$u(k) = \overline{R}[x(k)] = \lim_{K\to\infty} R[x(k),k] = x(k)$$

δηλαδή μηδενική κατανάλωση. Στην περίπτωση αυτή η μέγιστη τιμή του κριτηρίου $V[\mathbf{x}(k)]$ είναι άπειρη. Ας σημειωθεί ότι στην περίπτωση $K\to\infty$ οι προαναφερθείσες προϋποθέσεις εκπληρούνται και κατά συνέπεια ο βέλτιστος κανόνας ελέγχου και το βέλτιστο κριτήριο είναι μη μεταβλητά με το χρόνο k.

■

Παράδειγμα 3.9: Θεωρούμε το σύστημα

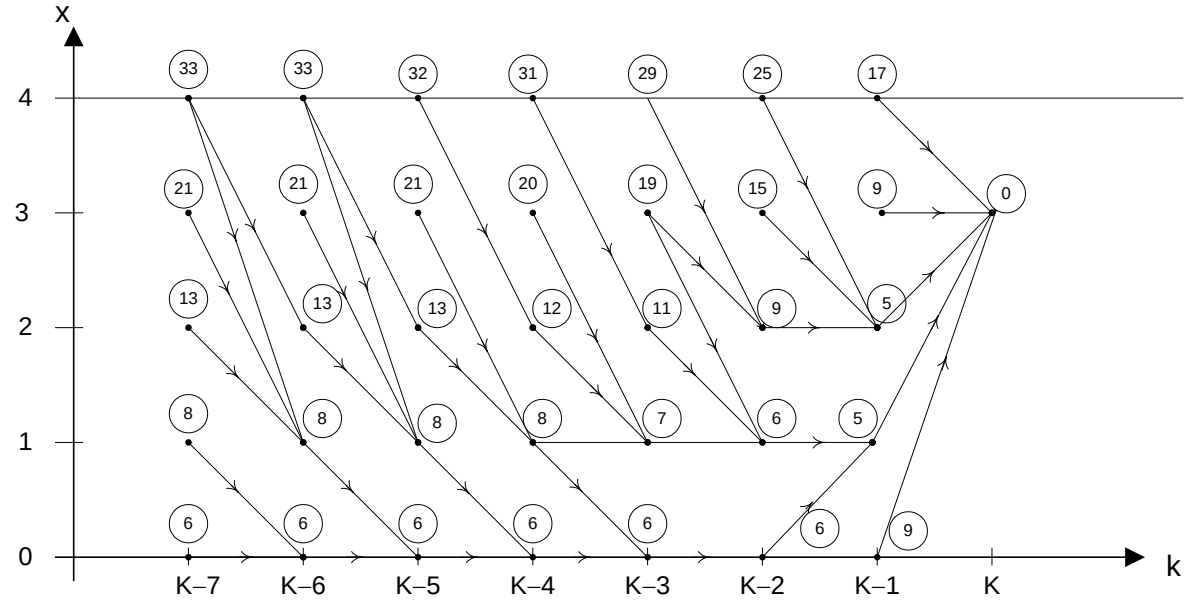
$$x(k+1) = x(k) + u(k)$$

με έλεγχο $u(k)\in\{..., -2, -1, 0, 1, 2,...\}$, επιτρεπτή περιοχή κατάστασης $x(k)\in[0, 4]$ και κριτήριο κόστους

$$J = \sum_{k=0}^{\infty}\left[x(k)^2 + u(k)^2\right].$$

Το Σχ. 3.12 εμπεριέχει τον γραφικό προσδιορισμό του μη μεταβλητού βέλτιστου κανόνα ελέγχου $u(k) = \overline{R}[x(k)]$. (Στο Σχ. 3.12 αναπαριστώνται μόνον οι βέλτιστες μεταβάσεις.) Προς το σκοπό αυτό αρχίζουμε τους υπολογισμούς από έναν τυχαίο τελικό στόχο, π.χ. $x(K) = 3$, και προχωρούμε όπισθεν όπως στο Κεφάλαιο 3.3. Η στάσιμη λύση είναι ανεξάρτητη του επιλεχθέντος τελικού στόχου (βλ. Πρόβλ. 3.1).

Πραγματικά παρατηρούμε ότι, καθώς απομακρυνόμαστε από τον τελικό στόχο, η επίδραση του επιλεχθέντος τελικού σημείου εξασθενεί σταδιακά και τελικά εξαφανίζεται. Παρατηρούμε επίσης, ότι ο βέλτιστος έλεγχος και το ελάχιστο κόστος σε κάθε διακριτό σημείο κατάστασης συγκλίνουν προς σταθερές τιμές που, για $k<K-6$, δεν μεταβάλλονται πλέον με το χρόνο. Ο μη μεταβλητός βέλτιστος έλεγχος του προβλήματος είναι



Σχήμα 3.12: Λύση του Παραδ. 3.9.

x	4	3	2	1	0
u	–3 ή –2	–2	–1	–1	0

■

Ας δούμε τώρα και ένα παράδειγμα ΠΒΕ με ελεύθερο τελικό χρόνο K που οδηγεί σε μη μεταβλητά αποτελέσματα.

Παράδειγμα 3.10: Θεωρούμε το ίδιο δυναμικό σύστημα του Παραδ. 3.7 με τους ίδιους περιορισμούς και την ίδια διακριτοποίηση. Ζητούμε όμως τώρα τον μη μεταβλητό κανόνα ελέγχου που μεταφέρει οποιαδήποτε επιτρεπτή αρχική κατάσταση σε ελάχιστο χρόνο (βλ. (3.30)) στο τελικό σημείο $x(K) = x_e = 0$.

Η λύση του προβλήματος απεικονίζεται στο Σχ. 3.13. Παρατηρούμε στην γραφική επίλυση του Σχ. 3.13 (αναπαριστώνται μόνον οι βέλτιστες μεταβάσεις) ότι ο ελάχιστος χρόνος μεταφοράς και ο βέλτιστος έλεγχος εξαρτώνται από την τιμή x(k) της υπό μεταφοράν κατάστασης, όχι όμως από τον χρόνο k. Ο βέλτιστος κανόνας ελέγχου του Σχ. 3.13 μπορεί να εκφρασθεί ως

$$u(k) = \mathrm{sign}\left[x_e - x(k)\right]$$

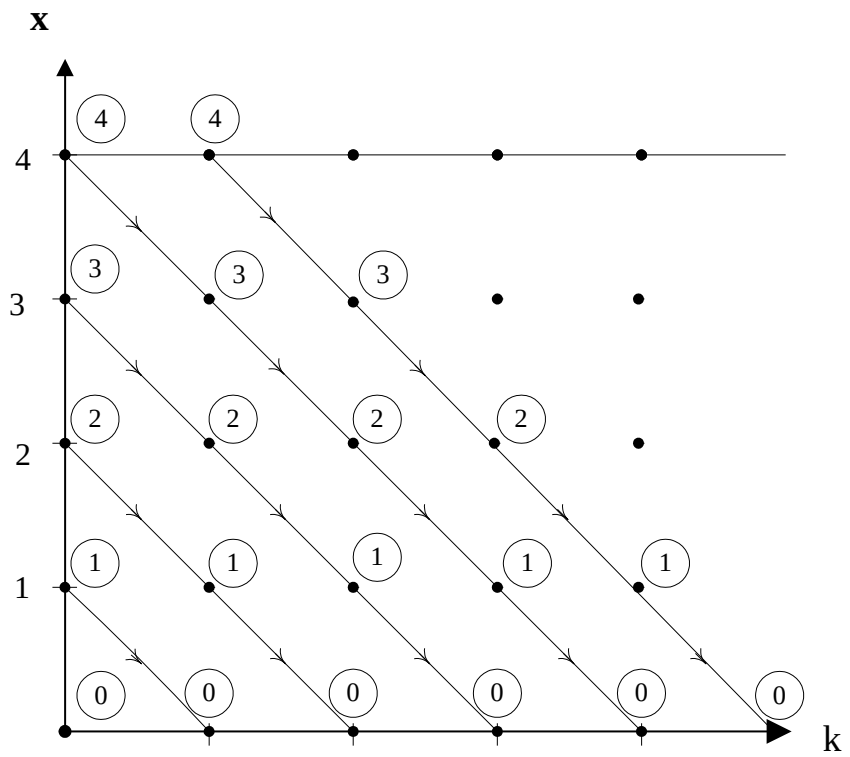
και ο ελάχιστος τελικός χρόνος

$$V\left[x(k)\right] = \left|x_e - x(k)\right|.$$

Η συναρτησιακή εξίσωση (3.32) για το παρόν παράδειγμα είναι (για K>0)

$$V\left[x(k)\right] = \min_{u(k)\in U}\left\{1 + V\left[x(k) + u(k)\right]\right\}.$$

Αντικαθιστώντας την V[x(k)] στη συναρτησιακή εξίσωση έχουμε



Σχήμα 3.13: Λύση του Παραδ. 3.10.

$$|x_e - x| = \min_{u \in U}\left(1 + |x_e - x - u|\right).$$

Η ελαχιστοποίηση της δεξιάς πλευράς δίδει ακριβώς τον προαναφερθέντα βέλτιστο κανόνα ελέγχου. Δι' αντικαταστάσεως στην παραπάνω ισότητα έχουμε

$$|x_e - x| = 1 + |x_e - x - \mathrm{sign}(x_e - x)|$$

η οποία προφανώς ισχύει για K>0, δηλαδή $x \neq x_e$. Άρα η λύση του παραδείγματος ικανοποιεί την συναρτησιακή εξίσωση (3.32).

■

Προβλήματα

3.1 Επιλύστε το πρόβλημα του Παραδ. 3.9 χρησιμοποιώντας διαφορετικούς εικονικούς τελικούς στόχους

a) $x(K) = 4$

b) $x(K) = 0$

c) $x(K) \in \{0,3,4\}$.

Συγκρίνετε την εκάστοτε λύση με αυτήν του Παραδ. 3.9.

3.2 Επιλύστε το πρόβλημα του Παραδ. 3.7 με πανομοιότυπη εξίσωση κατάστασης, περιορισμούς, τελικό στόχο και διακριτοποίηση, επιχειρώντας όμως ελαχιστοποίηση του κριτηρίου κόστους

$$J = \sum_{k=0}^{K-1} \left[1 + \beta \cdot |u(k)|\right] \text{ , K ελεύθερος}$$

για β = 1 και β = 4. Ερμηνεύστε τη σημασία του κριτηρίου και τον ρόλο της παραμέτρου βάρους β.

3.3 Θεωρούμε το δυναμικό σύστημα

$$x(k+1) = x(k) + u(k)$$

με τους περιορισμούς

$$u(k) \in U = \{-3, -2, -1, 0, 1, 2, 3\}$$
$$x(k) \in X = [5, 0]$$

τον τελικό στόχο x(k) = 3 και το κριτήριο κόστους

$$J = \sum_{k=0}^{K-1} \left[1 + 0.2u(k)^2\right], \text{ K ελεύθερος.}$$

Υπολογίστε γραφικά τον μη μεταβλητό κανόνα βέλτιστου ελέγχου και την μη μεταβλητή συνάρτηση ελάχιστου κόστους. Αποδείξτε ότι η λύση ικανοποιεί την συναρτησιακή εξίσωση του Bellman.

3.4 Θεωρούμε το εξής ΠΒΕ:

Ελαχιστοποίηση του κριτηρίου κόστους

$$J = \frac{1}{2} \sum_{k=0}^{2} u(k)^2$$

λαμβάνοντας υπ' όψη

$$x(k+1) = x(k) + u(k), \quad x(0) = 2, x(3) = 0.$$

Κατασκευάστε διακριτά πλέγματα για την μεταβλητή κατάστασης με διαστήματα

α) Δx = 1, Δu = 1

β) Δx = 2/3, Δu = 2/3.

Περιορίστε με κατάλληλο τρόπο τον αριθμό των αντίστοιχων μεταβάσεων. Υπολογίστε για τις περιπτώσεις α) και β) την λύση του ΠΒΕ μέσω Διακριτού Δυναμικού Προγραμματισμού και συγκρίνετε τις αντίστοιχες ελάχιστες τιμές του κριτηρίου κόστους καθώς και τους κανόνες ελέγχου.

3.5 Μια βιοτεχνία καταναλώνει κάθε μήνα 3 τόνους καυσίμων. Η προμήθεια καυσίμων σε ποσότητες (τόνους) u(k) ∈ {0, 1, 2, 3, 4, 5} γίνεται στην αρχή κάθε μήνα k και το κόστος προμήθειας είναι $\varphi_2[u(k)] = u(k) \cdot [10 - u(k)]$. Σε περίπτωση περισσεύματος στο τέλος ενός μήνα υπάρχει δυνατότητα αποθήκευσης μέχρι την μέγιστη χωρητικότητα X(k) = 5–k. Η αποθήκευση ποσότητας x(k) για κάθε μήνα k κοστίζει $\varphi_1[x(k), k] = k \cdot x(k)$. Ζητείται ο βέλτιστος προγραμματισμός u(k), k=0,...,4, που ελαχιστοποιεί το ολικό κόστος

προμήθειας και αποθήκευσης καυσίμων $\varphi_1+\varphi_2$ για μια περίοδο 5 μηνών, όπου τα αποθηκευμένα καύσιμα στην αρχή και το τέλος της περιόδου είναι μηδενικά.

(a) Διατυπώστε το πρόβλημα βελτιστοποίησης με το κριτήριο κόστους, την καταστατική εξίσωση και όλους τους περιορισμούς.

(b) Σχεδιάστε σε αντίστοιχο διάγραμμα όλες της επιτρεπτές μεταβάσεις (προσοχή: x(0) = x(5) = 0 !) και αναγράψτε τα αντίστοιχα κόστη.

(c) Επιλύστε γραφικά το πρόβλημα προσδιορίζοντας την βέλτιστη πολιτική προμηθειών u(k), k = 0,...,4, και το ελάχιστο ολικό κόστος.

3.6 Δίδεται το δυναμικό σύστημα

$$x(k+1) = x(k) + u(k)$$

με έλεγχο $u(k) \in \{-1, 0\}$. Υπολογίστε γραφικά σε αντίστοιχο διάγραμμα τον βέλτιστο κανόνα ελέγχου u(k) = R[x(k), k] για $x(k) \in \{0,1,2,3,4\}$ και $k=0,1,2,3$, προς ελαχιστοποίηση του κριτηρίου κόστους

$$J = x(4)^2 + \sum_{k=0}^{3} k \cdot u(k)^2.$$

3.7 Μια αντιπροσωπία ηλεκτρικών ειδών θέλει να μεταφέρει αεροπορικώς κάποια είδη της από Αθήνα στον τοπικό αντιπρόσωπό της στα Χανιά. Το φορτίο θα αποτελείται από u_0 βίντεο, u_1 στερεοφωνικά, u_2 τηλεοράσεις και u_3 ραδιοκασετόφωνα. Το βάρος b_i και η αξία w_i ανά είδος προς φόρτωση είναι $b_0 = 2$, $b_1 = 5$, $b_2 = 4$, $b_3 = 1$ και $w_0 = 10$, $w_1 = 15$, $w_2 = 7$, $w_3 = 5$ μονάδες αντίστοιχα. Δεδομένου ότι το αεροπλάνο θα φέρει κι άλλα φορτία, η αεροπορική εταιρεία έθεσε τους εξής όρους στην αντιπροσωπία:

- Η μέγιστη ποσότητα φόρτωσης για τα είδη βίντεο, στερεοφωνικά και τηλεοράσεις είναι 2 κομμάτια/είδος.
- Η μέγιστη ποσότητα φόρτωσης για το είδος ραδιοκασετόφωνα είναι ελεύθερη.
- Το ολικό βάρος του φορτίου δεν πρέπει να υπερβαίνει τις 25 μονάδες βάρους.

Από την πλευρά της η αντιπροσωπία, κάνοντας δεκτούς τους όρους της αεροπορικής εταιρείας, έθεσε επιπλέον τον περιορισμό, το φορτίο να περιλαμβάνει τουλάχιστον 1 κομμάτι από κάθε είδος.

Δεδομένων των άνω, η αντιπροσωπία ζητά να προσδιορίσει την κατανομή (ή κατανομές) του φορτίου που μεγιστοποιεί τη συνολική αξία του.

(a) Διατυπώστε το πρόβλημα της αντιπροσωπίας σαν πρόβλημα βέλτιστου ελέγχου με το κριτήριο αξίας του φορτίου (κριτήριο κόστους), την καταστατική εξίσωση και όλους του περιορισμούς.

(b) Σχεδιάστε σε αντίστοιχο διάγραμμα όλες τις επιτρεπτές μεταβάσεις και αναγράψτε τα αντίστοιχα κόστη.

(c) Επιλύστε γραφικά το πρόβλημα προσδιορίζοντας τη βέλτιστη κατανομή (ή κατανομές) του φορτίου u(k), k = 0,...,3, και τη μέγιστη ολική αξία του.

3.8 Μια μικρή αερολέσχη μισθώνει δύο επαγγελματίες πιλότους, τους οποίους χρησιμοποιεί ως εκπαιδευτές ερασιτεχνών πιλότων έναντι αμοιβής. Η αερολέσχη έχει οργανώσει δύο επίπεδα εκπαίδευσης. Το πρώτο απευθύνεται σε αρχάριους και το έχει αναθέσει στον 1ο πιλότο, ενώ το δεύτερο απευθύνεται σε προχωρημένους και το έχει αναθέσει στον 2ο πιλότο. Για τα μαθήματα η αερολέσχη έχει διαθέσει ένα μόνο από τα αεροπλάνα της το οποίο οι δύο εκπαιδευτές πιλότοι χρησιμοποιούν εναλλακτικά. Το κάθε μάθημα διαρκεί μία ώρα και ο κανονισμός λειτουργίας της αερολέσχης ορίζει ότι μαθήματα παραδίδονται συνεχώς επί 6 ώρες ημερησίως. Στην αρχή της κάθε ώρας η διεύθυνση της αερολέσχης αποφασίζει ποιός εκπαιδευτής, θα κάνει μάθημα. Η απόφαση αυτή περιγράφεται από τη μεταβλητή u που ορίζεται ως εξής:

$$u(k) = \begin{cases} 1 \text{ αν τη χρονική στιγμή k αποφασιστεί να κάνει μάθημα ο 1ος εκπαιδευτής} \\ 0 \text{ αν τη χρονική στιγμή k αποφασιστεί να κάνει μάθημα ο 2ος εκπαιδευτής} \end{cases}$$

Επίσης, με $x_i(k)$, i = 1,2, (όπου $x_i(k) \geq 0$) συμβολίζονται οι ώρες μαθήματος που έχει κάνει μέχρι τη χρονική στιγμή k χωρίς διακοπή ο πιλότος i, για τις οποίες τα συμβόλαια που έχει υπογράψει η αερολέσχη με τους εκπαιδευτές πιλότους ορίζουν ότι μπορούν να είναι το πολύ 2 ώρες όσον αφορά τον 1ο πιλότο και 3 ώρες όσον αφορά τον 2ο. Αν για παράδειγμα ο 1ος πιλότος μέχρι την χρονική στιγμή k έχει κάνει 2 ώρες μαθήματος χωρίς διακοπή (δηλαδή $x_1(k) = 2$), την χρονική στιγμή k+1 δεν μπορεί να παραδώσει μάθημα.

Η αερολέσχη εισπράτει από κάθε μάθημα αρχαρίων 2.000 € και από κάθε μάθημα προχωρημένων 2.200 €, ενώ ο λογιστής της έχει υπολογίσει τα έξοδά της ως εξής:

a) *αμοιβές εκπαιδευτών*
$\varphi_1[u(k)] = 30\cdot u(k) + 45\cdot[1-u(k)]$ (σε €)

b) *ειδικά επιδόματα εκπαιδευτών*
$\varphi_2[\mathbf{x}(k),u(k)] = 30\cdot\{0.1\cdot x_1(k)\cdot u(k) + 0.5\cdot x_2(k)\cdot [1-u(k)]\}$ (σε €)

c) *έξοδα συντήρησης και ανεφοδιασμού*
$\varphi_3[u(k)] = 60\cdot\{0.05\cdot u(k) + 0.1\cdot[1-u(k)]\}$ (σε €).

Στην αρχή της κάθε εργάσιμης ημέρας κανένας εκπαιδευτής δεν έχει παραδώσει μάθημα, δηλαδή $x_1(0) = x_2(0) = 0$. Ζητείται η βέλτιστη πολιτική της αερολέσχης που θα μεγιστοποιήσει τα ημερήσια κέρδη της (έσοδα μείον έξοδα) από την παράδοση μαθημάτων αεροπλοήγησης υπό την παραδοχή ότι υπάρχει πάντα ζήτηση για μαθήματα.

(a) Διατυπώστε το πρόβλημα βελτιστοποίησης με την καταστατική εξίσωση, το αντικειμενικό κριτήριο των ημερήσιων κερδών της αερολέσχης και όλους τους περιορισμούς.

(b) Πόσες και ποιες διακριτές δυνατότητες αποφάσεων έχει η διεύθυνση της αερολέσχης στην αρχή του κάθε ωριαίου μαθήματος; Πόσες και ποιες επιτρεπτές τιμές έχει το διάνυσμα κατάστασης **x**;

(c) Έστω $\mathbf{X}_1$, $\mathbf{X}_2$, $\mathbf{X}_3$,… οι διακριτές επιτρεπτές τιμές κατάστασης, π.χ. $\mathbf{X}_1 = (0,0)$, $\mathbf{X}_2 = (0,1)$, $\mathbf{X}_3 = (0,2)$,… Αναγράψτε σε αντίστοιχο διάγραμμα όλες τις διακριτές επιτρεπτές τιμές και σχεδιάστε όλες τις επιτρεπτές μεταβάσεις με τα αντίστοιχα κόστη.

(d) Επιλύστε γραφικά το πρόβλημα, προσδιορίζοντας τη βέλτιστη πολιτική (ή τις βέλτιστες πολιτικές) ημερήσιων μαθημάτων και το μέγιστο συνολικό ημερήσιο κέρδος της αερολέσχης.

3.9 Ένα πλοίο ξεκινά από ένα λιμάνι A με προορισμό ένα λιμάνι B στο οποίο πρέπει να φτάσει μετά από K χρονικές περιόδους (βλέπε το παρακάτω σχήμα). Στην πορεία του είναι γνωστό ότι θα συναντήσει μια περιοχή ανέμων, η οποία είναι αναπτυγμένη σε N νοητές ζώνες έντασης γύρω από ένα γνωστό σημείο C με συντεταγμένες (c,c).

Η θέση του πλοίου κάθε χρονική περίοδο k, όπου k = 0, 1,...,K–1, περιγράφεται από την ακόλουθη καταστατική εξίσωση

$$\mathbf{x}(k+1) = \mathbf{x}(k)+\mathbf{u}(k)-\mathbf{d}(\mathbf{x}(k))$$

όπου:

- $\mathbf{x}(k)^T = [x_1(k)\ x_2(k)]$ το διάνυσμα θέσης του πλοίου με $x_i(k) \in \{0,1,...,X_i\}$, i = 1,2
- $\mathbf{u}(k)^T = [u_1(k)\ u_2(k)]$ το διάνυσμα μεταβολής της θέσης του πλοίου σαν αποτέλεσμα αντίστοιχης ενεργοποίησης του κινητήρα του με $u_i(k) \in \{0,1,...,U\}$, i = 1,2
- $\mathbf{d}(\mathbf{x}(k))^T = [d_1(\mathbf{x}(k))\ d_2(\mathbf{x}(k))]$ το διάνυσμα μεταβολής θέσης του πλοίου σαν αποτέλεσμα της αντίστασης του αέρα και των κυμάτων όπου

$$d_1(\mathbf{x}(k)) = \begin{cases} \pm(D+1) & \text{αν } \{x_1(k)\in[c-N,c+N] \text{ και } x_2(k) = c \mu N\} \\ \pm(D+2) & \text{αν } \{x_1(k)\in[c-(N-1),c+(N-1)] \text{ και } x_2(k) = c \mu (N-1)\} \\ & \dots \\ \pm(D+N) & \text{αν } \{x_1(k)\in[c-1,c+1] \text{ και } x_2(k) = c \mu 1\} \\ 0 & \text{αν } x_1(k) = x_2(k) = c \\ D & \text{άλλως} \end{cases}$$

$$d_2(\mathbf{x}(k)) = \begin{cases} \pm(D+1) & \text{αν } \{x_2(k)\in[c-N,c+N] \text{ και } x_1(k) = c \pm N\} \\ \pm(D+2) & \text{αν } \{x_2(k)\in[c-(N-1),c+(N-1)] \text{ και } x_1(k) = c \pm (N-1)\} \\ & \dots \\ \pm(D+N) & \text{αν } \{x_2(k)\in[c-1,c+1] \text{ και } x_1(k) = c \pm 1\} \\ 0 & \text{αν } x_2(k) = x_1(k) = c \\ D & \text{άλλως} \end{cases}$$

με D γνωστή σταθερά που αντιπροσωπεύει την αντίσταση του αέρα και των κυμάτων σε κανονικές καιρικές συνθήκες.

Επίσης πρέπει πάντα

$$u_i(k) - d_i(\mathbf{x}(k)) > 0, \quad i = 1,2$$

δηλαδή το πλοίο δεν επιτρέπεται να ακινητοποιηθεί ή να οπισθοδρομήσει.

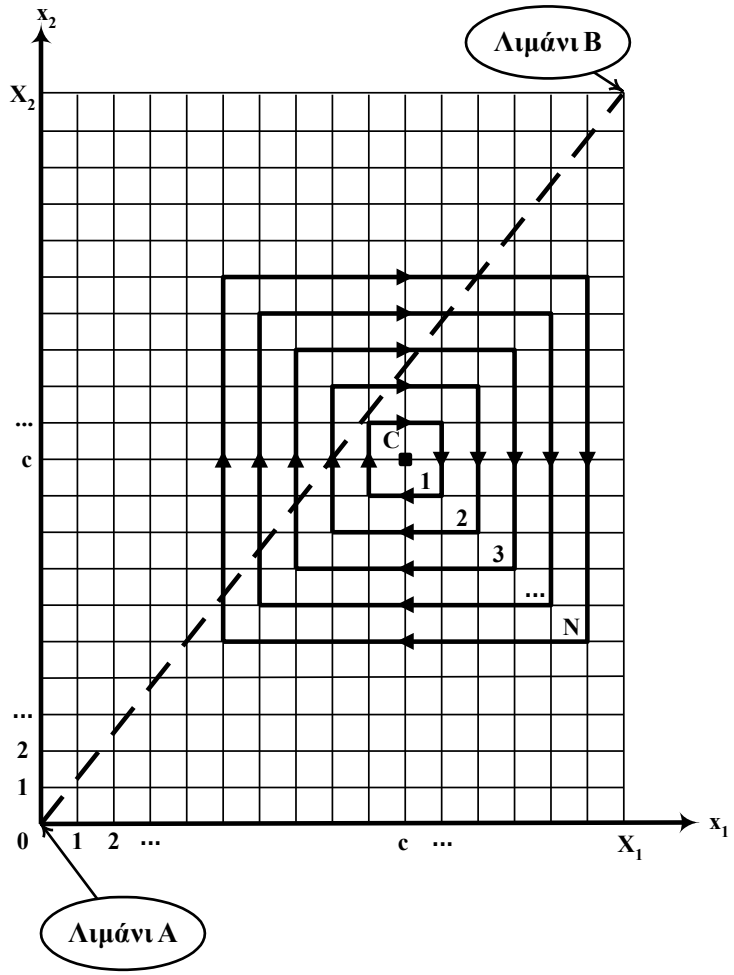
Το αντικειμενικό κριτήριο προς ελαχιστοποίηση δίδεται από τη σχέση

$$J = \sum_{k=0}^{K-1} \left[u_1(k)^2 + u_2(k)^2 \right]$$

και αντιπροσωπεύει το κόστος καυσίμων.

(a) Διατυπώστε το πρόβλημα καθορισμού της πορείας του πλοίου σαν πρόβλημα βελτιστοποίησης με την καταστατική εξίσωση, το κριτήριο κόστους και όλους τους περιορισμούς.

(b) Να αναπτυχθεί ένα διάγραμμα ροής του προγράμματος που να επιλύει το παραπάνω πρόβλημα. Το πρόγραμμα θα πρέπει να δέχεται ως δεδομένα εισόδου τις τιμές K, X_1, X_2, U, D, c, N και να υπολογίζει τον βέλτιστο κανόνα ελέγχου και τη βέλτιστη τροχιά **u**(k) για τη μεταφορά του πλοίου από το λιμάνι A στο λιμάνι B.



3.10 Θεωρούμε το δυναμικό σύστημα

$$x(k+1)=x(k)+u(k)\ ,\ x(0) = 0$$

με τους περιορισμούς

$$u(k) \in U = \{1, 0\}$$
$$x(k) \in X = \{0, 1, 2, 3\}.$$

Ζητείται η βέλτιστη τροχιά ελέγχου για τη μεταφορά της αρχικής κατάστασης σε οποιοδήποτε από τα τελικά σημεία

$$[K, x(K)] = \{(3,3),(4,2),(4,1)\},$$

ελαχιστοποιώντας το αντικειμενικό κριτήριο

$$J = K + \sum_{k=0}^{K-1} k u(k)^2 .$$

Προσδιορίστε πάνω σε αντίστοιχο σχήμα όλες τις δυνατές μεταβάσεις από μία βαθμίδα στην επόμενή της καθώς και το βέλτιστο κανόνα ελέγχου του προβλήματος.

3.11 Η θερμοκρασία *t* ενός δωματίου μεταβάλλεται χρονικά σύμφωνα με την παρακάτω σχέση

$$t(k+1) = 0.5 \cdot t(k) + 0.5 \cdot d(k) + u(k)$$

όπου d είναι η εξωτερική θερμοκρασία και u είναι η επιρροή του κλιματιστικού εντός του δωματίου. Επειδή στο δωμάτιο φυλάσσονται ηλεκτρονικές συσκευές, η θερμοκρασία του πρέπει να παραμένει μεταξύ 19°C και 21°C. Η εξωτερική θερμοκρασία d(k) μεταβάλλεται κατά τη διάρκεια μίας ημέρας σύμφωνα με τον Πίνακα της επόμενης σελίδας, ενώ το κλιματιστικό που χρησιμοποιείται έχει δυνατότητα μεταβολής της θερμοκρασίας εντός της περιοχής $u(k) \in [-0.5^0\text{C}, 0.5^0\text{C}]$. Προσδιορίστε το βέλτιστο κανόνα ελέγχου για το u(k) του κλιματιστικού, για K=3, έτσι ώστε στην αρχή και στο τέλος της ημέρας η θερμοκρασία του δωματίου να είναι 20 ^{o}C, δηλαδή t(0)=t(3)=20 ^{o}C. Ως αντικειμενική συνάρτηση προς ελαχιστοποίηση θεωρήστε την

$$J = \sum_{k=0}^{2} \left[[t(k) - 20]^2 + u(k)^2 \right].$$

(a) Διατυπώστε το παραπάνω πρόβλημα με την αντικειμενική συνάρτηση, την καταστατική εξίσωση και όλους τους περιορισμούς.

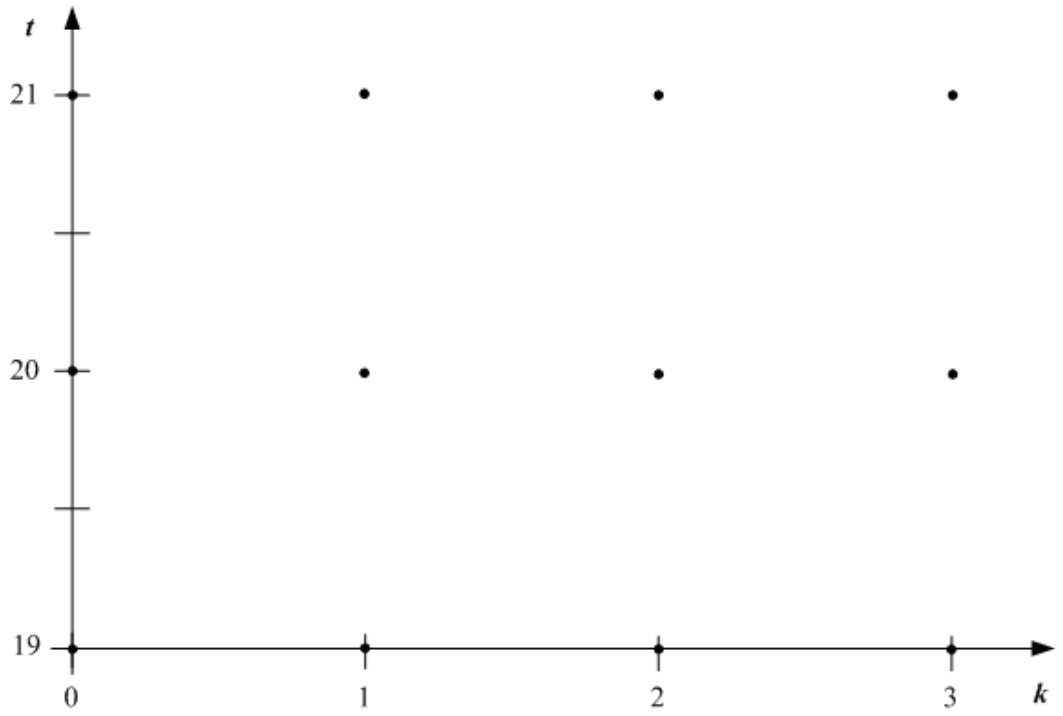
(b) Θεωρούμε τα διακριτά διαστήματα Δt=1 και Δu=0.5. Προσδιορίστε στο παρακάτω σχήμα όλες τις επιτρεπτές μεταβάσεις από κάθε διακριτή κατάσταση καθώς και το βέλτιστο κανόνα ελέγχου του προβλήματος, μέσω του αλγορίθμου Δυναμικού Προγραμματισμού.

Υπόδειξη 1: Μεταβάσεις που οδηγούν σε μη διακριτές τιμές της μεταβλητής κατάστασης γίνονται δεκτές.
Υπόδειξη 2: Για τον υπολογισμό του κόστους μετάβασης V[t(k),k] σε σημεία t(k) που δεν είναι διακριτά, χρησιμοποιήστε γραμμική παρεμβολή.

Υπόδειξη 3: Για τον υπολογισμό της βέλτιστης τροχιάς ελέγχου u(k), σε περίπτωση που αυτή διέρχεται από μη διακριτά σημεία t(k), χρησιμοποιείστε γραμμική παρεμβολή.

Πίνακας: Εξωτερική θερμοκρασία δωματίου.

k	**k=0 (πρωί)**	**k=1 (μεσημέρι)**	**k=2 (βράδυ)**
d(k)	19	22	20



3.12 Θεωρούμε το ΠΒΕ διακριτού χρόνου που αφορά την ελαχιστοποίηση του κριτηρίου κόστους

$$J = \sum_{k=0}^{2} k \cdot u(k)^2$$

υπό τους περιορισμούς

$$x(k+1) = x(k) + u(k)\,,\ \ x(3) = 2\,,$$

για οποιαδήποτε (τυχαία) αρχική κατάσταση $x(0)$.

(a) Υπολογίστε μέσω της διαδικασίας Δυναμικού Προγραμματισμού (χωρίς διακριτοποίηση) το βέλτιστο κανόνα ελέγχου $u(k)$ που μεταφέρει κάθε κατάσταση $x(k) \in \mathbf{R},\ k = 0,1,2$, στον τελικό στόχο.

(b) Ποιες είναι οι βέλτιστες τροχιές ελέγχου και κατάστασης για τη μεταφορά της $x(0) = -20$ στη $x(3) = 2$.

4. ΒΕΛΤΙΣΤΟΣ ΕΛΕΓΧΟΣ ΣΥΝΕΧΟΥΣ ΧΡΟΝΟΥ

Το κεφάλαιο αυτό αναφέρεται στην εφαρμογή της Αρχής του Βέλτιστου στο πρόβλημα βέλτιστου ελέγχου συνεχούς χρόνου (βλ. Παραρτ. Β.1).

4.1 Διατύπωση του Προβλήματος

Θεωρούμε την ελαχιστοποίηση του κριτηρίου κόστους

$$J = \vartheta[\mathbf{x}(t_e), t_e] + \int_0^{t_e} \varphi(\mathbf{x}, \mathbf{u}, t)\,dt \tag{4.1}$$

με τελικό χρόνο t_e ελεύθερο ή καθορισμένο, λαμβάνοντας υπόψη την καταστατική εξίσωση

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u}, t) \tag{4.2}$$

με την αρχική κατάσταση $\mathbf{x}(0) = \mathbf{x}_0$ και τον τελικό στόχο

$$\mathbf{g}\,[\mathbf{x}(t_e), t_e] = \mathbf{0}. \tag{4.3}$$

Εννοείται ότι όλα τα αποτελέσματα του κεφαλαίου αυτού είναι εφαρμόσιμα στην ειδική περίπτωση απουσίας ενός τελικού στόχου. Η επιτρεπτή περιοχή ελέγχου ορίζεται

$$\mathbf{u}(t) \in \mathcal{U}[\mathbf{x}(t), t] = \{\mathbf{u}(t) \,|\, \mathbf{h}(\mathbf{x}, \mathbf{u}, t) \le \mathbf{0}\} \tag{4.4}$$

όπου υποθέτουμε ότι ο Ιακωβιανός πίνακας $\partial\mathbf{h}/\partial\mathbf{u}$ είναι πλήρους βαθμού. Περιορισμοί κατάστασης δεν θα ληφθούν υπόψη στα πλαίσια του κεφαλαίου αυτού. Τέλος, προϋποθέτουμε την ύπαρξη απόλυτου ελάχιστου του διατυπωθέντος προβλήματος.

Η αριθμητική επίλυση του ΠΒΕ συνεχούς χρόνου μέσω Δυναμικού Προγραμματισμού είναι δυνατή μετά από διακριτοποίηση του προβλήματος ως προς το χρόνο (βλ. π.χ. την (Α.7)) και εφαρμογή των μεθόδων του Κεφαλαίου 3. Στο παρόν κεφάλαιο θα περιορισθούμε σε μερικά θεωρητικά αποτελέσματα της εφαρμογής της Αρχής του Βέλτιστου στο ΠΒΕ συνεχούς χρόνου. Τα αποτελέσματα αυτά, πέρα από την θεωρητική τους σημασία, οδηγούν στην αναλυτική λύση σχετικά απλών ΠΒΕ.

4.2 Η Εξίσωση Hamilton-Jacobi-Bellman

Όπως στο Κεφάλαιο 3.2, ορίζουμε καταρχήν το **υπολειπόμενο κόστος** ή **κόστος μεταφοράς** J_t για τη μεταφορά μιας κατάστασης $\mathbf{x}(t)$ στον τελικό στόχο (4.3) ως εξής (προσοχή στο κάτω όριο του ολοκληρώματος)

$$J_t = [\mathbf{x}(t_e), t_e] + \int_t^{t_e} \varphi(\mathbf{x}, \mathbf{u}, \tau)\,d\tau. \tag{4.5}$$

Όπως και στο ΠΒΕ διακριτού χρόνου, το ελάχιστο κόστος μεταφοράς $J_t^* = \min J_t$ (ικανοποιώντας και όλους τους περιορισμούς) για ένα συγκεκριμένο πρόβλημα εξαρτάται αποκλειστικά από την υπό μεταφορά κατάσταση $\mathbf{x}(t)$ και από την χρονική στιγμή t. Ονομάζουμε αυτό το ελάχιστο κόστος $V(\mathbf{x}, t)$ και έχουμε για κάποιο t', $t < t' < t_e$,

$$V(\mathbf{x}, t) = \min J_t = \min \left[\int_t^{t'} \varphi(\mathbf{x}, \mathbf{u}, \tau)\, d\tau + J_{t'} \right] \qquad (4.6)$$

όπου το ελάχιστο νοείται μεταξύ όλων των τροχιών $\mathbf{u}(\tau)$, $\tau \in [t, t_e]$, που ικανοποιούν τις (4.2)-(4.4). Εφαρμόζοντας την Αρχή του Βέλτιστου στην (4.6) έχουμε

$$V(\mathbf{x}, t) = \min \left[\int_t^{t'} \varphi(\mathbf{x}, \mathbf{u}, \tau)\, d\tau + V(\mathbf{x}', t') \right] \qquad (4.7)$$

όπου $\mathbf{x}' = \mathbf{x}(t')$, και η ελαχιστοποίηση νοείται τώρα μόνο για την επιτρεπτή μετάβαση από την $\mathbf{x}(t)$ στην $\mathbf{x}(t')$, δηλαδή για τον έλεγχο $\mathbf{u}(\tau)$, $\tau \in [t, t']$.

Στο σημείο αυτό, για να μπορέσουμε να συνεχίσουμε την επεξεργασία του προβλήματος, πρέπει να προϋποθέσουμε ότι η συνάρτηση $V(\mathbf{x},t)$ είναι συνεχής και παραγωγίσιμη. Η παραδοχή αυτή δεν ισχύει σε κάποια ΠΒΕ, πράγμα που βέβαια περιορίζει καταρχήν το πεδίο εφαρμογής των παρακάτω αποτελεσμάτων. Υπάρχει όμως παρ' όλα αυτά η δυνατότητα εφαρμογής της θεωρίας του Δυναμικού Προγραμματισμού αν ο χώρος $(\mathbf{x}, t)$ θεωρηθεί χωριζόμενος σε επί μέρους περιοχές, εντός των οποίων η συνάρτηση $V(\mathbf{x}, t)$ είναι συνεχής και παραγωγίσιμη.

Θεωρώντας λοιπόν την $V(\mathbf{x}, t)$ συνεχή και παραγωγίσιμη, μπορούμε να προβούμε με $\Delta t = t' - t$ στον ακόλουθο ανάπτυγμα Taylor

$$V(\mathbf{x}', t') = V(\mathbf{x}, t) + \frac{\partial V(\mathbf{x}, t)^T}{\partial \mathbf{x}} \dot{\mathbf{x}}(t)\, \Delta t + \frac{\partial V(\mathbf{x}, t)}{\partial t} \Delta t + R(\Delta t^2). \qquad (4.8)$$

Επί πλέον ισχύει κατά γραμμική προσέγγιση

$$\int_t^{t'} \varphi(\mathbf{x}, \mathbf{u}, t) d\tau = \varphi(\mathbf{x}, \mathbf{u}, t) \cdot \Delta t. \qquad (4.9)$$

Αντικαθιστώντας τις (4.2), (4.8), (4.9) στην (4.7) έχουμε, παραλείποντας τα υπόλοιπα υψηλών βαθμών $R(\Delta t^2)$,

$$V(\mathbf{x}, t) = \min\left[\varphi(\mathbf{x}, \mathbf{u}, t)\Delta t + V(\mathbf{x}, t) + \frac{\partial V(\mathbf{x}, t)^T}{\partial \mathbf{x}} \mathbf{f}(\mathbf{x}, \mathbf{u}, t)\Delta t + \frac{\partial V(\mathbf{x}, t)}{\partial t} \Delta t \right]. \qquad (4.10)$$

Παρατηρούμε τώρα ότι η $V(\mathbf{x},t)$ είναι ανεξάρτητη του $\mathbf{u}(\tau)$, $\tau \in [t,t']$, και επομένως μπορεί να εξαχθεί από τον υπό ελαχιστοποίηση όρο και να απλοποιηθεί με την αριστερή πλευρά της ισότητας (4.10). Για τον ίδιο λόγο μπορεί να εξαχθεί από την ελαχιστοποίηση και ο όρος $\partial V(\mathbf{x},t)/\partial t \cdot \Delta t$. Εν συνεχεία απλοποιείται ο Δt από όλους τους όρους της ισότητας. Σαν τελική

παρατήρηση, ο υπό ελαχιστοποίηση όρος εξαρτάται από τον έλεγχο $\mathbf{u}(t)$ μόνο της χρονικής στιγμής t και επομένως η ελαχιστοποίηση νοείται μεταξύ των $\mathbf{u}(t) \in \mathcal{U}(\mathbf{x},t)$. Εν κατακλείδι έχουμε από την (4.10)

$$0 = \frac{\partial V(\mathbf{x},t)}{\partial t} + \min_{\mathbf{u} \in \mathcal{U}(\mathbf{x},t)} [\varphi(\mathbf{x},\mathbf{u},t) + \frac{\partial V(\mathbf{x},t)}{\partial \mathbf{x}}^T \mathbf{f}(\mathbf{x},\mathbf{u},t)]. \qquad (4.11)$$

Αυτή η μη γραμμική μερική διαφορική εξίσωση πρώτου βαθμού είναι μια γενίκευση της από την Φυσική γνωστής εξίσωσης Hamilton-Jacobi, και φέρει στην θεωρία βελτιστοποίησης το όνομα **εξίσωση Hamilton-Jacobi-Bellman** (αν και είχε δημοσιευθεί ήδη από το 1935 από τον **Κωνσταντίνο Καραθεοδωρή**). Για την επίλυση της εξίσωσης ως προς τη συνάρτηση $V(\mathbf{x},t)$ απαιτείται μια συνοριακή τιμή που προφανώς είναι

$$V[\mathbf{x}(t_e), t_e] = \vartheta[\mathbf{x}(t_e), t_e] \qquad (4.12)$$

για όλα τα $(\mathbf{x}(t_e), t_e)$ που ικανοποιούν την (4.3). Εισάγοντας τους ορισμούς

$$H(\mathbf{x},\mathbf{u},\frac{\partial V}{\partial \mathbf{x}},t) = \varphi(\mathbf{x},\mathbf{u},t) + \frac{\partial V(\mathbf{x},t)}{\partial \mathbf{x}}^T f(\mathbf{x},\mathbf{u},t) \qquad (4.13)$$

$$H^*(\mathbf{x},\frac{\partial V}{\partial \mathbf{x}},t) = \min_{\mathbf{u} \in \mathcal{U}(\mathbf{x},t)} H(\mathbf{x},\mathbf{u},\frac{\partial V}{\partial \mathbf{x}},t) \qquad (4.14)$$

έχουμε την ακόλουθη εναλλακτική μορφή της εξίσωσης Hamilton-Jacobi-Bellman

$$\frac{\partial V}{\partial t} = -H^*(\mathbf{x},\frac{\partial V}{\partial \mathbf{x}},t) \ . \qquad (4.15)$$

Ας σημειωθεί ότι η συνάρτηση H φέρει το όνομα **συνάρτηση Hamilton** στην θεωρία του Λογισμού Μεταβολών (βλ. Παραρτ. B2).

Όπως στο Κεφάλαιο 3.4γ, το ελάχιστο κόστος μεταφοράς $V(\mathbf{x}, t)$ είναι ανεξάρτητο του χρόνου t, αν ο τελικός χρόνος t_e είναι ελεύθερος ή άπειρος, οι συναρτήσεις του προβλήματος φ, **f**, **h**, **g** είναι ανεξάρτητες του χρόνου και η συνάρτηση φ είναι θετικά ημιορισμένη. Στην περίπτωση αυτή έχουμε επομένως $\partial V/\partial t = 0$ και η (4.15) απλοποιείται ως εξής

$$H^*(\mathbf{x},\frac{\partial V}{\partial \mathbf{x}},t) = 0 \quad \forall t \in [0, t_e]. \qquad (4.16)$$

Η εξίσωση Hamilton-Jacobi-Bellman αποτελεί μια **ικανή και αναγκαία συνθήκη** για ένα **απόλυτο ελάχιστο** του προβλήματος υπό την προϋπόθεση ότι η ελαχιστοποίηση (4.14) της συνάρτησης Hamilton αποδίδει για κάθε $t \in [0,t_e]$ ένα μοναδικό απόλυτο ελάχιστο.

Η σημασία της εξίσωσης Hamilton-Jacobi-Bellman περιορίζεται από το γεγονός ότι η αναλυτική επίλυση αυτής της μη γραμμικής μερικής διαφορικής εξίσωσης δεν είναι δυνατή παρά για απλά προβλήματα. Στις περιπτώσεις όμως επίλυσής της, έχουμε σαν άμεσο αποτέλεσμα έναν **βέλτιστο κανόνα ελέγχου**. Αυτό γίνεται φανερό αν παρατηρήσουμε

καταρχήν ότι η αναλυτική ελαχιστοποίηση (4.14) της συνάρτησης Hamilton έχει σαν αποτέλεσμα

$$\mathbf{u}^*(t) = \mathbf{R}(\mathbf{x}, \frac{\partial V}{\partial \mathbf{x}}, t). \qquad (4.17)$$

Αν, λοιπόν, η V(**x**,t) είναι γνωστή (μετά από αντικατάσταση του u* και αναλυτική λύση της (4.11) ή της (4.15)) τότε μπορεί η V(**x**, t) να αντικατασταθεί με την σειρά της στην (4.17) δίνοντας **u***(t) = **R**(**x**, t), δηλαδή πράγματι έναν βέλτιστο κανόνα ελέγχου.

Παράδειγμα 4.1: Θεωρούμε το ακόλουθο ΠΒΕ:

Ελαχιστοποίηση

$$J = \frac{1}{2}\int_0^\infty (x^2+u^2)dt$$

λαμβάνοντας υπόψη

$$\dot{x} = x^3+u \ , \quad x(0) = x_0.$$

Για το πρόβλημα αυτό μπορούμε να επιλύσουμε αναλυτικά την εξίσωση Hamilton-Jacobi-Bellman καταλήγοντας σ' ένα βέλτιστο κανόνα ελέγχου ως εξής. Η συνάρτηση Hamilton (4.13) για το παρόν πρόβλημα είναι

$$H(x,u, \frac{\partial V}{\partial x}) = \frac{1}{2}\ (x^2+u^2) + \frac{\partial V}{\partial x}\ (x^3+u).$$

Το μοναδικό απόλυτο ελάχιστο της συνάρτησης αυτής με την έννοια της (4.14) είναι

$$u^* = -\frac{\partial V}{\partial x} \qquad (4.18)$$

εξού

$$H^*(x, \frac{\partial V}{\partial x}) = -\frac{1}{2}\ (\frac{\partial V}{\partial x})^2 + \frac{\partial V}{\partial x} x^3 + \frac{1}{2} x^2. \qquad (4.19)$$

Καθότι το πρόβλημα δεν εξαρτάται από τον χρόνο και ο τελικός χρόνος είναι άπειρος, ισχύει $\partial V/\partial t = 0$. Από τις (4.16), (4.19) έχουμε τότε

$$(\frac{\partial V}{\partial x})^2 - 2x^3 (\frac{\partial V}{\partial x}) - x^2 = 0.$$

Η θετική λύση αυτής της εξίσωσης δευτέρου βαθμού δίδει

$$\frac{\partial V}{\partial x} = x^3 + x\sqrt{x^4+1}$$

και δι' αντικαταστάσεως στην (4.18) έχουμε τον μη γραμμικό, μη μεταβλητό βέλτιστο κανόνα ελέγχου

$$u^*(x) = -x^3 - x\sqrt{x^4+1}\,.$$

Αντικαθιστώντας τον κανόνα ελέγχου στην καταστατική εξίσωση, έχουμε το βέλτιστα ελεγχόμενο δυναμικό σύστημα

$$\dot{x} = -x\sqrt{x^4+1}\,.$$

Το σύστημα αυτό είναι ασυμπτωτικά ευσταθές περί την $x = 0$ για κάθε αρχική κατάσταση x_0 (βλ. Πρόβλ. 4.1 για μια γενίκευση του προβλήματος αυτού). Σημειώνεται εδώ ότι, αν είχαμε επιλέξει την αρνητική λύση της τετραγωνικής εξίσωσης ως προς ($\partial V/\partial x$), θα είχαμε καταλήξει σε ένα ασταθές ελεγχόμενο σύστημα $\dot{x} = x\sqrt{x^4+1}$, γι' αυτό η αρνητική λύση αποκλείεται.

■

Σε μερικές περιπτώσεις όπου η αναλυτική επίλυση της εξίσωσης Hamilton-Jacobi-Bellman είναι δύσκολη, μπορεί να επιχειρηθεί μία κατά προσέγγιση λύση που οδηγεί σε έναν υποβέλτιστο κανόνα ελέγχου όπως θα δούμε στο Παράδ. 4.2. Πέραν αυτού, η εξίσωση Hamilton-Jacobi-Bellman μπορεί να χρησιμοποιηθεί για την επαλήθευση κανόνων ελέγχου που είναι υποθετικά βέλτιστοι. Τέλος, η αριθμητική επίλυση της εξίσωσης μπορεί επίσης να οδηγήσει σε έναν κανόνα ελέγχου με μορφή πίνακος.

Παράδειγμα 4.2: Θεωρούμε και πάλι το πρόβλημα του Παραδ. 4.1, επιχειρούμε όμως τώρα μια κατά προσέγγιση λύση με την υπόθεση

$$V(x) = p_0 + p_1x + \frac{1}{2!}\,p_2x^2 + \frac{1}{3!}\,p_3x^3 + \frac{1}{4!}\,p_4x^4$$

όπου p_0, p_1, p_2, p_3, p_4 είναι άγνωστες παράμετροι ελέγχου. Εξ αυτού έχουμε

$$\frac{\partial V}{\partial x} = p_1 + p_2x + \frac{1}{2!}\,p_3x^2 + \frac{1}{3!}\,p_4x^3.$$

Με τις σχέσεις αυτές, την (4.18) ή (4.19) και $\partial V/\partial t = 0$, θεωρούμε την (4.15) και έχουμε, αν αποκλείσουμε τους όρους άνω του τετάρτου βαθμού,

$$p_1^2 + 2p_1p_2x + (p_2^2 + p_1p_3 - 1)x^2 + (\frac{1}{3}\,p_1p_4 + p_2p_3 - 2p_1)x^3$$
$$+ (\frac{1}{4}\,p_3^2 + \frac{1}{3}\,p_2p_4 - 2p_2)x^4 = 0.$$

Επειδή αυτή η ισότητα πρέπει να ισχύει για κάθε x, θα πρέπει και οι πέντε προσθετέοι να μηδενισθούν, εξού

$$p_1 = 0,\quad p_2 = 1,\quad p_3 = 0,\quad p_4 = 6.$$

Τότε έχουμε $\partial V/\partial x = x + x^3$ και από την (4.18) καταλήγουμε στον υποβέλτιστο, μη γραμμικό κανόνα ελέγχου

$$u(t) = -\,x - x^3.$$

Δι' αντικαταστάσεως αυτού του κανόνα ελέγχου στην καταστατική εξίσωση έχουμε το ελεγχόμενο δυναμικό σύστημα

$$\dot{x} = -x$$

που είναι ασυμπτωτικά ευσταθές περί την $x = 0$ για κάθε αρχική κατάσταση x_0.

■

Προβλήματα

4.1 Προσδιορίστε μέσω αναλυτικής επίλυσης της εξίσωσης Hamilton-Jacobi-Bellman τον βέλτιστο κανόνα ελέγχου για την ελαχιστοποίηση του κριτηρίου κόστους

$$J = \frac{1}{2}\int_0^\infty [\varphi(x) + u^2]dt,$$

όπου $\varphi(x)$ θετικά ημιορισμένη, λαμβάνοντας υπόψη την καταστατική εξίσωση

$$\dot{x} = f_1(x) + f_2(x) \cdot u.$$

(Υπόδειξη: Ακολουθήστε τα βήματα του Παραδ. 4.1. Ο βέλτιστος κανόνας ελέγχου είναι

$$u^*(t) = -\frac{f_1(x)}{f_2(x)} - \text{sign}(x)\sqrt{\frac{f_1(x)^2}{f_2(x)^2} + \varphi(x)}.)$$

4.2 Θεωρούμε την ελαχιστοποίηση του κριτηρίου κόστους

$$J = \frac{1}{2}\int_0^\infty (x^2e^{2x} + u^2)dt$$

λαμβάνοντας υπόψη την καταστατική εξίσωση

$$\dot{x} = \alpha x + e^{-x}u.$$

(a) Αναγράψτε τη συνάρτηση Hamilton $H(x,u,\partial V/\partial x)$ για το παρόν πρόβλημα και υπολογίστε το μοναδικό απόλυτο ελάχιστό της ως προς u και την αντίστοιχη ελάχιστη τιμή $H^*(x, \partial V/\partial x)$.

(b) Υπολογίστε μέσω της συνάρτησης Hamilton-Jacobi-Bellman τον ζητούμενο κανόνα βέλτιστου ελέγχου. Αναγράψτε την καταστατική εξίσωση του βέλτιστα ελεγχόμενου δυναμικού συστήματος.

4.3 Ζητείται ελαχιστοποίηση του κριτηρίου κόστους

$$J = \frac{1}{2}\int_{0}^{+\infty} [(x-\alpha)^2 + (u-\beta)^2]dt$$

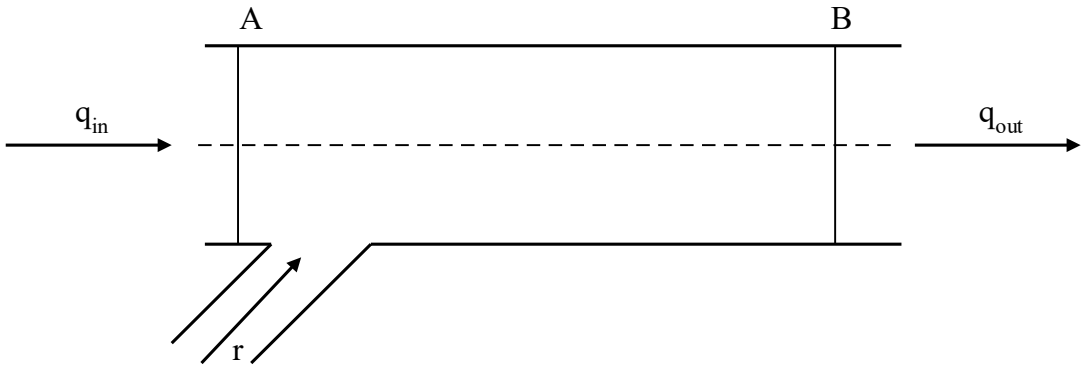
όπου α, β γνωστές παράμετροι, υπό την καταστατική εξίσωση

$$\dot{x} = \ln x + u, \quad x(0) = x_0.$$

(a) Αναγράψτε τη συνάρτηση Hamilton $H(x, u, \partial V/\partial x)$ για το πρόβλημα αυτό και υπολογίστε το μοναδικό απόλυτο ελάχιστό της ως προς u, καθώς και την αντίστοιχη ελάχιστη τιμή της.

(b) Υπολογίστε μέσω της συνάρτησης Hamilton-Jacobi-Bellman τον βέλτιστο κανόνα ελέγχου που ελαχιστοποιεί το κριτήριο κόστους και προσδιορίστε την εξίσωση του βέλτιστα ελεγχόμενου δυναμικού συστήματος.

4.4 Θεωρούμε το απεικονιζόμενο τμήμα αυτοκινητοδρόμου AB που περιέχει μια



ράμπα εισόδου. Η διαφορική εξίσωση του αριθμού x(t) οχημάτων εντός του τμήματος είναι

$$\dot{x} = q_{in} - q_{out} + r$$

όπου $q_{in}(t)$ είναι δεδομένη και σταθερά ροή οχημάτων (οχημ./h) στην αρχή του τμήματος και r(t) είναι η **ελέγξιμη** ροή οχημάτων της ράμπας εισόδου. Για τη ροή q_{out} στο τέλος του τμήματος ισχύει

$$q_{out} = \alpha \cdot x \cdot \left(1 - \frac{x}{b}\right)$$

όπου α, b είναι γνωστές παράμετροι. Ζητείται ο μη μεταβλητός βέλτιστος κανόνας ελέγχου

$$r(t) = R[x(t)]$$

που ελαχιστοποιεί το κριτήριο κόστους

$$J = \frac{1}{2}\int_0^\infty [(x - x_\varepsilon)^2 + (r - d)^2]\, dt$$

όπου x_ε είναι δεδομένος επιθυμητός αριθμός οχημάτων και d είναι δεδομένη, σταθερή προσέλευση οχημάτων στην ράμπα.

(a) Αναγράψτε την συνάρτηση Hamilton $H(x, r, \partial V/\partial x)$ για το παρόν πρόβλημα. Υπολογίστε το μοναδικό απόλυτο ελάχιστο της H ως προς r και την αντίστοιχη ελάχιστη τιμή $H^*(x, \partial V/\partial x)$.

(b) Επιλύστε την συνάρτηση Hamilton-Jacobi-Bellman.

(c) Υπολογίστε τον ζητούμενο κανόνα βέλτιστου ελέγχου.

(d) Έστω ότι η q_{in} δεν είναι σταθερά και δεδομένη, αλλά μεταβλητή με τον χρόνο και μετρήσιμη σε πραγματικό χρόνο. Ισχύει η παραπάνω επίλυση; Αιτιολογείστε την απάντησή σας.

4.5 Ζητείται η ελαχιστοποίηση του κριτηρίου κόστους

$$J = \frac{1}{2}\int_0^\infty \left[(x - \alpha)^2 + (u - \beta)^2\right] dt$$

όπου α, β γνωστές παράμετροι, υπό τον περιορισμό

$$\dot{x} = 2e^x + u + \gamma\,,\ \ x(0) = x_0\,,$$

όπου γ γνωστή σταθερή παράμετρος.

(a) Αναγράψτε τη συνάρτηση Hamilton $H(x, u, \partial V/\partial x)$ για το πρόβλημα αυτό και υπολογίστε το μοναδικό απόλυτο ελάχιστό της ως προς u, καθώς και την αντίστοιχη ελάχιστη τιμή της.

(b) Υπολογίστε μέσω της συνάρτησης Hamilton-Jacobi-Bellman το βέλτιστο κανόνα ελέγχου που ελαχιστοποιεί το κριτήριο κόστους και προσδιορίστε την καταστατική εξίσωση του βέλτιστα ελεγχόμενου δυναμικού συστήματος.

5. ΕΦΑΡΜΟΓΕΣ

Στο κεφάλαιο αυτό θα εξετάσουμε μερικές επιλεγμένες εφαρμογές της μεθοδολογίας του Δυναμικού Προγραμματισμού.

5.1 Το Πρόβλημα Αποθήκευσης

Υποθέτουμε ότι μια επιχείρηση χρειάζεται για μια περίοδο K χρονικών διαστημάτων (ημερών, μηνών,...) τις ποσότητες z(k), k = 0,...,K−1, ενός συγκεκριμένου είδους. Η προμήθεια του είδους αυτού σε ποσότητες u(k) ≥ 0 γίνεται πάντα στην αρχή κάθε χρονικού διαστήματος k και η μέγιστη δυνατή προμήθεια είναι u(k) ≤ U. Σε περίπτωση περισσεύματος στο τέλος ενός χρονικού διαστήματος, υπάρχει δυνατότητα αποθήκευσης μέχρι τη μέγιστη χωρητικότητα X. Η αποθηκευμένη ποσότητα στην αρχή της περιόδου είναι $x(0) = x_0$ και στο τέλος της πρέπει να είναι $x(K) = x_e$. Το περί ου ο λόγος είδος παρέχεται ενδεχομένως, ανάλογα με την συγκεκριμένη εφαρμογή, μόνο σε διακριτές μονάδες (κιβώτια, τεμάχια,...).

Η αποθήκευση ποσότητας x(k) σε κάθε χρονικό διάστημα κοστίζει $\varphi_1[x(k)]$. Το κόστος προμήθειας $\varphi_2[u(k),k]$ εξαρτάται από την ποσότητα u(k) και το χρόνο k (πληθωρισμός, εποχικές τιμές κλπ.). Για δεδομένες ανάγκες z(k), k=0,...,K−1, ζητείται ο βέλτιστος προγραμματισμός u(k), k = 0,...,K−1, που ελαχιστοποιεί το ολικό κόστος λαμβάνοντας υπόψη τους παραπάνω περιορισμούς.

Το πρόβλημα αποθήκευσης μπορεί να διατυπωθεί με τη μορφή ΠΒΕ διακριτού χρόνου ως εξής:

Ελαχιστοποίηση του κριτηρίου κόστους

$$J = \sum_{k=0}^{K-1} \{\varphi_1[x(k)] + \varphi_2[u(k),k]\} \tag{5.1}$$

λαμβάνοντας υπόψη την καταστατική εξίσωση

$$x(k+1) = x(k) + u(k) - z(k), \quad x(0) = x_0, \quad x(K) = x_e \tag{5.2}$$

και τους περιορισμούς

$$0 \le x(k) \le X \tag{5.3}$$

$$0 \le u(k) \le U. \tag{5.4}$$

Επιπλέον, όλες οι μεταβλητές u(k), x(k), z(k) του προβλήματος παίρνουν ενδεχομένως μόνον ακέραιες τιμές.

Το βασικό αυτό πρόβλημα αποθήκευσης έχει βέβαια διάφορες παραλλαγές και γενικεύσεις. Μια παραλλαγή παρουσιάζεται αν αφαιρέσουμε τον περιορισμό x(k) ≥ 0 και θεωρήσουμε ότι η μη ικανοποίηση της ανάγκης z(k) σε κάποιο χρονικό διάστημα οδηγεί σε αντίστοιχη ζημία $\varphi_3[z(k)-x(k)-u(k)]$ όπου φ_3 είναι μια συνάρτηση με την ιδιότητα $\varphi_3(\eta) = 0 \ \forall \eta \le 0$. Στην

περίπτωση αυτή προσθέτουμε βέβαια την ζημιά φ_3 στο κριτήριο κόστους (5.1). Επίσης, αφού αρνητική αποθήκευση δεν νοείται, αντικαθίσταται η καταστατική εξίσωση (5.2) από

$$x(k+1) = \max\{x(k) + u(k) - z(k), 0\}. \qquad (5.5)$$

Μια γενίκευση του προβλήματος παρουσιάζεται αν θεωρήσουμε μια σειρά i=1,…,n από αναγκαία είδη με αντίστοιχες ποσότητες $z_i(k)$, $u_i(k)$, $x_i(k)$. Το κριτήριο κόστους (5.1) είναι τότε

$$J = \sum_{k=0}^{K-1} \sum_{i=1}^{n} \{\varphi_{1,i}\,[x_i(k)] + \varphi_{2,i}\,[u_i(k),k]\} \qquad (5.6)$$

και έχουμε συνολικά n καταστατικές εξισώσεις, μια για κάθε είδος i,

$$x_i(k+1) = x_i(k) + u_i(k) - z_i(k), \quad x_i(0) = x_{0,i}, \quad x_i(K) = x_{e,i} \qquad (5.7)$$

και n σύνολα περιορισμών

$$0 \le x_i(k) \qquad (5.8)$$

$$0 \le u_i(k) \le U_i. \qquad (5.9)$$

Βέβαια, θα είχαμε n ανεξάρτητα προβλήματα αποθήκευσης αν δεν υπήρχε κάποια αλληλεξάρτηση μεταξύ τους. Αυτή η αλληλεξάρτηση προέρχεται από το γεγονός ότι υπάρχει για όλα τα είδη ένας κοινός χώρος αποθήκευσης με μέγιστη χωρητικότητα X. Έχουμε έτσι ένα **πρόβλημα κοινών πόρων** εξ αιτίας του ακόλουθου περιορισμού που αντικαθιστά την δεξιά πλευρά της (5.3)

$$\sum_{i=1}^{n} x_i(k) \le X. \qquad (5.10)$$

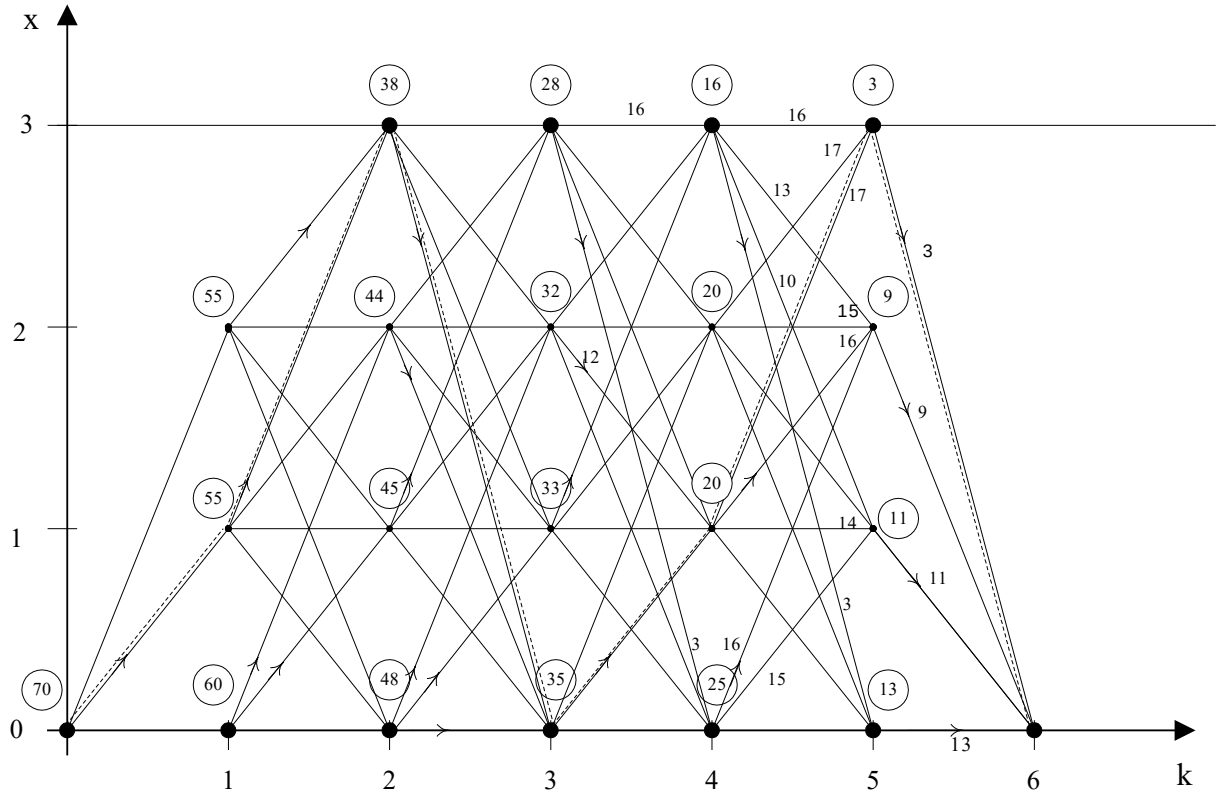
Παράδειγμα 5.1: Θεωρούμε το ακόλουθο αριθμητικό παράδειγμα ενός προβλήματος αποθήκευσης (5.1) - (5.4):

$z(k) = 3\ \forall k \in [0, K-1]$, $X = 3$, $U = 5$, $x_0 = x_e = 0$, $K = 6$, $\varphi_1[x(k)] = x(k)$.

Η συνάρτηση $\varphi_2[u(k), k]$ είναι ανεξάρτητη του χρόνου και δίδεται από τον παρακάτω πίνακα

u	0	1	2	3	4	5
$\varphi_2(u)$	0	7	10	13	15	16

Το Σχήμα 5.1 δίδει τον υπολογισμό της λύσης του προβλήματος και την βέλτιστη τροχιά κατάστασης (διακεκομμένη γραμμή). Είναι ενδιαφέρον να παρατηρήσουμε ότι ο βέλτιστος σχεδιασμός ευνοεί βέλτιστες τροχιές κατάστασης με οδοντωτό σχήμα (βλ. διακεκομμένη γραμμή στο Σχ. 5.1). Ο λόγος αυτής της επιλογής είναι σαφής. Ο βέλτιστος σχεδιασμός επιχειρεί να επωφεληθεί από την έκπτωση που παρουσιάζει ο παραπάνω πίνακας για μεγαλύτερες προμήθειες (η τιμή ανά κομμάτι μειώνεται σταδιακά και είναι για u = 1, 2, 3, 4, 5 ίση προς 7, 5, 4.33, 3.75, 3.2, αντίστοιχα), επιχειρεί δηλαδή την προμήθεια μεγάλων ποσοτήτων που οδηγούν σε μηδενική προμήθεια σ' ένα επόμενο διάστημα. Η πορεία του βέλτιστου ελέγχου είναι



Σχήμα 5.1: Επίλυση του Παραδ. 5.1.

k	0	1	2	3	4	5
u*(k)	4	5	0	4	5	0

Βέβαια η περιγραφείσα πολιτική του βέλτιστου ελέγχου οφείλεται στο σχετικά χαμηλό κόστος αποθήκευσης. Με υψηλότερο κόστος αποθήκευσης ή με μεγαλύτερη χωρητικότητα ή με οποιαδήποτε άλλη αλλαγή τιμών του προβλήματος, αλλάζει αντίστοιχα και η βέλτιστη πολιτική. Η σημασία της βελτιστοποίησης γενικότερα και του Δυναμικού Προγραμματισμού ειδικότερα έγκειται στο γεγονός ότι οποιαδήποτε αλλαγή δεδομένων οδηγεί **αυτόματα** στην αντίστοιχη βέλτιστη πολιτική ελαχιστοποίησης του κόστους. Την βέλτιστη αυτή πολιτική μπορεί να επιχειρήσει να ερμηνεύσει **εκ των υστέρων** ο σχεδιαστής, π.χ. να ερμηνεύσει τους λόγους παραγωγής οδοντωτής τροχιάς στο Σχ. 5.1.

■

5.2 Το Πρόβλημα Αντικατάστασης

Θεωρούμε για μια επιχείρηση την χρήση μιας μηχανής (ή ανταλλακτικού, ή αυτοκινήτου,...) και θέτουμε το ερώτημα αν, και σε ποιές χρονικές στιγμές μέσα σε ένα χρονικό ορίζοντα K διαστημάτων πρέπει να αντικατασταθεί η μηχανή με μια καινούργια. Η απόφαση u(k) για την αντικατάσταση ή όχι της μηχανής παίρνεται στο τέλος κάθε χρονικού διαστήματος k = 0,1,..., K−1. Τα έξοδα αντικατάστασης $\varphi_1[x(k),k]$ εξαρτώνται από την παλαιότητα x(k) της αποσυρομένης μηχανής και από τον χρόνο k. Για κάθε χρονικό διάστημα έχουμε έξοδα λειτουργίας $\varphi_2[x(k),k]$ που εξαρτώνται επίσης από την παλαιότητα x(k) της υπό χρήση μηχανής και από τον χρόνο k. Ζητείται η βέλτιστη πολιτική αντικατάστασης για τον χρονικό ορίζοντα K διαστημάτων και για δεδομένη αρχική παλαιότητα $x(0) = x_0$ της μηχανής.

Ορίζοντας την μεταβλητή ελέγχου

$$u(k) = \begin{cases} 1 & \text{για μη αντικατάσταση} \\ 0 & \text{για αντικατάσταση} \end{cases} \tag{5.11}$$

έχουμε την καταστατική εξίσωση

$$x(k+1) = 1 + x(k) \cdot u(k) \tag{5.12}$$

με επιτρεπτές περιοχές ελέγχου και κατάστασης

$$u(k) \in \{0,1\} \tag{5.13}$$

$$x(k) \in \{1, 2, 3,..., X\} \tag{5.14}$$

αντίστοιχα. Στην (5.14) υποθέτουμε ότι, για λόγους ασφαλείας, η μηχανή απαγορεύεται να χρησιμοποιηθεί και πρέπει να αντικατασταθεί το αργότερο μετά την πάροδο X διαστημάτων.

Το υπό ελαχιστοποίηση κριτήριο κόστους είναι

$$J = \sum_{k=0}^{K-1} \{[1-u(k)] \cdot \varphi_1[x(k), k] + \varphi_2[x(k), k]\} \tag{5.15}$$

με δεδομένο χρονικό ορίζοντα K. Στο κριτήριο κόστους (5.15) μπορεί να εισαχθεί και ένας όρος τελικού κόστους $\vartheta[x(K)]$ που να αντικατοπτρίζει την αξία της (μεταχειρισμένης) μηχανής κατά την τελική χρονική στιγμή K, η οποία εξαρτάται βέβαια από την παλαιότητα x(K), βλ. π.χ. το ακόλουθο Παράδ. 5.2.

Παράδειγμα 5.2: Θεωρούμε ένα πρόβλημα αντικατάστασης με τα εξής στοιχεία. Τα έξοδα λειτουργίας $\varphi_2[x(k), k]$ είναι ανεξάρτητα του χρόνου και καθορίζονται σε συνάρτηση με την παλαιότητα x(k) της μηχανής ως εξής

x	1	2	3	4	5	6	7
$\varphi_2(x)$	1	1	1	2	4	5	5

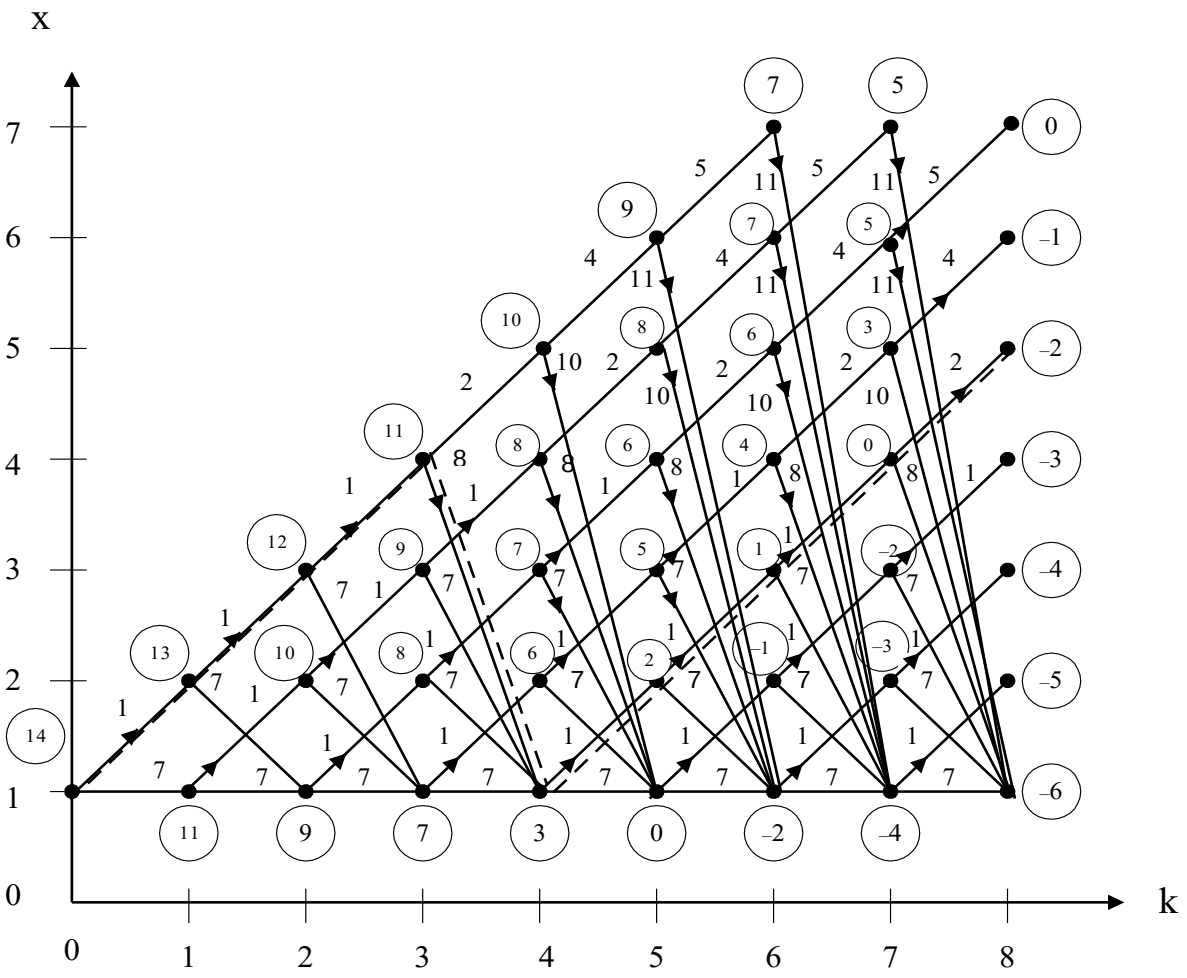
Τα έξοδα αντικατάστασης $\varphi_1[x(k), k]$ είναι ανεξάρτητα της παλαιότητας της μηχανής και του χρόνου και ίσα προς 6. Στην αρχή της περιόδου έχει αγορασθεί μια καινούργια μηχανή, έχουμε δηλαδή x(0) = 1. Για λόγους ασφαλείας, η παλαιότητα της μηχανής δεν μπορεί να υπερβεί τα 7 διαστήματα. Ζητείται η βέλτιστη πολιτική αντικατάστασης της μηχανής για μια περίοδο K = 8 διαστημάτων. Η τελική κατάσταση x(K) είναι ελεύθερη, έχουμε όμως ένα τελικό κόστος

$$\vartheta[x(K)] = x(K) - 7$$

που αντικατοπτρίζει την αξία της υπό χρήση μηχανής στο τέλος της θεωρούμενης περιόδου.

Το Σχ. 5.2 δίδει τον υπολογισμό της λύσης του προβλήματος. Η αρχική μηχανή χρησιμοποιείται επί τέσσερα διαστήματα και αντικαθίσταται πριν από τον διπλασιασμό του κόστους λειτουργίας της που θα επέρχονταν στο πέμπτο διάστημα σύμφωνα με τον παραπάνω πίνακα. Η νέα μηχανή δεν αντικαθίσταται πλέον μέχρι το τελικό, όγδοο, διάστημα.

■



Σχήμα 5.2: Επίλυση του Παραδ. 5.2.

Το πρόβλημα αντικατάστασης γενικεύεται αν θεωρήσουμε μια εγκατάσταση αποτελούμενη από n μηχανές (ή ανταλλακτικά, ...) με αντίστοιχη παλαιότητα $x_i(k)$, έξοδα αντικατάστασης $\varphi_{1,i}[x_i(k), k]$ και μέγιστη επιτρεπτή παλαιότητα X_i, $i = 1,...,n$. Έχουμε τότε n καταστατικές εξισώσεις, μία για κάθε μηχανή i,

$$x_i(k+1) = 1 + x_i(k) \cdot u_i(k) \qquad (5.16)$$

και n σύνολα περιορισμών

$$u_i(k) \in \{0, 1\} \qquad (5.17)$$

$$x_i(k) \in \{1, 2,..., X_i\}. \qquad (5.18)$$

Βέβαια, θα είχαμε n ανεξάρτητα προβλήματα αντικατάστασης αν δεν υπήρχε κάποια αλληλεξάρτηση μεταξύ τους. Αυτή η αλληλεξάρτηση προέρχεται από την παραδοχή ότι τα έξοδα λειτουργίας εξαρτώνται από όλες τις $x_i(k)$ (δηλαδή από το διάνυσμα κατάστασης $\mathbf{x}(k)$), δηλαδή έχουμε $\varphi_2[\mathbf{x}(k), k]$. Στο γενικευμένο αυτό πρόβλημα ζητούμε την βέλτιστη πολιτική $\mathbf{u}(k)$ αντικατάστασης κάθε μηχανής, έτσι ώστε να ελαχιστοποιηθεί το ολικό κόστος

$$J = \sum_{k=0}^{K-1} \{\varphi_2[\mathbf{x}(k), k] + \sum_{i=1}^{n} [1-u_i(k)]\varphi_{1,i}[x_i(k),k]\} \qquad (5.19)$$

για δεδομένο χρονικό ορίζοντα K.

Παράδειγμα 5.3: Θεωρούμε ένα συγκρότημα παραγωγής αποτελούμενο από δύο μηχανές 1 και 2. Η μέγιστη επιτρεπτή παλαιότητα των μηχανών είναι $X_1 = 4$ και $X_2 = 2$ για $k \le K-1$. Κατά συνέπεια έχουμε 8 δυνατές τιμές του διανύσματος κατάστασης **x**. Η αποδοτικότητα (κέρδος) $\varphi_2[\mathbf{x}(k)]$ του συγκροτήματος παραγωγής είναι ανεξάρτητη του χρόνου και εξαρτάται από την παλαιότητα των δύο μηχανών ως εξής

Κατάσταση **x**	1 1	2 1	3 1	4 1	1 2	2 2	3 2	4 2
Κέρδος φ_2 [**x**(k)]	10	10	5	3	8	7	4	1

Η αντικατάσταση των μηχανών 1 και 2 κοστίζει 6 και 4 αντίστοιχα. Ζητείται η βέλτιστη πολιτική αντικατάστασης που μεγιστοποιεί το συνολικό κέρδος

$$J = \sum_{k=K-\kappa}^{K-1} \{\varphi_2[\mathbf{x}(k)] - 6[1-u_1(k)] - 4[1-u_2(k)]\}.$$

Για την επίλυση του προβλήματος παρατηρούμε ότι υπάρχουν το πολύ 4 δυνατές αποφάσεις **u**(k) σε κάθε βαθμίδα, δηλαδή $\mathbf{u}(k) \in \{\mathbf{U}_0, \mathbf{U}_1, \mathbf{U}_2, \mathbf{U}_3\}$ όπου

$$\mathbf{U}_0 = \begin{bmatrix}1\\1\end{bmatrix},\ \mathbf{U}_1 = \begin{bmatrix}0\\1\end{bmatrix},\ \mathbf{U}_2 = \begin{bmatrix}1\\0\end{bmatrix},\ \mathbf{U}_3 = \begin{bmatrix}0\\0\end{bmatrix}.$$

Παρατηρούμε επίσης ότι στην προτελευταία βαθμίδα K−1 η βέλτιστη απόφαση για κάθε επιτρεπτή κατάσταση **x**(K−1) είναι $\mathbf{u}(K-1) = \mathbf{U}_0$ διότι:

- Ελλείψει τελικού κριτηρίου κόστους για την βαθμίδα K στο αντικειμενικό κριτήριο, η αντικατάσταση των μηχανών στην βαθμίδα K−1 θα μείωνε το κέρδος χωρίς κανένα όφελος.
- Ο περιορισμός μέγιστης επιτρεπτής παλαιότητας των μηχανών δεν ισχύει για την τελική βαθμίδα K και άρα δεν υπάρχει ενδεχόμενη αναγκαιότητα αντικατάστασης των μηχανών στην βαθμίδα K−1.

Ο Πίνακας 5.1 δίδει την λύση του προβλήματος για τις βαθμίδες K−κ, κ = 1,..., 5.

■

5.3 Στατικά Προβλήματα Κατανομής

Στο κεφάλαιο αυτό θα δώσουμε καταρχήν τρία παραδείγματα προβλημάτων κατανομής και στη συνέχεια την γενική τους μορφή και την εφαρμογή του Δυναμικού Προγραμματισμού.

Κατάσταση	1 1	2 1	3 1	4 1	1 2	2 2	3 2	4 2
Βαθμίδα Κ–1 βέλτ. έλεγχος	10 $\mathbf{U}_0$	10 $\mathbf{U}_0$	5 $\mathbf{U}_0$	3 $\mathbf{U}_0$	8 $\mathbf{U}_0$	7 $\mathbf{U}_0$	4 $\mathbf{U}_0$	1 $\mathbf{U}_0$
Βαθμίδα Κ–2 $\mathbf{U}_0$ $\mathbf{U}_1$ $\mathbf{U}_2$ $\mathbf{U}_3$ βελτ. έλεγχος	10+7=17 10–6+8 10–4+10 10–10+10 $\mathbf{U}_0$	10+4=14 10–6+8 10–4+5 10–10+10 $\mathbf{U}_0$	5+1 5–6+8=7 5–4+3 5–10+10 $\mathbf{U}_1$	— 3–6+8=5 — 3–10+10 $\mathbf{U}_1$	— — 8–4+10=14 8–10+10 $\mathbf{U}_2$	— — 7–4+5=8 7–10+10 $\mathbf{U}_2$	— — 4–4+3 4–10+10=4 $\mathbf{U}_3$	— — — 1–10+10=1 $\mathbf{U}_3$
Βαθμίδα Κ–3 $\mathbf{U}_0$ $\mathbf{U}_1$ $\mathbf{U}_2$ $\mathbf{U}_3$ βελτ. έλεγχος	10+8 10–6+14 10–4+14=20 10–10+17 $\mathbf{U}_2$	10+4 10–6+14=18 10–4+7 10–10+17 $\mathbf{U}_1$	5+1 5–6+14=13 5–4+5 5–10+17 $\mathbf{U}_1$	— 3–6+14=11 — 3–10+17 $\mathbf{U}_1$	— — 8–4+14=18 8–10+17 $\mathbf{U}_2$	— — 7–4+7 7–10+17=14 $\mathbf{U}_3$	— — 4–4+5 7–10+17=11 $\mathbf{U}_3$	— — — 1–10+17=8 $\mathbf{U}_3$
Βαθμίδα Κ–4 $\mathbf{U}_0$ $\mathbf{U}_1$ $\mathbf{U}_2$ $\mathbf{U}_3$ βελτ. έλεγχος	10+14=24 10–6+18 10–4+18=24 10–10+20 $\mathbf{U}_0$ ή $\mathbf{U}_2$	10+11 10–6+18=22 10–4+13 10–10+20 $\mathbf{U}_1$	5+8 5–6+18=17 5–4+11 5–10+20 $\mathbf{U}_1$	— 3–6+18=15 — 3–10+20 $\mathbf{U}_1$	— — 8–4+18=22 8–10+20 $\mathbf{U}_2$	— — 7–4+13 7–10+20=17 $\mathbf{U}_3$	— — 4–4+11 4–10+20=14 $\mathbf{U}_3$	— — — 1–10+20=11 $\mathbf{U}_3$
Βαθμίδα Κ–5 $\mathbf{U}_0$ $\mathbf{U}_1$ $\mathbf{U}_2$ $\mathbf{U}_3$ βελτ. έλεγχος	10+17 10–6+22 10–4+22=28 10–10+24 $\mathbf{U}_2$	10+14 10–6+22=26 10–4+17 10–10+24 $\mathbf{U}_1$	5+11 5–6+22=21 5–4+15 5–10+24 $\mathbf{U}_1$	— 3–6+22=19 — 3–10+24 $\mathbf{U}_1$	— — 8–4+22=26 8–10+24 $\mathbf{U}_2$	— — 7–4+17 7–10+24=21 $\mathbf{U}_3$	— — 4–4+15 4–10+24=18 $\mathbf{U}_3$	— — — 1–10+24=15 $\mathbf{U}_3$

Πίνακας 5.1: Επίλυση του Παραδ. 5.3.

α) Το Πρόβλημα Αξιοπιστίας

Ένα μηχάνημα λειτουργεί απρόσκοπτα αν όλα τα K εξαρτήματά του λειτουργούν απρόσκοπτα. Η πιθανότητα βλάβης του εξαρτήματος k είναι q_k. Για να αυξήσουμε την πιθανότητα R καλής λειτουργίας του μηχανήματος (δηλαδή την αξιοπιστία του), διαθέτουμε u(k) πανομοιότυπα κομμάτια από κάθε εξάρτημα k. Κατά συνέπεια, αναστολή της λειτουργίας του εξαρτήματος k εμφανίζεται αν υποστούν βλάβη όλα τα u(k) διαθέσιμα κομμάτια. Με την παραδοχή ότι οι πιθανότητες παρουσίας βλάβης στα διαθέσιμα κομμάτια κάθε εξαρτήματος είναι ανεξάρτητες μεταξύ τους, έχουμε την αξιοπιστία του μηχανήματος

$$R = \prod_{k=0}^{K-1} [1 - q_k^{u(k)}]. \qquad (5.20)$$

Η τιμή ανά κομμάτι του εξαρτήματος k είναι c_k. Αν το κόστος του μηχανήματος δεν πρέπει να υπερβαίνει C, ζητείται η κατανομή u(k), k = 0,...,K−1, που μεγιστοποιεί την αξιοπιστία (5.20) του μηχανήματος. Για να αποφύγουμε την παρουσία του γινομένου στο κριτήριο κόστους (5.20) και να το μετατρέψουμε σε άθροισμα, θα επιχειρήσουμε την μεγιστοποίηση του κριτηρίου J = logR, πράγμα που οδηγεί στο ίδιο ακριβώς αποτέλεσμα (γιατί;). Έχουμε λοιπόν μεγιστοποίηση του

$$J = \sum_{k=0}^{K-1} \log[1-q_k^{u(k)}] \qquad (5.21)$$

αντί του (5.20), με τους περιορισμούς

$$\sum_{k=0}^{K-1} c_k\, u(k) \le C \qquad (5.22)$$

$$u(k) \in \{1,2,...\}. \qquad (5.23)$$

Μια εναλλακτική εκδοχή του προβλήματος είναι η ελαχιστοποίηση του κόστους του μηχανήματος για δεδομένη ελάχιστη επιτρεπτή αξιοπιστία R_{min}, πράγμα που επιτυγχάνεται αλλάζοντας τους ρόλους του κριτηρίου (5.21) και του περιορισμού (5.22).

β) Το Πρόβλημα Κατανομής Εργασίας

Ένας φοιτητής θεωρεί ότι αφιερώνοντας u(k) ημέρες προετοιμασίας στο μάθημα k θα πετύχει στις εξετάσεις τον βαθμό φ[u(k),k], όπου φ είναι μια αύξουσα συνάρτηση ως προς u(k). Συνήθως η φ είναι μια μη γραμμική συνάρτηση με σχήμα S (γιατί;). Ο φοιτητής επιθυμεί την μεγιστοποίηση του γενικού βαθμού του

$$J = \sum_{k=0}^{K-1} \varphi[u(k),k] \qquad (5.24)$$

αφιερώνοντας στην συνολική προετοιμασία του ένα μέγιστο αριθμό U ημερών, δηλαδή

$$\sum_{k=0}^{K-1} u(k) \le U. \qquad (5.25)$$

Έχουμε επίσης τον προφανή περιορισμό

$$u(k) \ge 0, \quad k = 0,..., K-1. \qquad (5.26)$$

Αν η επιτυχία σε **κάθε** μάθημα αποτελεί προϋπόθεση επιτυχίας στις εξετάσεις, έχουμε επί πλέον

$$\varphi[u(k),k)] \ge \beta, \quad k = 0,\ldots, K-1 \qquad (5.27)$$

όπου $\beta > 0$ είναι η βάση της βαθμολογίας[1]. Ζητείται η βέλτιστη κατανομή των ημερών προετοιμασίας στα αντίστοιχα μαθήματα που μεγιστοποιεί το κριτήριο (5.24). Σε μια εναλλακτική εκδοχή του προβλήματος, ζητείται ο ελάχιστος συνολικός χρόνος προετοιμασίας έτσι ώστε ο γενικός βαθμός να μην είναι μικρότερος από ένα δεδομένο όριο. Άλλες εφαρμογές του προβλήματος κατανομής εργασίας (π.χ. κατανομή αριθμού εργαζομένων) απαιτούν ακέραιες τιμές της μεταβλητής u(k).

γ) Το Πρόβλημα Κατανομής Φορτίου

Το Παραδ. 3.6 παρουσίασε ένα πρόβλημα κατανομής φορτίου. Σε γενική μορφή έχουμε ένα φορτηγό με μέγιστη χωρητικότητα (ή μέγιστο βάρος φορτίου) B, και K προς φόρτωση είδη με όγκο (ή βάρος) b_k και αξία w_k, $k = 0, \ldots, K-1$, ανά τεμάχιο του κάθε είδους. Ζητείται η κατανομή του φορτίου (δηλαδή ο αριθμός τεμαχίων u(k) ανά είδος) που μεγιστοποιεί την συνολική αξία του φορτίου.

δ) Γενική Επίλυση

Η γενική μορφή των προβλημάτων βέλτιστης κατανομής έχει ως εξής:

Ελαχιστοποίηση

$$J = \sum_{k=0}^{K-1} \varphi[u(k),k] \qquad (5.28)$$

υπό τους περιορισμούς

$$\sum_{k=0}^{K-1} G[u(k),k] \le U \qquad (5.29)$$

$$u(k) \ge u_{min}(k) \qquad (5.30)$$

και ενδεχομένως u(k) ακέραιος.

[1] Ειδικά για το μάθημα Δυναμικός Προγραμματισμός πρέπει να ισχύει $\varphi[u(k),k] \ge 9$, διαφορετικά ο φοιτητής αδυνατεί να επιλύσει το πρόβλημα κατανομής εργασίας του.

Τα προβλήματα βέλτιστης κατανομής είναι καταρχήν στατικά και η εφαρμογή του Δυναμικού Προγραμματισμού δεν είναι προφανής. Μπορούμε όμως, εκμεταλλευόμενοι την αθροιστική μορφή του κριτηρίου κόστους, να θεωρήσουμε ότι κάθε απόφαση u(k) κατανομής για την κατηγορία k αντιστοιχεί σε μια βαθμίδα. Στην περίπτωση αυτή οι "διαχρονικοί" περιορισμοί (5.29) μας δίδουν κατά το Κεφ. 3.4 α) την καταστατική εξίσωση

$$x(k+1) = x(k) + G[u(k),k], \quad x(0) = 0 \qquad (5.31)$$

με τον περιορισμό

$$x(K) \le U. \qquad (5.32)$$

Το πρόβλημα κατανομής ανάγεται λοιπόν στο πρόβλημα ελαχιστοποίησης του (5.28) λαμβάνοντας υπόψη τους περιορισμούς (5.30) - (5.32), και παίρνει έτσι την μορφή ενός ΠΒΕ διακριτού "χρόνου". Αν ο περιορισμός (5.29) έχει την μορφή ισότητας, πρέπει και ο περιορισμός (5.32) να θεωρηθεί σαν ισότητα.

Παράδειγμα 5.4: Ένας φοιτητής θέλει να κατανείμει τον συνολικό χρόνο U = 6 εβδομάδων στην προετοιμασία των εξετάσεων για K = 4 μαθήματα και θεωρεί ότι ο αναμενόμενος βαθμός για κάθε μάθημα δίδεται από την συνάρτηση

$$\varphi[u(k),k] = 10/[1+\exp(\alpha_k - u(k))]$$

όπου η παράμετρος α_k καθορίζει τον βαθμό δυσκολίας του μαθήματος k και u(k) είναι ο αντίστοιχος αριθμός εβδομάδων προετοιμασίας. Για το παρόν παράδειγμα έχουμε $\alpha_0 = \alpha_1 = \alpha_2 = 1$, $\alpha_3 = 2$. Στο μάθημα k = 3 πρέπει ο φοιτητής να επιτύχει οπωσδήποτε και άρα

$$10/(1+\exp[2-u(3)] \ge 5 \Leftrightarrow u(3) \ge 2.$$

Επιπλέον, ο φοιτητής δεν θέλει να αφιερώσει σε κανένα μάθημα περισσότερο από το μισό του συνολικού διαθέσιμου χρόνου, δηλαδή $u(k) \le 3$, k = 0,1,2,3.

Το πρόβλημα του φοιτητή διατυπώνεται λοιπόν ως εξής:

Μεγιστοποίηση

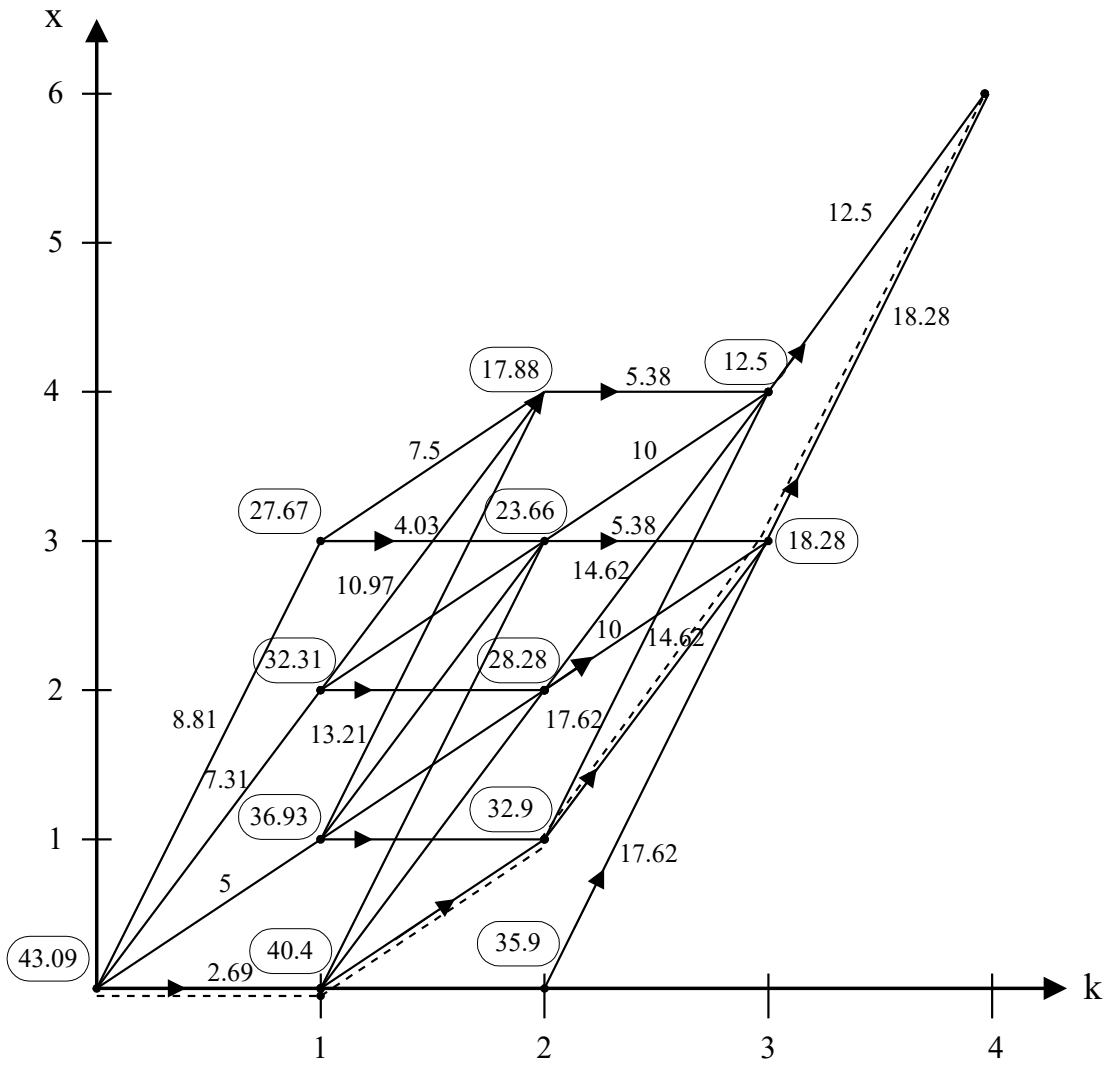
$$J = \sum_{k=0}^{K-1} w_k \, 10/[1+\exp(\alpha_k - u(k))]$$

υπό τους περιορισμούς

$$x(k+1) = x(k) + u(k), \; x(0) = 0, \; x(4) = 6$$
$$3 \ge u(k) \ge 0, \; k = 0,1,2,3$$
$$u(3) \ge 2.$$

Επιπλέον θα θεωρήσουμε την κατανομή σε ακέραιο αριθμό εβδομάδων. Η παράμετρος $w_k = 1+0.5k$ εκφράζει το βάρος του αντίστοιχου μαθήματος στον γενικό βαθμό των εξετάσεων.

Η λύση του προβλήματος απεικονίζεται στο Σχ. 5.3. Η βέλτιστη πολιτική κατανομής χρόνου εργασίας είναι u(0) = 0, u(1) = 1, u(2) = 2, u(3) = 3 και οδηγεί στον γενικό βαθμό J = 43,09 με μέγιστο δυνατό J = 70. Παρατηρούμε ότι η λύση κάνει βέλτιστη χρήση:
– του βάρους κάθε μαθήματος
– της μη γραμμικής μορφής της συνάρτησης φ



Σχήμα 5.3: Επίλυση του Παραδ. 5.4.

για την μεγιστοποίηση του γενικού βαθμού.

■

5.4 Γραμμικός-Τετραγωνικός (ΓΤ) Έλεγχος

Ο **Γραμμικός-Τετραγωνικός Έλεγχος** είναι μια μεθοδολογία της δυναμικής βελτιστοποίησης με ιδιαίτερη σημασία για την θεωρία και την πρακτική της επιστήμης Αυτομάτου Ελέγχου καθότι παρέχει τη δυνατότητα ανάπτυξης βέλτιστων κανόνων ελέγχου για γραμμικά συστήματα MIMO (πολλαπλών εισόδων και εξόδων) απεριόριστων διαστάσεων. Στο κεφάλαιο αυτό θα περιοριστούμε στην παρουσίαση των πολύ βασικών στοιχείων της μεθοδολογίας αυτής.

α) ΓΤ Έλεγχος Συνεχούς Χρόνου

Θεωρούμε το γραμμικό δυναμικό σύστημα

$$\dot{\mathbf{x}} = \mathbf{A}(t)\mathbf{x} + \mathbf{B}(t)\mathbf{u}, \quad \mathbf{x}(0) = \mathbf{x}_0 \tag{5.33}$$

με τα διανύσματα κατάστασης $\mathbf{x}\in R^n$ και ελέγχου $\mathbf{u}\in R^m$ και τους χρονικά μεταβλητούς **πίνακες συστήματος** $\mathbf{A}(t)\in R^{n\times n}$ και **εισόδου** $\mathbf{B}(t)\in R^{n\times m}$. Ζητείται ο βέλτιστος κανόνας ελέγχου για την ελαχιστοποίηση του **τετραγωνικού** κριτηρίου κόστους

$$J = \frac{1}{2}\|\mathbf{x}(T)\|_{\mathbf{S}}^2 + \frac{1}{2}\int_0^T [\|\mathbf{x}(t)\|_{\mathbf{Q}(t)}^2 + \|\mathbf{u}(t)\|_{\mathbf{R}(t)}^2]dt \tag{5.34}$$

με δεδομένο τελικό χρόνο T και συμμετρικούς πίνακες βαρών $\mathbf{S} \geq \mathbf{0}$, $\mathbf{Q}(t) \geq \mathbf{0}$, $\mathbf{R}(t) > \mathbf{0}$, όπου οι δύο τελευταίοι πίνακες μπορούν να μεταβάλλονται με τον χρόνο t. Χάριν συντομίας, η μεταβλητή t θα παραλείπεται εν μέρει στο παρόν κεφάλαιο.

Διευκρινίζουμε τον ορισμό της **τετραγωνικής μορφής**

$$T(\mathbf{x}) = \|\mathbf{x}\|_{\mathbf{Q}}^2 = \mathbf{x}^T\mathbf{Q}\mathbf{x} \tag{5.35}$$

με συμμετρικό πίνακα **Q**. Ονομάζουμε τον πίνακα **Q**

θετικά ορισμένο,	αν $T(\mathbf{x}) > 0$	$\forall \mathbf{x} \neq \mathbf{0}$
θετικά ημιορισμένο,	αν $T(\mathbf{x}) \geq 0$	$\forall \mathbf{x}$
αρνητικά ορισμένο,	αν $T(\mathbf{x}) < 0$	$\forall \mathbf{x} \neq \mathbf{0}$
αρνητικά ημιορισμένο	αν $T(\mathbf{x}) \leq 0$	$\forall \mathbf{x}$

και σημειώνουμε τις περιπτώσεις αυτές με $\mathbf{Q} > \mathbf{0}$, $\mathbf{Q} \geq \mathbf{0}$, $\mathbf{Q} < \mathbf{0}$, $\mathbf{Q} \leq \mathbf{0}$ αντίστοιχα.

Με το κριτήριο κόστους (5.34) επιχειρείται προφανώς η βέλτιστη μεταφορά της αρχικής κατάστασης $\mathbf{x}_0$ σε τιμές κοντά στο μηδέν. Η βέλτιστη αυτή μεταφορά λαμβάνει όμως υπόψη και το κόστος του ελέγχου μέσω του τελευταίου όρου στο κριτήριο (5.34). Η τελική κατάσταση $\mathbf{x}(T)$ είναι ελεύθερη, τιμωρούνται όμως αποκλίσεις από το μηδέν μέσω του τελικού κόστους.

Για την επίλυση του προβλήματος θεωρούμε την εξίσωση Hamilton-Jacobi-Bellman (4.11) και έχουμε με την (5.33)

$$0 = \frac{\partial V}{\partial t} + \min_{\mathbf{u}\in R^m}[\frac{1}{2}\|\mathbf{x}\|_{\mathbf{Q}}^2 + \frac{1}{2}\|\mathbf{u}\|_{\mathbf{R}}^2 + \frac{\partial V^T}{\partial \mathbf{x}}(\mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u})]. \tag{5.36}$$

Η παραπάνω ελαχιστοποίηση δίδει

$$\mathbf{u}^* = -\mathbf{R}^{-1}\mathbf{B}^T\frac{\partial V}{\partial \mathbf{x}}. \tag{5.37}$$

Με την υπόθεση

$$V(\mathbf{x},t) = \frac{1}{2}\|\mathbf{x}\|_{\mathbf{P}(t)}^2 \tag{5.38}$$

όπου $\mathbf{P}(t)$ συμμετρικός πίνακας, έχουμε

$$\frac{\partial V}{\partial t} = \frac{1}{2}\|\mathbf{x}\|_{\dot{\mathbf{P}}(t)}^2 \tag{5.39}$$

$$\frac{\partial V}{\partial \mathbf{x}} = \mathbf{P}(t)\mathbf{x}. \tag{5.40}$$

Δι' αντικαταστάσεως των (5.37), (5.39), (5.40) στην (5.36) έχουμε

$$\frac{1}{2}\|\mathbf{x}\|^2_{\dot{\mathbf{P}}(t)} + \frac{1}{2}\|\mathbf{x}\|^2_{\mathbf{Q}} + \frac{1}{2}\left\|\mathbf{R}^{-1}\mathbf{B}^T\mathbf{P}\mathbf{x}\right\|^2_{\mathbf{R}} + \mathbf{x}^T\mathbf{P}^T\mathbf{A}\mathbf{x} - \mathbf{x}^T\mathbf{P}^T\mathbf{B}\mathbf{R}^{-1}\mathbf{B}^T\mathbf{P}\mathbf{x} = 0. \tag{5.41}$$

Λόγω συμμετρίας του πίνακα **P** έχουμε

$$\mathbf{x}^T\mathbf{P}^T\mathbf{A}\mathbf{x} = \frac{1}{2}\mathbf{x}^T\mathbf{P}\mathbf{A}\mathbf{x} + \frac{1}{2}\mathbf{x}^T\mathbf{A}^T\mathbf{P}\mathbf{x} \tag{5.42}$$

που μπορεί να αντικατασταθεί στην (5.41). Η ισότητα (5.41) πρέπει να ισχύει για κάθε **x** και άρα ο πίνακας **P** πρέπει να ικανοποιεί την ακόλουθη μη γραμμική διαφορική εξίσωση

$$\dot{\mathbf{P}} = -\mathbf{P}\mathbf{A} - \mathbf{A}^T\mathbf{P} + \mathbf{P}\mathbf{B}\mathbf{R}^{-1}\mathbf{B}^T\mathbf{P} - \mathbf{Q} \tag{5.43}$$

που φέρει το όνομα **διαφορική εξίσωση Riccati**. Σαν συνοριακή τιμή της (5.43) εξάγεται από τις (4.12), (5.34) και (5.38) η

$$\mathbf{P}(T) = \mathbf{S}. \tag{5.44}$$

Ο πίνακας $\mathbf{P} \in R^{n\times n}$ φέρει το όνομα **πίνακας Riccati**. Ο αριθμητικός υπολογισμός του **P** μέσω της διαφορικής εξίσωσης (5.43) με την συνοριακή τιμή (5.44) είναι εύκολος ακόμη και για συστήματα μεγάλων διαστάσεων. Να σημειωθεί ότι, λόγω συμμετρίας του **P**, η επίλυση της (5.43) απαιτεί ολοκλήρωση ενός συστήματος αποτελούμενου από n(n+1)/2 (και όχι n^2) εξισώσεις. Μετά τον υπολογισμό του **P**(t) και αντικατάστασή του στις (5.40) και (5.37) έχουμε τον βέλτιστο κανόνα ελέγχου

$$\mathbf{u}(t) = -\mathbf{R}^{-1}(t)\,\mathbf{B}(t)^T\,\mathbf{P}(t)\,\mathbf{x}(t) = -\mathbf{K}(t)\,\mathbf{x}(t) \tag{5.45}$$

όπου **K**(t) είναι ο **πίνακας ελέγχου** του γραμμικού κανόνα (5.45), έχουμε δηλαδή

$$\mathbf{K}(t) = \mathbf{R}(t)^{-1}\,\mathbf{B}(t)^T\,\mathbf{P}(t). \tag{5.46}$$

Παρατηρούμε ότι οι **P**(t) και **K**(t) εξαρτώνται από τα μεγέθη του προβλήματος **A**, **B**, **Q**, **R**, **S** και T, όχι όμως από τις μετρήσεις πραγματικού χρόνου **x**. Κατά συνέπεια, ο υπολογισμός των **P**(t) και **K**(t) μέσω (5.43), (5.44), (5.46) μπορεί να γίνει εκ των προτέρων (off-line), στην φάση ανάπτυξης του κανόνα ελέγχου, ενώ οι υπολογισμοί στη φάση εφαρμογής του κανόνα ελέγχου (on-line) περιορίζονται στην εκτέλεση της (5.45).

Παράδειγμα 5.5: Θεωρούμε το δυναμικό σύστημα

$$\dot{x} = u,\ \ x(0) = x_0,\ \ x(T) = 0 \tag{5.47}$$

και το κριτήριο κόστους

$$J = \frac{1}{2}\int_0^T [x^2 + r\cdot u^2] \tag{5.48}$$

με δεδομένα T και r > 0. Το πρόβλημα είναι προφανώς ΓΤ με A = 0, B = 1, S = ∞ (γιατί;), Q = 1, R = r. Από την (5.43) έχουμε

$$\dot{P} = \frac{P^2}{r} - 1. \qquad (5.49)$$

Η επίλυση αυτής της διαφορικής εξίσωσης δίδει

$$P(t) = \sqrt{r}\ \coth\left[-(t+c)/\sqrt{r}\right] \qquad (5.50)$$

όπου c μια σταθερά. Από την (5.44) έχουμε P(T) = ∞, εξού c = –T στην (5.50) και άρα

$$P(t) = \sqrt{r}\ \coth[(T-t)/\sqrt{r}\,]. \qquad (5.51)$$

Με την (5.46) έχουμε

$$K(t) = \coth[(T-t)/\sqrt{r}\,]/\sqrt{r} \qquad (5.52)$$

και άρα τον βέλτιστο έλεγχο (5.45)

$$u(t) = -K(t) \cdot x(t). \qquad (5.53)$$

■

Το γενικό ΓΤ πρόβλημα δίδει μη μεταβλητό κανόνα ελέγχου αν εκπληρούνται οι ακόλουθες προϋποθέσεις:

(i) Οι πίνακες **A**, **B**, **Q**, **R** είναι μη μεταβλητοί.
(ii) Ο τελικός χρόνος είναι άπειρος: T→ ∞.
(iii) Το σύστημα [**A**, **B**] είναι ελέγξιμο (βλ. Παράρτ. Α.4).
(iv) Το σύστημα [**A**, **C**] είναι παρατηρήσιμο (βλ. Παράρτ. Α.4), όπου **C** είναι τυχαίος πίνακας που εκπληροί $\mathbf{C}^T\mathbf{C} = \mathbf{Q}$.

Υπό τις προϋποθέσεις αυτές μπορεί να αποδειχθεί ότι η ολοκλήρωση της διαφορικής εξίσωσης Riccati (5.43) οδηγεί σε σύγκλιση του **P** προς μία μοναδική στάσιμη τιμή $\overline{\mathbf{P}} \geq \mathbf{0}$, εφ' όσον χρησιμοποιηθεί μία τυχαία συνοριακή τιμή $\mathbf{P}(T) \geq \mathbf{0}$. Η στάσιμη τιμή $\overline{\mathbf{P}}$ είναι βέβαια ανεξάρτητη από την χρησιμοποιηθείσα συνοριακή τιμή **P**(T). Οι ιδιότητες αυτές παρέχουν έναν πρώτο τρόπο (αναλυτικού ή αριθμητικού) υπολογισμού της θετικά ημιορισμένης στάσιμης τιμής $\overline{\mathbf{P}}$ μέσω ολοκλήρωσης της (5.43).

Ένας άλλος τρόπος υπολογισμού του $\overline{\mathbf{P}}$ είναι μέσω επίλυσης της **στάσιμης εξίσωσης** Riccati που εξάγεται από την (5.43) θέτοντας $\dot{\mathbf{P}} = \mathbf{0}$

$$\overline{\mathbf{P}}\,\mathbf{A} + \mathbf{A}^T\overline{\mathbf{P}} - \overline{\mathbf{P}}\,\mathbf{B}\mathbf{R}^{-1}\mathbf{B}^T\overline{\mathbf{P}} + \mathbf{Q} = \mathbf{0}. \qquad (5.54)$$

Η αλγεβρική αυτή εξίσωση είναι μη γραμμική, μπορεί όμως να αποδειχθεί ότι, υπό τις παραπάνω προϋποθέσεις, έχει μία μοναδική θετικά ημιορισμένη λύση $\overline{\mathbf{P}} \geq \mathbf{0}$, η οποία αντιστοιχεί στην ζητούμενη στάσιμη τιμή.

Με τον $\overline{\mathbf{P}}$ έχουμε τον μη μεταβλητό κανόνα ελέγχου

$$\mathbf{u}(t) = -\mathbf{R}^{-1}\mathbf{B}^T\,\overline{\mathbf{P}}\,\mathbf{x}(t) = -\mathbf{K}\mathbf{x}(t) \qquad (5.55)$$

με τον **στάσιμο πίνακα ελέγχου K.**

Παράδειγμα 5.6: Θεωρούμε το πρόβλημα του Παραδ. 5.5 με $T\to\infty$. Οι προϋποθέσεις μη μεταβλητού κανόνα ελέγχου εκπληρούνται και πράγματι έχουμε από την (5.51)

$$\overline{P} = \lim_{T\to\infty} P(t) = \sqrt{r}.$$

Εναλλακτικά μπορούμε να θεωρήσουμε την στάσιμη εξίσωση Riccati που δίδει

$$\overline{P}^2/r - 1 = 0$$

με την μη αρνητική λύση $\overline{P} = \sqrt{r}$. Ο στάσιμος συντελεστής ελέγχου είναι άρα $K = 1/\sqrt{r}$ και το βέλτιστα ελεγχόμενο σύστημα έχει την καταστατική εξίσωση

$$\dot{x} = -x/\sqrt{r}$$

με την ιδιοτιμή $\lambda = -1/\sqrt{r}$, είναι δηλαδή ασυμπτωτικά ευσταθές για κάθε $x\in R$. Το Σχ. 5.4 δίδει τις βέλτιστες τροχιές ελέγχου και κατάστασης για διάφορες τιμές της παραμέτρου βάρους r. Πως ερμηνεύονται οι τροχιές αυτές;

■

β) ΓΤ Έλεγχος Διακριτού Χρόνου

Θεωρούμε το γραμμικό δυναμικό σύστημα διακριτού χρόνου

$$\mathbf{x}(k+1) = \mathbf{A}(k)\,\mathbf{x}(k) + \mathbf{B}(k)\,\mathbf{u}(k),\ \ \mathbf{x}(0) = \mathbf{x}_0 \qquad (5.56)$$

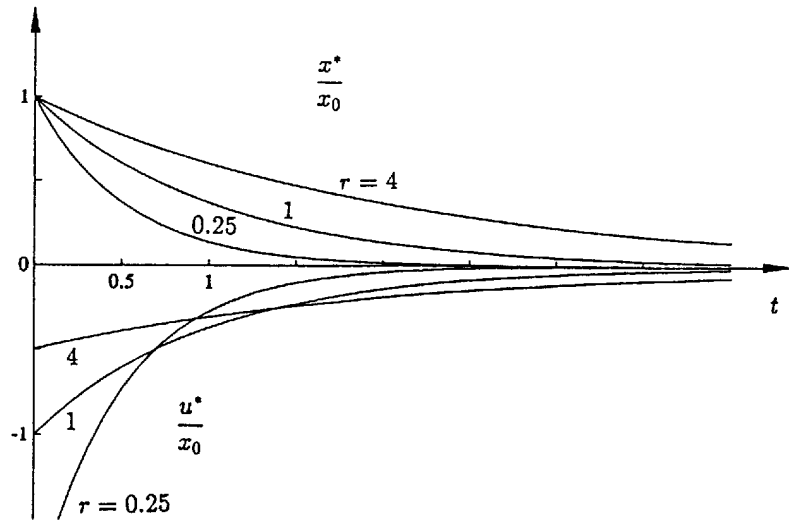
με τα διανύσματα κατάστασης $\mathbf{x}\in R^n$ και ελέγχου $\mathbf{u}\in R^m$ και τους χρονικά μεταβλητούς **πίνακες συστήματος** $\mathbf{A}(k)\in R^{n\times n}$ και **εισόδου** $\mathbf{B}(k)\in R^{n\times m}$. Ζητείται ο βέλτιστος κανόνας ελέγχου για την ελαχιστοποίηση του κριτηρίου κόστους

$$J = \frac{1}{2}\|\mathbf{x}(K)\|_{\mathbf{S}}^2 + \frac{1}{2}\sum_{k=0}^{K-1}[\|\mathbf{x}(k)\|_{\mathbf{Q}(k)}^2 + \|\mathbf{u}(k)\|_{\mathbf{R}(k)}^2] \qquad (5.57)$$

με δεδομένο τελικό χρόνο K και συμμετρικούς πίνακες βαρών $\mathbf{S}\geq\mathbf{0}$, $\mathbf{Q}(k)\geq\mathbf{0}$, $\mathbf{R}(k)\geq\mathbf{0}$ και με την παραδοχή ότι οι πίνακες $\mathbf{B}(K-1)^T\mathbf{S}\mathbf{B}(K-1)+\mathbf{R}(K-1)$ και $\mathbf{B}(k)^T\mathbf{Q}(k+1)\,\mathbf{B}(k)+\mathbf{R}(k)$, $k = 0,\ldots, K-2$, είναι πλήρους βαθμού.

Για την επίλυση του προβλήματος θεωρούμε την αναδρομική εξίσωση του Bellman (3.9) και έχουμε

$$V(\mathbf{x},k) = \min_{\mathbf{u}\in\mathbf{R}^m}\left\{\frac{1}{2}\|\mathbf{x}\|_{\mathbf{Q}}^2 + \frac{1}{2}\|\mathbf{u}\|_{\mathbf{R}}^2 + V(\mathbf{A}\mathbf{x}+\mathbf{B}\mathbf{u},\ k+1)\right\}. \qquad (5.58)$$



Σχήμα 5.4. Βέλτιστες τροχιές για το Παραδ. 5.6.

Για την παραπάνω ελαχιστοποίηση έχουμε δια παραγωγισμού

$$\mathbf{R}(\mathrm{k})\,\mathbf{u}(\mathrm{k}) + \mathbf{B}(\mathrm{k})^{\mathrm{T}} \frac{\partial \mathrm{V}[\mathbf{x}(\mathrm{k}+1), \mathrm{k}+1]}{\partial \mathbf{x}(\mathrm{k}+1)} = \mathbf{0}. \tag{5.59}$$

Με την υπόθεση

$$\mathrm{V}(\mathbf{x}, \mathrm{k}) = \frac{1}{2}\|\mathbf{x}\|^2_{\mathbf{P}(\mathrm{k})} \tag{5.60}$$

όπου $\mathbf{P}(\mathrm{k})$ συμμετρικός πίνακας, έχουμε

$$\frac{\partial \mathrm{V}[\mathbf{x}(\mathrm{k}), \mathrm{k}]}{\partial \mathbf{x}(\mathrm{k})} = \mathbf{P}(\mathrm{k})\,\mathbf{x}(\mathrm{k}) \tag{5.61}$$

και δι' αντικαταστάσεως των (5.56), (5.61) στην (5.59) καταλήγουμε στον βέλτιστο κανόνα ελέγχου

$$\mathbf{u}(\mathrm{k}) = -\mathbf{L}(\mathrm{k})\,\mathbf{x}(\mathrm{k}) \tag{5.62}$$

όπου ο πίνακας ελέγχου $\mathbf{L}(\mathrm{k})$ υπολογίζεται από

$$\mathbf{L}(\mathrm{k}) = [\mathbf{B}(\mathrm{k})^{\mathrm{T}}\,\mathbf{P}(\mathrm{k}+1)\,\mathbf{B}(\mathrm{k}) + \mathbf{R}(\mathrm{k})]^{-1}\,\mathbf{B}(\mathrm{k})^{\mathrm{T}}\,\mathbf{P}(\mathrm{k}+1)\,\mathbf{A}(\mathrm{k}). \tag{5.63}$$

Για τον υπολογισμό του **πίνακα Riccati διακριτού χρόνου** $\mathbf{P}(\mathrm{k})$ έχουμε από τις (5.58), (5.60), (5.62), (5.63) την ακόλουθη **εξίσωση διαφορών Riccati**

$$\mathbf{P}(\mathrm{k}) = \mathbf{A}(\mathrm{k})^{\mathrm{T}}\,\mathbf{P}(\mathrm{k}+1)\,\mathbf{A}(\mathrm{k}) + \mathbf{Q}(\mathrm{k}) - \mathbf{L}(\mathrm{k})^{\mathrm{T}}\,\mathbf{B}(\mathrm{k})^{\mathrm{T}}\,\mathbf{P}(\mathrm{k}+1)\,\mathbf{A}(\mathrm{k}) \tag{5.64}$$

με την συνοριακή τιμή

$$\mathbf{P}(K) = \mathbf{S}. \qquad (5.65)$$

Η επίλυση των (5.63), (5.64), (5.65) είναι εύκολη μέσω υπολογιστού ακόμη και για προβλήματα μεγάλων διαστάσεων. Ο υπολογισμός αυτός γίνεται εκτός γραμμής (off-line) καθότι οι εξισώσεις είναι ανεξάρτητες των μετρήσεων πραγματικού χρόνου.

Παράδειγμα 5.7: Ζητούμε τον βέλτιστο έλεγχο για την ελαχιστοποίηση του τετραγωνικού κριτηρίου

$$J = \frac{1}{2}Sx(K)^2 + \frac{1}{2}\sum_{k=0}^{K-1} [x(k)^2 + r \cdot u(k)^2]$$

όπου $S \geq 0$, $r \geq 0$, λαμβάνοντας υπόψη το γραμμικό σύστημα

$$x(k+1) = x(k) + u(k), \quad x(0) = x_0.$$

Για την επίλυση του ΓΤ αυτού προβλήματος έχουμε από τις (5.63), (5.64), (5.65) με A = 1, B = 1, Q = 1, R = r

$$L(k) = \frac{P(k+1)}{r + P(k+1)}, \quad P(k) = \frac{r + (r+1)\cdot P(k+1)}{r + P(k+1)}, \quad P(K) = S$$

και τον βέλτιστο έλεγχο

$$u(k) = -L(k)\, x(k).$$

Ο Πίνακας 5.2 δίδει τις τιμές P(k), L(k) για K = 4, r = 1. Το Σχ. 5.5 δείχνει τις αντίστοιχες βέλτιστες τροχιές x(k), u(k) για την περίπτωση $S \to \infty$ και $x_0 = 2$.

■

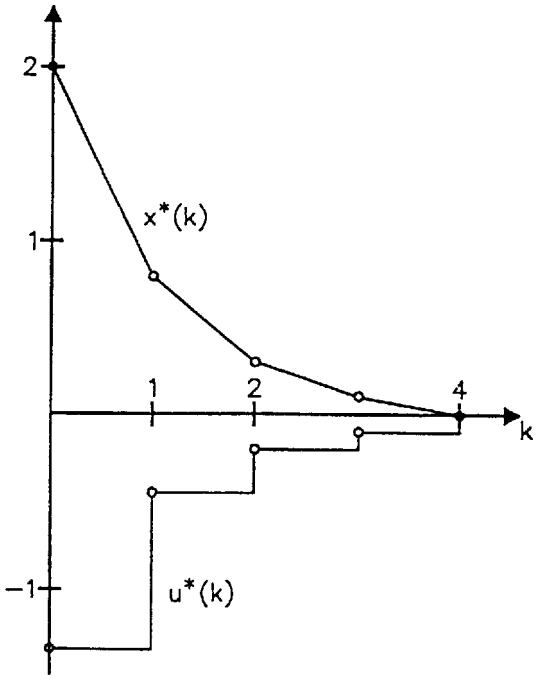
Το γενικό ΓΤ πρόβλημα διακριτού χρόνου δίδει μη μεταβλητό κανόνα ελέγχου αν εκπληρούνται οι ακόλουθες προϋποθέσεις:

(i) Οι πίνακες **A**, **B**, **Q**, **R** είναι μη μεταβλητοί.
(ii) Ο τελικός χρόνος είναι άπειρος: $K \to \infty$.
(iii) Το σύστημα [**A**, **B**] είναι ελέγξιμο (βλ. Παράρτ. Α.4).
(iv) Το σύστημα [**A**, **C**] είναι παρατηρήσιμο (βλ. Παράρτ. Α.4), όπου **C** είναι τυχαίος πίνακας που εκπληροί $\mathbf{C}^T\mathbf{C} = \mathbf{Q}$.

Υπό τις προϋποθέσεις αυτές μπορεί να αποδειχθεί ότι ο αναδρομικός υπολογισμός του **P**(K) μέσω της εξίσωσης διαφορών Riccati (5.64) οδηγεί σε σύγκλιση προς μία μοναδική στάσιμη τιμή $\overline{\mathbf{P}} \geq \mathbf{0}$ εφ' όσον χρησιμοποιηθεί μια τυχαία συνοριακή τιμή $\mathbf{P}(T) \geq \mathbf{0}$. Η τιμή $\overline{\mathbf{P}}$ μπορεί επομένως να υπολογισθεί μέσω της (5.64) ή, εναλλακτικά, ως η μοναδική θετικά ημιορισμένη λύση της ακόλουθης **στάσιμης εξίσωσης Riccati διακριτού χρόνου**

K	4	3	2	1	0
P(k)	S	$\frac{1+2S}{1+S}$	$\frac{3+5S}{2+3S}$	$\frac{8+13S}{5+8S}$	$\frac{21+34S}{13+21S}$
L(k)		$\frac{S}{1+S}$	$\frac{1+2S}{2+3S}$	$\frac{3+5S}{5+8S}$	$\frac{8+13S}{13+21S}$

Πίνακας 5.2. Βέλτιστες τιμές για το Παράδ. 5.7.



Σχήμα 5.5: Βέλτιστες τροχιές για το Παράδ. 5.7.

$$\overline{\mathbf{P}} = \mathbf{A}^T \overline{\mathbf{P}}\, \mathbf{A} + \mathbf{Q} - \mathbf{L}^T \mathbf{B}^T \overline{\mathbf{P}}\, \mathbf{A} \tag{5.66}$$

με τον **στάσιμο πίνακα ελέγχου**

$$\mathbf{L} = (\mathbf{B}^T\ \overline{\mathbf{P}}\, \mathbf{B} + \mathbf{R})^{-1}\, \mathbf{B}^T\ \overline{\mathbf{P}}\, \mathbf{A}. \tag{5.67}$$

Καταλήγουμε έτσι στον μη μεταβλητό κανόνα ελέγχου

$$\mathbf{u}(k) = -\mathbf{L}\ \mathbf{x}(k). \tag{5.68}$$

Παράδειγμα 5.8: Θεωρούμε το πρόβλημα του Παραδ. 5.7, τώρα όμως για $K \to \infty$. Οι προαναφερθείσες προϋποθέσεις για μη μεταβλητό κανόνα ελέγχου εκπληρούνται και έχουμε από τις (5.66), (5.67)

$$\overline{P} = \overline{P} + 1 - \frac{\overline{P}^2}{r + \overline{P}} \Rightarrow \overline{P}^2 - \overline{P} - r = 0 \Rightarrow \overline{P} = \frac{1}{2} + \sqrt{\frac{1}{4} + r}$$

(όπου αποκλείσθηκε η αρνητική λύση) και

$$L = \frac{\overline{P}}{r+\overline{P}} = \frac{1+\sqrt{1+4r}}{2r+1+\sqrt{1+4r}}.$$

Το βέλτιστα ελεγχόμενο σύστημα έχει επομένως την καταστατική εξίσωση

$$x(k+1) = (1-L)x(k) = \frac{2r}{2r+1+\sqrt{1+4r}}\,x(k)$$

με πόλο λ(r) που εκπληροί $0 \le \lambda(r) < 1 \;\; \forall r \ge 0$, είναι δηλαδή ασυμπτωτικά ευσταθές. Για r = 0 έχουμε x(k+1) = 0, δηλαδή σύγκλιση στο μηδέν μέσα σε ένα και μόνο διακριτό διάστημα, ανεξάρτητα της αρχικής τιμής κατάστασης, με έλεγχο u(k) = –x(k).

■

Προβλήματα

5.1 Στο Παραδ. 5.1, ο υπολογισθείς βέλτιστος σχεδιασμός ανταποκρίνεται στην εβδομαδιαία πολιτική αγορών μιας επιχείρησης που επαναλαμβάνεται πανομοιότυπα κάθε εβδομάδα. (Θα ήταν αυτό δυνατό αν οι συναρτήσεις του προβλήματος δεν ήταν μη μεταβλητές με το χρόνο;) Θεωρούμε όμως τώρα ότι δεν υπάρχει αποκλίνων λόγος για το δεδομένο $x_0 = 0$, $x_e = 0$. Υπολογίστε τον βέλτιστο εβδομαδιαίο σχεδιασμό με ελεύθερη αρχική και τελική κατάσταση υπό τον περιορισμό $x_0 = x_e$ (γιατί;).

5.2 Υπολογίστε την χείριστη πολιτική προμηθειών για το πρόβλημα αποθήκευσης του Παραδ. 5.1. Ερμηνεύστε γιατί η πολιτική αυτή μεγιστοποιεί το ολικό κόστος προμήθειας και αποθήκευσης.

5.3 Ένας φοιτητής θέλει να κατανείμει τον συνολικό χρόνο U ημερών στην προετοιμασία των εξετάσεων για K μαθήματα και θεωρεί ότι ο αναμενόμενος βαθμός για κάθε μάθημα δίδεται από την συνάρτηση

$$\varphi[u(k), k] = 10[1-\exp(-\alpha_k\, u(k))]$$

όπου η παράμετρος α_k καθορίζει τον βαθμό δυσκολίας του μαθήματος k και u(k) είναι ο αντίστοιχος αριθμός εβδομάδων προετοιμασίας. Ο φοιτητής δεν θέλει να αφιερώσει σε κανένα μάθημα περισσότερο από το μισό του συνολικού διαθέσιμου χρόνου. Στόχος του φοιτητή είναι η μεγιστοποίηση του μέσου βαθμού

$$J = \sum_{k=0}^{K-1} w_k\, \varphi[u(k), k]$$

όπου w_k είναι παράμετροι βάρους των μαθημάτων.

(a) Διατυπώστε το πρόβλημα του φοιτητή σαν πρόβλημα βέλτιστου ελέγχου με το κριτήριο κόστους, την καταστατική εξίσωση και όλους τους περιορισμούς.

(b) Αναπτύξτε ένα πρόγραμμα που δέχεται σαν δεδομένα εισόδου τις τιμές U, K, (α_k, w_k, k = 0,..., K–1) και υπολογίζει την βέλτιστη κατανομή u(k), k = 0,..., K–1, και τον μέγιστο μέσο βαθμό.

5.4 Θεωρούμε το Πρόβλημα Αντικατάστασης για ένα συγκρότημα παραγωγής που αποτελείται από δύο μηχανές 1 και 2 με μέγιστη επιτρεπτή παλαιότητα 3 και 2 χρόνια αντίστοιχα. Στο τέλος κάθε χρόνου παίρνεται η απόφαση για αντικατάσταση μίας, καμίας ή αμφοτέρων των μηχανών. Η αποδοτικότητα του συγκροτήματος παραγωγής είναι $\varphi[\mathbf{x}(k)] = 10 - x_1(k) \cdot x_2(k)$ όπου x_1, x_2 η παλαιότητα κάθε μηχανής (σε χρόνια). Η αντικατάσταση των μηχανών 1 και 2 κοστίζει 5 και 3 αντίστοιχα, δεν επιτρέπεται όμως η αντικατάσταση της μηχανής 1 αν έχει ένα μόνο χρόνο λειτουργίας. Ζητείται η βέλτιστη πολιτική αντικατάστασης που μεγιστοποιεί το συνολικό κέρδος (δηλαδή την αποδοτικότητα μείον τα έξοδα αντικατάστασης) για χρονικό ορίζοντα 4 χρόνων. Στην αρχή έχουμε δύο καινούργιες μηχανές, δηλαδή $x_1(0) = x_2(0) = 1$ ενώ η τελική κατάσταση είναι ελεύθερη.

(a) Πόσες και ποιές διακριτές δυνατότητες αποφάσεων έχουμε γενικά στο τέλος κάθε χρονιάς; Πόσες και ποιές επιτρεπτές διακριτές τιμές έχει το διάνυσμα κατάστασης **x**;

(b) Έστω $\mathbf{X}_1$, $\mathbf{X}_2$,... οι διακριτές επιτρεπτές τιμές κατάστασης, π.χ. $\mathbf{X}_1 = (1,1)$, $\mathbf{X}_2 = (2,1)$ κλπ. Αναγράψτε στο αντίστοιχο διάγραμμα όλες τις διακριτές επιτρεπτές τιμές και σχεδιάστε όλες τις επιτρεπτές μεταβάσεις με τα αντίστοιχα κόστη.

(c) Επιλύστε γραφικά το πρόβλημα προσδιορίζοντας την βέλτιστη πολιτική αντικατάστασης και το μέγιστο συνολικό κέρδος.

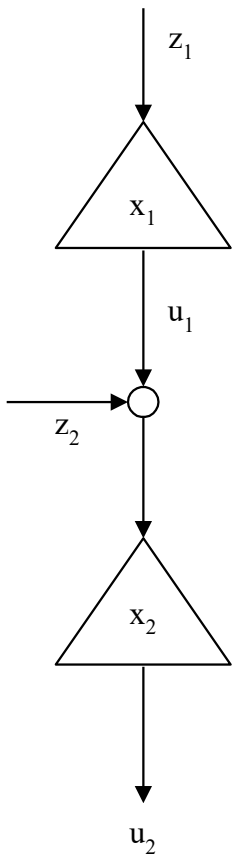
5.5 Θεωρούμε το παρακάτω δίκτυο φυσικών δεξαμενών νερού με γνωστές εξωτερικές εισροές $z_1(k)$, $z_2(k)$, k = 0,...,K–1, ελεγχόμενες εκροές $u_1(k)$, $u_2(k)$ και όγκους αποθηκευμένου νερού $x_1(k)$, $x_2(k)$ με περιορισμούς (i = 1,2)

$$0 \le u_i(k) \le U_i\,, \quad 0 \le x_i(k) \le X_i.$$

Ο αρχικός και τελικός όγκος αποθηκευμένου νερού είναι $\mathbf{x}(0) = \mathbf{x}(K) = 0.5 \cdot [X_1\ X_2]^T$. Κάθε εκροή κινεί γεννήτριες παραγωγής ηλεκτρικής ενέργειας σε ποσότητα $\sqrt{u_i(k)}, i = 1,2$. Ζητείται η τροχιά ελέγχου που μεγιστοποιεί την παραγωγή ηλεκτρικής ενέργειας σε δεδομένο χρονικό ορίζοντα K.

(a) Διατυπώστε το πρόβλημα βέλτιστου ελέγχου με το κριτήριο κόστους, τις καταστατικές εξισώσεις και όλους τους περιορισμούς.

(b) Αναπτύξτε ένα πρόγραμμα που δέχεται σαν δεδομένα εισόδου τις τιμές K, ($z_1(k)$, $z_2(k)$, k = 0,...,K–1), U_1, U_2, X_1, X_2 και υπολογίζει την βέλτιστη τροχιά ελέγχου.



5.6 Μία εμπορική επιχείρηση πωλεί στη διάρκεια ενός έτους $z_1(k)$ ψυγεία και $z_2(k)$ πλυντήρια ανά μήνα, όπου k = 0,...,11. Η προμήθεια των ειδών αυτών γίνεται στην αρχή κάθε μήνα και η μέγιστη δυνατή προμήθεια είναι $u_i(k) \le U_i$, i = 1,2. Σε περίπτωση περισσεύματος στο τέλος κάθε μήνα, υπάρχει δυνατότητα αποθήκευσης του αριθμού $x_1(k)$ και $x_2(k)$ των ψυγείων και πλυντηρίων, αντίστοιχα, μέχρι τη μέγιστη χωρητικότητα $x_1(k)+x_2(k) \le X$. Οι αποθηκευμένες ποσότητες στο τέλος του έτους πρέπει να είναι $x_1(12) = x_2(12) = 0.5$ X. Η αποθήκευση ποσότητας $x_i(k)$ κοστίζει κάθε μήνα $A_i x_i(k)$, i=1,2, ενώ το κόστος προμήθειας σε συνάρτηση της αντίστοιχης παραγγελθείσας ποσότητας είναι $B_i \cdot \ln[u_i(k)]$. Ζητείται η βέλτιστη πολιτική προμηθειών (κανόνας ελέγχου) $\mathbf{u}(k) = \mathbf{R}[\mathbf{x}(k),k]$ της επιχείρησης, που ελαχιστοποιεί το συνολικό κόστος προμήθειας και αποθήκευσης για κάθε επιτρεπτό $\mathbf{x}(k)$, k = 0,...,11.

(a) Διατυπώστε το πρόβλημα βελτιστοποίησης με τις καταστατικές εξισώσεις, το κριτήριο κόστους και όλους τους περιορισμούς.

(b) Αναπτύξτε ένα ψευδοκώδικα (διάγραμμα ροής) για την αριθμητική επίλυση του προβλήματος. Το πρόγραμμα να δέχεται σαν δεδομένα εισόδου τιμές
– $z_1(k)$, $z_2(k)$, k = 0,...,11
– U_1, U_2, X, A_1, A_2, B_1, B_2
και να υπολογίζει υπό μορφή πίνακα τον βέλτιστο κανόνα ελέγχου.

5.7 Θεωρούμε το δυναμικό σύστημα που περιγράφεται από την ακόλουθη διαφορική εξίσωση

$$\dot{x}(t) = -x(t) + u(t)$$

και το κριτήριο κόστους

$$J = \frac{1}{2} \cdot \int_0^T \left[3 \cdot x(t)^2 + u(t)^2\right] dt$$

με άπειρο τελικό χρόνο ($T \to \infty$).

Ζητείται ο μη μεταβλητός (γραμμικός-τετραγωνικός) κανόνας ελέγχου.

(a) Αναφέρετε τον πίνακα συστήματος, τον πίνακα εισόδου και τους πίνακες βαρών του προβλήματος.

(b) Αναγράψτε και επιλύστε τη στάσιμη εξίσωση Riccati.

(c) Υπολογίστε το στάσιμο πίνακα ελέγχου και αναγράψτε τον μη μεταβλητό κανόνα ελέγχου. Ποιά είναι η καταστατική εξίσωση του βέλτιστα ελεγχόμενου συστήματος;

Αν ο τελικός χρόνος T δεν είναι άπειρος, ισχύει η παραπάνω επίλυση; Αιτιολογείστε την απάντησή σας.

5.8 Έχετε μια μηχανή που δουλεύει συνεχώς μέσα σε μια γραμμή παραγωγής. Για λόγους ασφαλείας η ηλικία της δε μπορεί να ξεπεράσει τα X χρονικά διαστήματα. Σε κάθε χρονικό διάστημα k πρέπει να λάβετε την απόφαση για το αν θα γίνει αντικατάστασή της, συντήρησή της, ή τίποτε από τα δύο. Η απόφασή σας μορφοποιεί τον έλεγχο u(k) ως εξής:

u(k) = 1 αν γίνει αντικατάσταση
u(k) = –1 αν γίνει συντήρηση
u(k) = 0 αν δεν γίνει τίποτα.

Οι αποφάσεις αυτές πραγματοποιούνται χωρίς καθυστέρηση.

Έστω ότι x(k) είναι η ηλικία της μηχανής που βρίσκεται σε λειτουργία κατά τη χρονική στιγμή k. Αν σ' αυτή τη χρονική στιγμή επιλέξουμε αντικατάσταση, τότε θα έχουμε καινούργιο μηχάνημα και άρα x(k+1) = 1. Αν επιλέξουμε συντήρηση, τότε μειώνουμε την λειτουργική ηλικία του μηχανήματος κατά b χρονικά διαστήματα και άρα x(k+1) = x(k)–b+1. Η ηλικία της μηχανής δε μπορεί να πάρει αρνητικές τιμές και άρα η συντήρηση μπορεί να γίνει μόνον όταν έχει ξεπεραστεί ένα χρονικό όριο. Τέλος, αν επιλέξουμε να μην κάνουμε τίποτε, η ηλικία αυξάνει κατά ένα, οπότε x(k+1) = x(k)+1. Επειδή αρχίζουμε με μία νέα μηχανή, έχουμε x(0) = 1. Είναι προφανές ότι η καταστατική εξίσωση έχει ως εξής:

$$x(k+1) = \begin{cases} 1 & \text{αν } u(k) = 1 \\ x(k) + 1 & \text{αν } u(k) = 0 \\ x(k) + 1 - b & \text{αν } u(k) = -1. \end{cases}$$

Η απόφαση της κάθε χρονικής στιγμής θα πρέπει να λαμβάνει υπόψη τα παρακάτω κόστη:

i) Κόστος λειτουργίας $f_1[x(k),k]$
ii) Κόστος αντικατάστασης $f_2[x(k),k]$
iii) Κόστος συντήρησης $f_3[x(k),k]$.

Ο χρονικός ορίζοντας για τον οποίο ενδιαφερόμαστε είναι K ενώ μας είναι αδιάφορο ποιά θα είναι η ηλικία της μηχανής τη χρονική στιγμή K.

(a) Προσδιορίστε το όριο ηλικίας μετά το οποίο είναι δυνατόν να γίνει συντήρηση.

(b) Αναγράψτε το κριτήριο κόστους.

(c) Μορφοποιήστε το παραπάνω πρόβλημα σαν ένα πρόβλημα Βέλτιστου Ελέγχου.

(d) Αναπτύξτε το διάγραμμα ροής του προγράμματος που επιλύει το διατυπωθέν πρόβλημα. Σαν είσοδο θα πρέπει να δέχεται το χρονικό όριο ασφαλείας για τη λειτουργία της μηχανής X, τον χρονικό ορίζοντα K και την βελτίωση b που επιφέρει η συντήρηση. Η έξοδός του θα πρέπει να δίνει όλες τις βέλτιστες ακολουθίες αποφάσεων και το κόστος τους. Θεωρείστε ότι:

$f_1[x(k),k] = x(k)^2$, $f_2[x(k),k] = A\sqrt{k} + C$ και $f_3[x(k),k] = e^{\frac{x(k)}{2}}$.

5.9 Μία εταιρεία που εμπορεύεται ένα συγκεκριμένο προϊόν, θέλει να προγραμματίσει τη διακίνηση του προϊόντος για έναν ορίζοντα K χρονικών διαστημάτων. Η ζήτηση του προϊόντος στο διάστημα k είναι z(k), k=0,...,K−1. Στην αρχή κάθε χρονικού διαστήματος η εταιρεία κάνει μία προμήθεια του προϊόντος u(k), όπου u(k)≥0 και u(k) ακέραιος. Σε περίπτωση περισσεύματος υπάρχει δυνατότητα αποθήκευσης μέχρι X μονάδων προϊόντος. Στην αρχή του χρονικού ορίζοντα η αποθήκη είναι γεμάτη και υπάρχει η απαίτηση στο τέλος του να είναι πάλι γεμάτη. Η αποθήκευση της ποσότητας x(k) με x(k)≥0 σε κάθε χρονικό διάστημα κοστίζει $\varphi_1[x(k)] = x(k)$. Το κόστος της προμήθειας στην αρχή του διαστήματος k εξαρτάται τόσο από την παραγγελθείσα ποσότητα όσο και από την τρέχουσα χρονική περίοδο. Όσο μεγαλύτερη είναι η ποσότητα, τόσο καλύτερη είναι η τιμή προμήθειας, αλλά, λόγω πληθωρισμού, η τιμή αυξάνεται με τον χρόνο. Έτσι το κόστος παραγγελίας έχει τη μορφή

$$\varphi_2[u(k), k] = [1 + \alpha(k)] \ln[u(k) + 1]$$

όπου α(k) είναι ο πληθωρισμός της περιόδου k.

Έστω ότι ο χρονικός ορίζοντας για τον οποίο θέλουμε να σχεδιασθεί μια βέλτιστη πολιτική αποθέματος περιλαμβάνει K = 5 χρονικά διαστήματα, η χωρητικότητα της αποθήκης είναι X = 3 κομμάτια προϊόντος, η μέγιστη δυνατή ποσότητα παραγγελίας είναι U = 4 κομμάτια και η ζήτηση είναι σταθερή με z(k) = 2 για k = 0, 1,..., 4.

(a) Μορφοποιήστε το παραπάνω πρόβλημα σαν ένα πρόβλημα βέλτιστου ελέγχου, αναγράφοντας:
   (i) Την αντικειμενική συνάρτηση.
   (ii) Την καταστατική εξίσωση.
   (iii) Όλους τους περιορισμούς.

(b) Σχεδιάστε σε αντίστοιχο διάγραμμα όλες τις επιτρεπτές μεταβάσεις από βαθμίδα σε βαθμίδα, αναγράφοντας το αντίστοιχο κόστος μετάβασης.
   (Θεωρήστε ότι α(k) = 0.06 για k = 0, 1, …,4.)

(c) Προσδιορίστε την βέλτιστη πολιτική αποθεμάτων και παραγγελιών για όλον τον χρονικό ορίζοντα.

Αν τελικά η ζήτηση κατά το χρονικό διάστημα k = 2 δεν ήταν z(2) = 2 κομμάτια, αλλά ήταν z(2) = 3 κομμάτια, προσδιορίστε ποια είναι η βέλτιστη πολιτική αποθεμάτων και παραγγελιών από το χρονικό διάστημα k = 2 μέχρι το τελευταίο χρονικό διάστημα.

5.10 Ένας υπάλληλος εφορίας έχει πάνω στο γραφείο του δύο στοίβες με φακέλους. Η πρώτη στοίβα έχει συνολικά m φακέλους ενώ η δεύτερη n. Ο υπάλληλος αναζητεί ένα σημαντικό έγγραφο και θέλει να το βρει στο συντομότερο δυνατό χρόνο. Δεν ξέρει ακριβώς σε ποιόν φάκελο είναι, αλλά θεωρεί πως η πιθανότητα να βρίσκεται στο φάκελο i=1,…,m της πρώτης στοίβας είναι p_i, ενώ η πιθανότητα να βρίσκεται στον φάκελο j=1,...,n της δεύτερης στοίβας είναι q_j. Ξέρει επίσης ότι αν ψάξει όλους τους φακέλους θα το βρει σίγουρα, δηλαδή ισχύει:

$$\sum_{i=1}^{m} p_i + \sum_{j=1}^{n} q_j = 1 .$$

Σε κάθε στοίβα οι φάκελοι πρέπει να ελεγχθούν με την σειρά, δηλαδή δεν μπορεί π.χ. να ελεγχθεί ο φάκελος i=4 αν δεν έχει πρώτα ελεγχθεί ο φάκελος i=3. Ο υπάλληλος ξέρει πως για να ελέγξει ένα φάκελο χρειάζεται μια ώρα, οπότε το πολύ μετά από K=m+n ώρες θα έχει βρει το έγγραφο αφού θα έχει ψάξει όλους τους φακέλους. Στην αρχή κάθε ώρας k, με k=0,....,m+n−1, πρέπει να αποφασίσει σε ποια στοίβα u(k) θα συνεχίσει την ερευνά του, όπου $u(k) \in \{1,2\}$, ανάλογα με το αν θα επιλέξει την στοίβα 1 ή 2 αντίστοιχα.

Υποθέτουμε τώρα ότι m=3 και n=4. Ζητείται η τροχιά ελέγχου u(k), k=0,….,m+n−1, που ελαχιστοποιεί τον αναμενόμενο χρόνο εύρεσης του εγγράφου, ο οποίος δίνεται από την συνάρτηση:

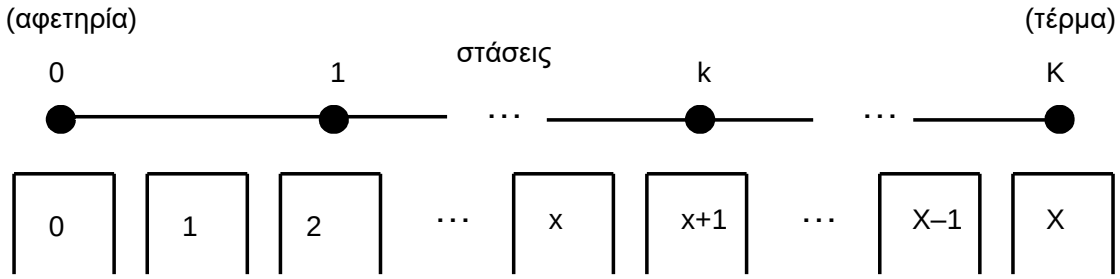
$$E = \sum_{k=0}^{m+n-1} (k+1) \cdot P(k) ,$$

με P(k) την πιθανότητα το έγγραφο να βρίσκεται στον φάκελο που ελέγχεται την ώρα k=0,….,m+n−1.

(a) Το πρόβλημα που αντιμετωπίζει ο υπάλληλος της εφορίας μπορεί να μοντελοποιηθεί σαν ένα πρόβλημα αποφάσεων πολλαπλών βαθμίδων. Θεωρούμε πως με (i,j) χαρακτηρίζεται η κατάσταση στην οποία έχουν ελεγχθεί i φάκελοι της στοίβας 1 και j φάκελοι της στοίβας 2 (μετά από i+j ώρες). Σχεδιάστε το πλήρες δίκτυο δυνατών μεταβάσεων του προβλήματος, ξεκινώντας από την κατάσταση (0,0) τη χρονική στιγμή k=0 (αρχή της πρώτης ώρας), και εφαρμόζοντας τις επιτρεπτές τιμές ελέγχου u(k), k=1,….,m+n−1, μέχρι τον έλεγχο όλων των φακέλων, δηλαδή μέχρι την τελική κατάσταση (m,n) κατά τον τελικό χρόνο m+n. Υπόδειξη: Κάθε κατάσταση αρκεί να εμφανίζεται μια μόνο φορά σε κάθε βαθμίδα.

(b) Προσδιορίστε τον βέλτιστο κανόνα ελέγχου για το παραπάνω πρόβλημα με p_1=0.1, p_2=0.4, p_3=0.05, q_1=0.2, q_2=0.1, q_3=0.05, q_4=0.1, καθώς και την τροχιά ελέγχου που ελαχιστοποιεί τον αναμενόμενο χρόνο E που θα παρέλθει μέχρι να βρεθεί το έγγραφο. Πόσος είναι ο χρόνος αυτός;

5.11 Μια γραμμή αστικού λεωφορείου διασχίζει X+1 οικοδομικά τετράγωνα ίσου μήκους και έχει σχεδιαστεί να περιλαμβάνει K+1 στάσεις k=0,…,K (όπου 0 η αφετηρία και K το τέρμα της γραμμής), με K<X (βλ. Σχήμα). Σε κάθε οικοδομικό τετράγωνο x=0,…,X υπάρχει ζήτηση d_x επιβατών. Ζητείται η βέλτιστη τοποθέτηση των στάσεων σε αντίστοιχα οικοδομικά τετράγωνα έτσι ώστε να ελαχιστοποιηθεί η συνολική απόσταση που πρέπει να διανύσουν οι επιβάτες d_x, x=0,1,…,X, προκειμένου να μεταβούν στην πλησιέστερη σε αυτούς στάση. Σημειώνουμε ότι οι στάσεις 0 και K (αφετηρία και τέρμα) είναι προκαθορισμένες στα τετράγωνα 0 και X, αντίστοιχα, και άρα απαιτείται η βέλτιστη τοποθέτηση των ενδιάμεσων K−1 στάσεων μόνον.



Οικοδομικά τετράγωνα

Για τη μοντελοποίηση του προβλήματος ορίζουμε ως μεταβλητή ελέγχου $u(k) \in \{1,2,\ldots,U\}$ την απόσταση (σε οικοδομικά τετράγωνα) μεταξύ των στάσεων k και k+1 και έχουμε έτσι τον "διαχρονικό" περιορισμό

$$\sum_{k=0}^{K-1} u(k) = X .$$

Η μέγιστη επιτρεπτή απόσταση μεταξύ δύο στάσεων είναι U και η ελάχιστη 1 οικοδομικό τετράγωνο. Σημειώνουμε με x(k) το οικοδομικό τετράγωνο όπου τοποθετείται η στάση k. Το αντικειμενικό κριτήριο δίδεται από:

$$J = \sum_{k=0}^{K-1} \phi\left[u(k), x(k), \mathbf{d}\right]$$

όπου $\mathbf{d}=[d_0\ d_1\ \ldots d_x]^T$ και η συνάρτηση ϕ υπολογίζεται ως εξής:

Οι επιβάτες των οικοδομικών τετραγώνων x(k) , k=0,1,…,K, (όπου υπάρχουν στάσεις) έχουν μηδενική διαδρομή και άρα δεν επιβαρύνουν το αντικειμενικό κριτήριο. Οι επιβάτες ενδιάμεσων οικοδομικών τετραγώνων μεταβαίνουν στην πλησιέστερη στάση και επιβαρύνουν το αντικειμενικό κριτήριο με την διανυθείσα απόσταση (σε οικοδομικά

τετράγωνα) επί τον αριθμό τους d_x. Καταλήγουμε έτσι στον ακόλουθο υπολογισμό των $\phi[u(k), x(k), \mathbf{d}]$:

u(k)=1 : $\phi = 0$ (δεν υπάρχει ενδιάμεσο τετράγωνο)

u(k)=2 : $\phi = d_{x(k)+1}$ (ένα ενδιάμεσο τετράγωνο, το x(k)+1)

u(k)=3 : $\phi = d_{x(k)+1}+d_{x(k)+2}$ (δύο ενδιάμεσα τετράγωνα, τα x(k)+1 και x(k)+2)

κ.ο.κ.

(a) Διαμορφώστε το διατυπωθέν πρόβλημα "καταμερισμού κοινών πόρων" ως πρόβλημα βέλτιστου ελέγχου με την καταστατική εξίσωση, το ως άνω αντικειμενικό κριτήριο και τους περιορισμούς ελέγχου και κατάστασης.

(b) Επιλύστε το πρόβλημα γραφικά για X=8 , K=4, U=3, $\mathbf{d}=[5\ 5\ 8\ 3\ 2\ 8\ 2\ 1\ 1]^T$.

5.12 Θεωρούμε ένα έργο το οποίο αποτελείται από K ανεξάρτητες εργασίες (δράσεις) οι οποίες εκτελούνται η μία μετά την άλλη. Το έργο θεωρείται επιτυχές αν και οι K εργασίες εκτελεστούν με επιτυχία. Η υλοποίηση της εργασίας k, k=1,...,K, ανατίθεται σε μια ομάδα u(k) εργαζομένων και η πιθανότητα $p_k[u(k)]$ επιτυχούς υλοποίησής της εξαρτάται από το μέγεθος u(k) της ομάδας. (Η πιθανότητα αυτή είναι ανεξάρτητη από τις πιθανότητες επιτυχούς υλοποίησης άλλων εργασιών.) Δεδομένου ότι υπάρχουν συνολικά W διαθέσιμοι εργαζόμενοι, μας ενδιαφέρει να καθορίσουμε την κατανομή των εργαζομένων στις K εργασίες που μεγιστοποιεί την πιθανότητα επιτυχίας του συνολικού έργου.

(a) Διατυπώστε το πρόβλημα σαν ένα πρόβλημα κατανομής (πόρων, εργασιών, κ.λπ.) αναγράφοντας (α) το κριτήριο κόστους και (β) την καταστατική εξίσωση και όλους τους περιορισμούς.

(b) Θεωρώντας K=3, W=5 και πιθανότητες επιτυχούς υλοποίησης των εργασιών k=1,2,3 συναρτήσει του μεγέθους της ομάδας u(k) όπως δίδονται από τον πίνακα

u(k)	$p_1[u(k)]$	$p_2[u(k)]$	$p_3[u(k)]$
0	0	0	0
1	0.5	0.4	0.3
2	0.6	0.7	0.6
3	0.7	0.8	0.8
4	0.8	0.9	0.85
5	0.85	0.95	0.85

επιλύστε το πρόβλημα γραφικά σε αντίστοιχο σχήμα.

Σημείωση: Για την επίλυση του προβλήματος υποθέστε ότι και οι 5 εργαζόμενοι πρέπει να έχουν κατανεμηθεί στις K εργασίες στο τέλος του έργου.

5.13 Επιλύστε το Πρόβλημα 3.12 με χρήση της μεθοδολογίας Γραμμικού - Τετραγωνικού Ελέγχου διακριτού χρόνου.

6. ΠΑΡΑΛΛΑΓΜΕΝΟΙ ΑΛΓΟΡΙΘΜΟΙ ΔΥΝΑΜΙΚΟΥ ΠΡΟΓΡΑΜΜΑΤΙΣΜΟΥ

6.1 Εισαγωγή

Όπως είδαμε στο Κεφάλαιο 3.3, το κεντρικό μειονέκτημα που περιορίζει σημαντικά το πεδίο πρακτικής εφαρμογής του Διακριτού Δυναμικού Προγραμματισμού (ΔΔΠ), είναι ο εκθετικά αυξανόμενος υπολογιστικός φόρτος αριθμητικής επίλυσης ΠΒΕ. Στην προσπάθειά τους να επιλύσουν πρακτικά προβλήματα μεσαίων ή υψηλών διαστάσεων χρησιμοποιώντας την μεθοδολογία του Δυναμικού Προγραμματισμού, πολλοί ερευνητές ανέπτυξαν και πρότειναν παραλλαγμένους αλγόριθμους Δυναμικού Προγραμματισμού που οδηγούν σε μείωση του υπολογιστικού φόρτου. Στο κεφάλαιο αυτό θα παρουσιάσουμε τους σημαντικότερους τρεις από τους αλγόριθμους αυτούς.

Όλοι οι παραλλαγμένοι αλγόριθμοι είναι αλγόριθμοι **διαδοχικών προσεγγίσεων**. Αυτό σημαίνει ότι υπολογίζονται σε μια σειρά **επαναλήψεων** διαδοχικές επιτρεπτές τροχιές ελέγχου $\mathbf{u}^{(1)}(k)$, $\mathbf{u}^{(2)}(k)$,... (πάντα για k=0,...,K−1) μέχρι την εκπλήρωση κάποιου **κριτηρίου σύγκλισης** που υποδηλώνει ικανοποιητική προσέγγιση της ζητούμενης λύσης του προβλήματος $\mathbf{u}^*(k)$, k = 0,...,K−1. Η επιτρεπτή **τροχιά εκκίνησης** $\mathbf{u}^{(0)}(k)$ του αλγόριθμου θεωρείται δεδομένη, πρέπει δηλαδή να καθορισθεί από τον χρήστη. Ο εκάστοτε αλγόριθμος περιγράφει τον ιδιαίτερο τρόπο μετασχηματισμού της προηγούμενης προσέγγισης $\mathbf{u}^{(l-1)}(k)$ στην νέα προσέγγιση $\mathbf{u}^{(l)}(k)$ και ο μετασχηματισμός αυτός εφαρμόζεται πανομοιότυπα σε κάθε επανάληψη μέχρι την ικανοποίηση του κριτηρίου σύγκλισης.

Για το κεφάλαιο αυτό θα θεωρήσουμε το γενικό ΠΒΕ διακριτού χρόνου όπως διατυπώθηκε στο Κεφάλαιο 3.1. Σημειώνουμε ότι, εφαρμόζοντας την τροχιά ελέγχου $\mathbf{u}^{(l)}(k)$ μιας επανάληψης l στην εξίσωση κατάστασης (3.2) με δεδομένη αρχική κατάσταση $\mathbf{x}_0$, μπορούμε να υπολογίσουμε την αντίστοιχη τροχιά κατάστασης $\mathbf{x}^{(l)}(k)$. Τόσο η τροχιά εκκίνησης $\mathbf{u}^{(0)}(k)$ όσο και οι διαδοχικές προσεγγίσεις $\mathbf{u}^{(l)}(k)$ είναι επιτρεπτές, δηλαδή η εκάστοτε τροχιά $\mathbf{u}^{(l)}(k)$ και η αντίστοιχη $\mathbf{x}^{(l)}(k)$ ικανοποιούν τους περιορισμούς (3.4), (3.5) και τον τελικό στόχο (3.3), ενδεχομένως βέβαια στην διακριτή τους μορφή.

Πρέπει να τονισθεί ότι η μείωση υπολογιστικού φόρτου που επιτυγχάνεται μέσω των παραλλαγμένων αλγορίθμων έχει κάποιο τίμημα ως προς τα γνωστά πλεονεκτήματα του ΔΔΠ. Ένα σημαντικό μειονέκτημα όλων των παραλλαγών αυτού του κεφαλαίου είναι η δυνατότητα σύγκλισης προς μια τροχιά ελέγχου που αποτελεί σχετικό και όχι απόλυτο ελάχιστο του προβλήματος, ή ακόμη και η δυνατότητα απόκλισης. Η εγγύηση σύγκλισης ή σύγκλισης προς ένα απόλυτο ελάχιστο δεν είναι επομένως δεδομένη, εκτός αν εκπληρούνται αντίστοιχες παραδοχές που μειώνουν την γενικότητα εφαρμογής των αλγορίθμων. Ας σημειωθεί όμως ότι, και χωρίς την αναφερθείσα θεωρητική εγγύηση, οι παραλλαγμένοι αλγόριθμοι βρίσκουν επιτυχή πρακτική εφαρμογή σε ένα ευρύ πεδίο προβλημάτων.

6.2 Διακριτός Διαφορικός Δυναμικός Προγραμματισμός (ΔΔΔΠ)

Ο ΔΔΔΠ επιλύει σε κάθε επανάληψη το διακριτοποιημένο πρόβλημα του Κεφαλαίου 3.1 με τους επί πλέον περιορισμούς

$$\left|x_i(k) - x_i^{(l-1)}(k)\right| \le \Delta_i^{(l)}(k),\ i = 1,\ldots,n \qquad (6.1)$$

όπου $\mathbf{x}^{(l-1)}$ είναι η προσδιορισθείσα βέλτιστη τροχιά στην προηγούμενη επανάληψη l–1. Με άλλα λόγια, το διακριτοποιημένο πρόβλημα επιλύεται σε κάθε επανάληψη σε ένα **διάδρομο** εύρους $\mathbf{\Delta}^{(l)}(k)$ περί την προηγούμενη προσέγγιση $\mathbf{x}^{(l-1)}(k)$, πράγμα που μειώνει σημαντικά τον αντίστοιχο υπολογιστικό φόρτο. Η επιλογή του εύρους $\mathbf{\Delta}^{(l)}$ του διαδρόμου, αλλά και των διακριτών διαστημάτων $\Delta\mathbf{u}^{(l)}$, $\Delta\mathbf{x}^{(l)}$, μπορεί να μεταβάλλεται σε κάθε επανάληψη, με τάση βέβαια πτωτική. Το κριτήριο σύγκλισης ικανοποιείται αν

$$\left\|\mathbf{u}^{(l)}(k) - \mathbf{u}^{(l-1)}(k)\right\|^{\infty} < \varepsilon \quad \forall k \in [0, K-1] \qquad (6.2)$$

όπου ε > 0 ένα όριο ανοχής και $\|\mathbf{u}(k)\|^{\infty} = \max_i(|u_i(k)|)$.

Παράδειγμα 6.1: Θεωρούμε το πρόβλημα του Παραδ. 3.3, δηλαδή ελαχιστοποίηση του

$$J = \frac{1}{2}\sum_{k=0}^{3}[x(k)^2 + u(k)^2]$$

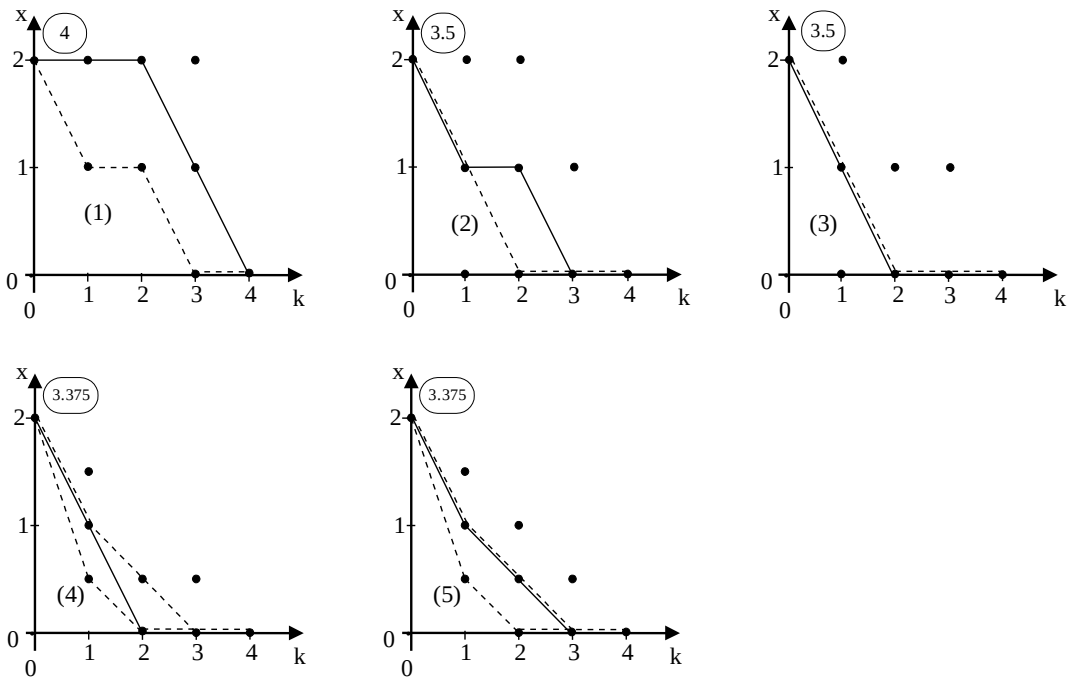
λαμβάνοντας υπόψη

$$x(k+1) = x(k)+u(k),\ \ x(0) = 2,\ \ x(4) = 0$$

$$\mathscr{X} = \{x(k) \mid 0 \le x(k) \le 2\},\ \mathscr{U} = \{u(k) \mid -2 \le u(k) \le 0\}.$$

Σαν τροχιά εκκίνησης καθορίζουμε την επιτρεπτή τροχιά $u^{(0)}(0) = u^{(0)}(1) = 0$, $u^{(0)}(2) = u^{(0)}(3) = -1$ που δίδει $x^{(0)}(0) = x^{(0)}(1) = x^{(0)}(2) = 2$, $x^{(0)}(3) = 1$, $x^{(0)}(4) = 0$ με κόστος 7.5. Το Σχ. 6.1 δίδει τη λύση του προβλήματος μέσω ΔΔΔΠ δείχνοντας για κάθε επανάληψη l = 1,...,5:

– Την προηγούμενη προσέγγιση κατάστασης $x^{(l-1)}(k)$ σε συνεχή γραμμή.
– Τα διακριτά σημεία του διαδρόμου επίλυσης.
– Την νέα προσέγγιση κατάστασης $x^{(l)}(k)$ σε διακεκομμένη γραμμή και το αντίστοιχο κόστος (σε κύκλο πάνω από την αρχική κατάσταση).



Σχήμα 6.1: Διαδοχικές προσεγγίσεις για το Παραδ. 6.1.

Σε όλες τις επαναλήψεις θεωρούμε εύρος διαδρόμου ίσο προς ένα διακριτό διάστημα Δx. Προφανώς το εύρος του διαδρόμου είναι μηδενικό για τις χρονικές στιγμές 0 και 4 λόγω καθορισμένης αρχικής και τελικής κατάστασης. Εννοείται ότι ο διάδρομος δεν επεκτείνεται σε απαγορευμένες περιοχές, δηλαδή περιορίζεται σε τιμές $0 \le x(k) \le 2$.

Οι πρώτες τρεις επαναλήψεις γίνονται με διακριτά διαστήματα Δx = Δu = 1. Η πρώτη επανάληψη μειώνει το κριτήριο κόστους από 7.5 στα 4 και η δεύτερη στα 3.5, ενώ η τρίτη δεν πετυχαίνει παραπέρα βελτίωση. Μειώνοντας τα διακριτά διαστήματα σε Δx = Δu = 0.5 πετυχαίνουμε μείωση του κόστους στα 3.375 στην τέταρτη επανάληψη ενώ η πέμπτη δεν αποφέρει άλλη βελτίωση. Εδώ μπορούμε να σταματήσουμε ή να μειώσουμε περισσότερο τα διακριτά διαστήματα για να βελτιώσουμε αντίστοιχα τη λύση.

■

Η εξοικονόμηση υπολογιστικού χρόνου μέσω ΔΔΔΠ εξαρτάται προφανώς από το εύρος του διαδρόμου **Δ** και τον αριθμό των επαναλήψεων, όπου βέβαια ο αριθμός των επαναλήψεων εξαρτάται εν γένει από το εύρος του διαδρόμου. Σ' ένα πρόβλημα με n = 5 και α = 100, επί παραδείγματι, η (3.20) δίδει χρόνο υπολογισμού ανάλογο προς $100^5 = 10^{10}$. Με εύρος διαδρόμου ίσο προς 5 διακριτά σημεία, έχουμε, για κάθε επανάληψη του ΔΔΔΠ, χρόνο υπολογισμού ανάλογο προς $5^5 = 3125$. Βέβαια και στον ΔΔΔΠ παραμένει το πρόβλημα εκθετικής αύξησης του υπολογιστικού φόρτου, έχουμε δηλαδή απλώς μια (σε μερικές εφαρμογές σημαντική) επέκταση των ορίων εφαρμογής του Δυναμικού Προγραμματισμού. Τα μειονεκτήματα του ΔΔΔΠ σε σχέση με τον ΔΔΠ είναι:

- **Σύγκλιση** μπορεί γενικά να αποδειχθεί μόνον αν γίνουν σχετικά περιοριστικές παραδοχές.
- Δεν μπορεί να αποκλεισθεί το ενδεχόμενο σύγκλισης προς **σχετικό ελάχιστο,** εκτός από ειδικές κατηγορίες προβλημάτων.
- **Κανόνας ελέγχου** υπολογίζεται μόνο για τον διάδρομο της τελικής επανάληψης και όχι για την ολότητα της επιτρεπτής περιοχής κατάστασης $\mathcal{X}(k)$.

6.3 Αυξάνων Δυναμικός Προγραμματισμός (ΑΔΠ)

Ο ΑΔΠ βρίσκει καταρχήν εφαρμογή μόνο σε ΠΒΕ με αντιστρεπτή καταστατική εξίσωση, όπου δηλαδή υπάρχει μια συνάρτηση $\mathbf{F}[\mathbf{x}(k), \mathbf{x}(k+1), k]$ έτσι ώστε για κάθε ζεύγος καταστάσεων $\mathbf{x}(k)$ και $\mathbf{x}(k+1)$ να έχουμε (βλ. και την (3.18))

$$\mathbf{u}(k) = \mathbf{F}[\mathbf{x}(k), \mathbf{x}(k+1), k] \qquad (6.3)$$

σαν μοναδική λύση αντιστροφής της καταστατικής εξίσωσης (3.2). Αντιστρεπτή καταστατική εξίσωση προϋποθέτει βέβαια m = n, δηλαδή ίσο αριθμό μεταβλητών ελέγχου και κατάστασης. Σε μερικές όμως περιπτώσεις προβλημάτων που δεν εκπληρούν αυτές τις παραδοχές, είναι δυνατό να εισαχθούν εικονικές μεταβλητές ελέγχου $\tilde{\mathbf{u}}(k)$ με τον επιπλέον περιορισμό $\tilde{\mathbf{u}}(k) = \mathbf{0}$, $k = 0,...,K-1$.

Σε κάθε επανάληψη του αλγόριθμου ΑΔΠ έχουμε

$$x_j^{(l)}(k) = x_j^{(l-1)}(k) \quad \forall k \in [0, K] \qquad (6.4)$$

για κάθε $j \neq i$, επιτρέπεται δηλαδή η αλλαγή μιας μόνο μεταβλητής κατάστασης αρ. i. Με τον τρόπο αυτό η διάσταση του προβλήματος Δυναμικού Προγραμματισμού που επιλύεται σε κάθε

επανάληψη, μειώνεται από n σε 1. Η μεταβλητή κατάστασης, της οποίας η αλλαγή επιτρέπεται, αλλάζει από επανάληψη σε επανάληψη, δηλαδή έχουμε i=1,...,n, στην συνέχεια επανατίθεται i=1,...,n κ.ο.κ.

Για τον προσδιορισμό των υπολογισμών κάθε επανάληψης παρατηρούμε ότι έχουμε λόγω της (6.4)

$$\mathbf{x}(k) = \mathbf{X}[\alpha(k),k] = \mathbf{x}^{(l-1)}(k) + \alpha(k)\mathbf{e}_i \qquad (6.5)$$

όπου το διάνυσμα $\mathbf{e}_i \in R^n$ έχει την συνιστώσα αρ. i ίση προς 1 και τις υπόλοιπες ίσες προς 0, ενώ α(k) εκλαμβάνεται ως η **βαθμωτή** κατάσταση της επανάληψης l. Έτσι, χρησιμοποιώντας την (6.5), μπορούμε να ορίσουμε την συνάρτηση

$$\hat{\mathbf{V}}\,[\alpha(k), k] = V[\mathbf{X}[\alpha(k), k], k]. \qquad (6.6)$$

Θέτοντας τώρα α(k+1) = α(k) + c(k), με αρχική κατάσταση α(0) = 0 (γιατί;) και θεωρώντας c(k) ως τον βαθμωτό έλεγχο της επανάληψης l, έχουμε από την (6.5)

$$\mathbf{x}(k+1) = \mathbf{X}[\alpha(k+1), k+1] = \mathbf{X}[\alpha(k) + c(k), k+1]$$

και με την (6.3) έχουμε επίσης

$$\mathbf{u}(k) = \mathbf{F}[\mathbf{x}(k), \mathbf{x}(k+1), k] = \mathbf{U}[\alpha(k), c(k), k, k+1].$$

Συμπερασματικά, οι μεταβλητές κατάστασης **x**(k) και ελέγχου **u**(k) του αρχικού προβλήματος μπορούν να εκφρασθούν ως συναρτήσεις της βαθμωτής κατάστασης α(k) και βαθμωτού ελέγχου c(k) για την επανάληψη l και να αντικατασταθούν σε όλες τις συναρτήσεις του προβλήματος. Έτσι προσδιορίζουμε

$$\hat{\varphi}[\alpha(k), c(k), k] = \varphi[\mathbf{x}(k), \mathbf{u}(k), k] \qquad (6.7)$$

και καταλήγουμε στην ακόλουθη αναδρομική εξίσωση Bellman για κάθε επανάληψη

$$\hat{V}[\alpha(k), k] = \min\{\hat{\varphi}[\alpha(k), c(k), k] + \hat{V}[\alpha(k) + c(k), k+1]\} \qquad (6.8)$$

όπου η ελαχιστοποίηση νοείται για $c(k) \in \mathcal{C}[\alpha(k),k]$. Η επιτρεπτή περιοχή ελέγχου $\mathcal{C}[\alpha(k), k]$ προσδιορίζεται από τους περιορισμούς του αρχικού προβλήματος (3.4), (3.5) μετά από αντικατάσταση των **x**(k), **u**(k), δηλαδή μέσω

$$\hat{\mathbf{h}}\,[\alpha(k), c(k),k] \le \mathbf{0} \qquad (6.9)$$

$$\hat{\mathbf{h}}^x\,[\alpha(k),k] \le \mathbf{0}. \qquad (6.10)$$

Η συνοριακή τιμή για την επίλυση της (6.8) είναι

$$\hat{V}[\alpha(K), K] = \hat{\vartheta}[\alpha(K), K] = \vartheta[X[\alpha(K), K], K] \qquad (6.11)$$

και ο τελικός στόχος

$$\hat{\mathbf{g}}[\alpha(K), K] = \mathbf{g}[\mathbf{X}[\alpha(K), K], K] = \mathbf{0}. \qquad (6.12)$$

Έτσι έχουμε σε κάθε επανάληψη την επίλυση ενός προβλήματος Δυναμικού Προγραμματισμού με βαθμωτή κατάσταση και βαθμωτό έλεγχο και επομένως αποφεύγεται η εκθετική αύξηση του υπολογιστικού φόρτου.

Η σύγκλιση του ΑΔΠ προς ένα σχετικό ελάχιστο μπορεί να αποδειχθεί με σχετικά ασθενείς παραδοχές. Η τάξη σύγκλισης είναι γραμμική, έχουμε δηλαδή για κάποιο A<1

$$\left\|\mathbf{u}^{(l)}(k)-\mathbf{u}^{*}(k)\right\|^{\infty} < A\cdot\left\|\mathbf{u}^{(l-1)}(k)-\mathbf{u}^{*}(k)\right\|^{\infty}. \qquad (6.13)$$

Σαν κριτήριο σύγκλισης μπορεί να χρησιμοποιηθεί το (6.2) εφόσον ικανοποιείται για μια σειρά n διαδοχικών επαναλήψεων. Ας σημειωθεί τέλος, ότι ο ΑΔΠ οδηγεί σε βέλτιστη τροχιά ελέγχου για μια συγκεκριμένη αρχική κατάσταση $\mathbf{x}(0)$ και όχι σε βέλτιστο κανόνα ελέγχου για κάθε $\mathbf{x}(k) \in \mathcal{X}(k)$.

Παράδειγμα 6.2: Θεωρούμε το δίκτυο φυσικών δεξαμενών νερού του Σχ. 6.2 με (γνωστή) εξωτερική εισροή z, ελεγχόμενες εκροές u_1, u_2 και ποσότητες αποθηκευμένου ύδατος x_1, x_2. Έχουμε τις καταστατικές εξισώσεις

$$x_1(k+1) = x_1(k) - u_1(k) + z(k)$$

$$x_2(k+1) = x_2(k) + u_1(k) - u_2(k)$$

και τους περιορισμούς (i = 1, 2)

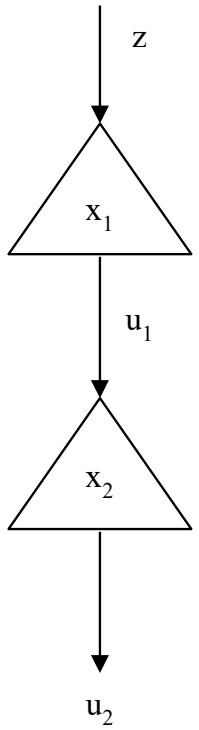
$$0 \le x_i(k) \le X_i \ , \ 0 \le u_i(k) \le U_i.$$

Κάθε εκροή κινεί γεννήτριες παραγωγής ηλεκτρικής ενέργειας σε ποσότητες $\varphi_i[u_i(k)]$ ανά χρονικό διάστημα. Ζητείται ο έλεγχος που μεγιστοποιεί την παραγωγή ηλεκτρικής ενέργειας σε δεδομένη περίοδο K

$$J=\sum_{k=0}^{K-1}\sum_{i=1}^{2}\varphi_i[u_i(k)]$$

με συνοριακές τιμές κατάστασης $\mathbf{x}(0) = \mathbf{x}(K) = \bar{\mathbf{x}}$.

Θεωρούμε το διακριτοποιημένο πρόβλημα με $\Delta x_i = \Delta u_i = 1$, i = 1, 2, και με αριθμητικές τιμές K = 4, $X_1 = X_2 = U_1 = U_2 = 3$, $\bar{x}_1 = \bar{x}_2 = 1$, $\varphi_i(u_i) = \sqrt{u_i}$, i=1, 2, και εισροή



Σχήμα 6.2: Σύστημα δύο φυσικών δεξαμενών.

k	0	1	2	3
z(k)	0	3	1	0

Για την επίλυση μέσω ΑΔΠ αντιστρέφουμε καταρχήν τις καταστατικές εξισώσεις

$u_1(k) = x_1(k) - x_1(k+1) + z(k)$

$u_2(k) = x_2(k) - x_2(k+1) + x_1(k) - x_1(k+1) + z(k)$

και θεωρούμε σαν τροχιά εκκίνησης ελέγχου $u_1^{(0)}(k) = u_2^{(0)}(k) = z(k),\ k = 0,1,2,3,$ με απόδοση $J^{(0)} = 5.46$, που δίδει $x_1^{(0)}(k) = x_2^{(0)}(k) = 1,\ k = 0,...,4.$

Για την 1η, 3η, 5η, ... επανάληψη έχουμε με την (6.5)

$x_1(k) = x_1^{(l-1)}(k) + \alpha(k),\ x_2(k) = x_2^{(l-1)}(k)$

δηλαδή για l = 1

$x_1(k) = 1+\alpha(k),\quad x_2(k) = 1,\quad k = 0,...,4.$

Αντικαθιστώντας στο κριτήριο απόδοσης έχουμε για l = 1

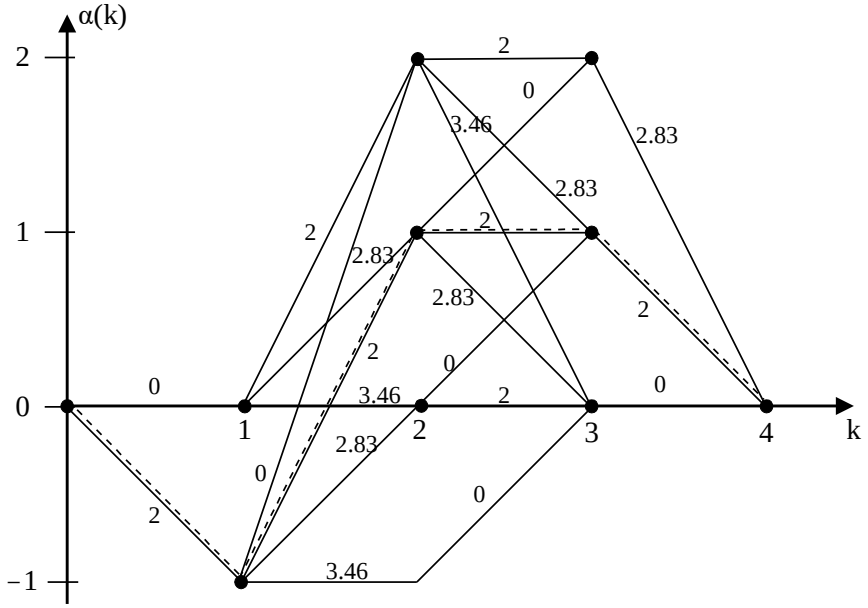
$$\hat{J}^{(1)} = \sum_{k=0}^{3} 2\sqrt{z(k) - c(k)}\,.$$

Επίσης έχουμε $\alpha(k+1) = \alpha(k) + c(k)$ με $\alpha(0) = \alpha(4) = 0$, και τους περιορισμούς

$-1 \le \alpha(k) \le 2$

$z(k)-3 \le c(k) \le z(k).$

Το Σχ. 6.3 δίδει τη γραφική λύση του μονοδιάστατου αυτού προβλήματος με τα αποτελέσματα του Πίνακα 6.1 και απόδοση $\hat{J}^{(1)} = 8$.



Σχήμα 6.3: Γραφική επίλυση 1ης επανάληψης του Παραδ. 6.2.

k	0	1	2	3	4
c(k)	–1	2	0	–1	
α(k)	0	–1	1	1	0
$x_1^{(1)}(k)$	1	0	2	2	1
$x_2^{(1)}(k)$	1	1	1	1	1
$u_1^{(1)}(k)$	1	1	1	1	
$u_2^{(1)}(k)$	1	1	1	1	

Πίνακας 6.1: Αποτελέσματα 1ης επανάληψης του Παραδ. 6.2.

Για την 2η, 4η, 6η, ... επανάληψη έχουμε

$$x_1(k) = x_1^{(l-1)}(k), \quad x_2(k) = x_2^{(l-1)}(k) + \alpha(k).$$

Στο συγκεκριμένο παράδειγμα διαπιστώνουμε πως οι επόμενες επαναλήψεις δεν επιφέρουν περαιτέρω βελτίωση της απόδοσης, δηλαδή ο αλγόριθμος ΑΔΠ έχει συγκλίνει σε μία και μόνη επανάληψη και μάλιστα στο απόλυτο μέγιστο. Εννοείται πως η άμεση αυτή σύγκλιση οφείλεται εδώ στις συγκεκριμένες αριθμητικές τιμές, ενώ κατά κανόνα απαιτούνται περισσότερες επαναλήψεις μέχρι την τελική σύγκλιση.

■

6.4 Διαφορικός Δυναμικός Προγραμματισμός (ΔΙΑΔΠ)

Ο ΔΙΑΔΠ αποφεύγει, όπως και ο ΑΔΠ, την εκθετική αύξηση του υπολογιστικού φόρτου, οδηγεί όμως επί πλέον σε τετραγωνική σύγκλιση σε μια περιοχή γύρω από το ζητούμενο ελάχιστο, δηλαδή ισχύει για κάποιο A < 1

$$\left\|\mathbf{u}^{(l)}(k) - \mathbf{u}^*(k)\right\|^{\infty} \le A[\left\|\mathbf{u}^{(l-1)}(k) - \mathbf{u}^*(k)\right\|^{\infty}]^2. \qquad (6.14)$$

Για ευκολότερη κατανόηση θα παρουσιάσουμε την μέθοδο ΔΙΑΔΠ στην περίπτωση βαθμωτής κατάστασης x(k) και βαθμωτού ελέγχου u(k). Η γενίκευση στην διανυσματική περίπτωση δεν παρουσιάζει μεθοδολογικές δυσκολίες, η παρουσίασή της όμως περιλαμβάνει σχετικά πολύπλοκους διανυσματικούς όρους.

Θα θεωρήσουμε καταρχήν το ΠΒΕ του Κεφαλαίου 3.1 χωρίς περιορισμούς ανισότητας (3.4), (3.5) και χωρίς τελικό στόχο (3.3). Θα κάνουμε επίσης την παραδοχή ότι οι συναρτήσεις ϑ, φ, f είναι επαρκώς παραγωγίσιμες για τους παρακάτω υπολογισμούς, πράγμα που περιορίζει βέβαια αντίστοιχα το πεδίο εφαρμογής της μεθόδου. Για την παρουσίαση του αλγόριθμου ορίζουμε καταρχήν

$$\delta x(k) = x(k) - x^{(l-1)}(k), \quad \delta u(k) = u(k) - u^{(l-1)}(k). \qquad (6.15)$$

Η βασική ιδέα του αλγόριθμου έγκειται στην επίλυση, σε κάθε επανάληψη l, μιας τετραγωνικής-γραμμικής προσέγγισης του αρχικού ΠΒΕ, λαμβανόμενης περί την τελευταία

λύση $x^{(l-1)}(k)$, $u^{(l-1)}(k)$. Το πλεονέκτημα της διαδικασίας αυτής προέρχεται από το γεγονός ότι η επίλυση της τετραγωνικής-γραμμικής προσέγγισης του ΠΒΕ (χωρίς περιορισμούς) είναι δυνατή αναλυτικά, χωρίς την χρήση διακριτοποίησης και άρα χωρίς εκθετική αύξηση του υπολογιστικού φόρτου.

Αρχίζουμε με την τελευταία βαθμίδα K−1, σύμφωνα με τον αλγόριθμο του Κεφαλαίου 3.2, θεωρώντας τον αναδρομικό τύπο (3.9) του Bellman. Ο υπό ελαχιστοποίηση όρος στην (3.9) για k = K−1 είναι $\varphi[x(K-1), u(K-1), K-1] + \vartheta[f[x(K-1), u(K-1), K-1]]$. Η τετραγωνική προσέγγιση του όρου αυτού περί τις τιμές $x^{(l-1)}(K-1)$, $u^{(l-1)}(K-1)$ δίδει

$$Q[x(K-1), u(K-1), K-1] = \frac{1}{2}\, D(K-1)\, \delta x(K-1)^2 + E(K-1)\, \delta x(K-1)\, \delta u(K-1)$$
$$+ \frac{1}{2}\, F(K-1)\, \delta u(K-1)^2 + G(K-1)\, \delta x(K-1) + H(K-1)\, \delta u(K-1) \quad (6.16)$$

όπου παραλείφθηκε ο σταθερός όρος μηδενικού βαθμού που δεν επηρεάζει το αποτέλεσμα της ελαχιστοποίησης. Οι παράγοντες D, E, F, G, H της (6.16) μπορούν εύκολα να υπολογιστούν. Δίδουμε εδώ για παράδειγμα τον τύπο υπολογισμού του D(K−1)

$$D(K-1) = \frac{\partial^2 \varphi[x(K-1), u(K-1), K-1]}{\partial x(K-1)^2} + \frac{\partial^2 \vartheta[x(K)]}{\partial x(K)^2}\left[\frac{\partial f[x(K-1), u(K-1), K-1]}{\partial x(K-1)}\right]^2$$
$$+ \frac{\partial \vartheta[x(K)]}{\partial x(K)}\, \frac{\partial^2 f[x(K-1), u(K-1), K-1]}{\partial x(K-1)^2} \quad (6.17)$$

όπου όλοι οι όροι υπολογίζονται για $x^{(l-1)}(k)$, $u^{(l-1)}(k)$.

Η ελαχιστοποίηση της τετραγωνικής προσέγγισης Q στην (6.16) ως προς u(K−1) δίδει

$$u(K-1) = u^{(l-1)}(K-1) + \delta u(K-1) \quad (6.18)$$

όπου

$$\delta u(K-1) = \alpha(K-1) + \beta(K-1)\, \delta x(K-1) \quad (6.19)$$

με

$$\alpha(K-1) = -F(K-1)^{-1}\, H(K-1), \quad \beta(K-1) = -F(K-1)^{-1}\, E(K-1). \quad (6.20)$$

Έτσι, η κατά προσέγγιση συνάρτηση $\hat{V}[x(K-1), K-1]$ (βλ. (3.9)) δίδεται (χωρίς τον σταθερό όρο) από την ελάχιστη τιμή της Q ως εξής

$$\hat{V}[x(K-1), K-1] = \frac{1}{2} A(K-1)\, \delta x(K-1)^2 + B(K-1)\, \delta x(K-1) \quad (6.21)$$

όπου

$$A(K-1) = D(K-1) - E(K-1)\, F(K-1)^{-1}\, E(K-1) \quad (6.22)$$

$$B(K-1) = G(K-1) - E(K-1)\, F(K-1)^{-1}\, H(K-1). \qquad (6.23)$$

Οι υπολογισμοί των κατωτέρων βαθμίδων k γίνονται κατά παρόμοιο τρόπο. Ας υποθέσουμε ότι έχει ήδη υπολογισθεί η κατά τετραγωνική προσέγγιση συνάρτηση $\hat{V}[x(k+1), k+1] = (1/2)A(k+1)\delta x(k+1)^2 + B(k+1)\delta x(k+1)$. Ορίζοντας $Q[x(k), u(k), k]$ σαν τετραγωνική προσέγγιση του όρου $\varphi[x(k), u(k), k] + \hat{V}[f[x(k), u(k), k], k+1]$ περί τις $x^{(l-1)}(k)$, $u^{(l-1)}(k)$ (πάντα χωρίς τον σταθερό όρο), έχουμε την αντίστοιχη σχέση της (6.16) για το διάστημα k

$$Q[x(k), u(k), k] = \frac{1}{2} D(k)\, \delta x(k)^2 + E(k)\, \delta x(k)\, \delta u(k) + \frac{1}{2} F(k)\, \delta u(k)^2 + G(k)\, \delta x(k) + H(k)\, \delta u(k) \qquad (6.24)$$

όπου οι παράγοντες D, E, F, G, H μπορούν εύκολα να υπολογισθούν, επί παραδείγματι

$$D(k) = \frac{\partial^2 \varphi[x(k),u(k),k]}{\partial x(k)^2} + A(k+1)\left(\frac{\partial f[x(k),u(k),k]}{\partial x(k)}\right)^2 + B(k+1)\frac{\partial^2 f[x(k),u(k),k]}{\partial x(k)^2} \qquad (6.25)$$

όπου όλοι οι όροι υπολογίζονται για $x^{(l-1)}(k)$, $u^{(l-1)}(k)$.

Με αυτούς τους παράγοντες μπορούμε τώρα να ελαχιστοποιήσουμε την Q, όπως και στο διάστημα K−1, ως προς u(k) και έχουμε

$$u^{(l)}(k) = u^{(l-1)}(k) + \alpha(k) + \beta(k)\, \delta x(k) \qquad (6.26)$$

όπου α(k), β(k) υπολογίζονται όπως στην (6.20). Από την ελαχιστοποίηση λαμβάνουμε επίσης την προσέγγιση $\hat{V}[x(k),k] = (1/2)\, A(k)\, \delta x(k)^2 + B(k)\, \delta x(k)$ όπου A(k), B(k) υπολογίζονται όπως στις (6.22), (6.23). Οι παράγοντες α(k), β(k) πρέπει να αποθηκεύονται σε κάθε βαθμίδα, ενώ οι A(k+1), B(k+1) δεν χρειάζονται πλέον και μπορούν να σβήνονται από τη μνήμη του υπολογιστού.

Μετά τους υπολογισμούς όλων των βαθμίδων μπορεί να παραχθεί ο νέος, βελτιωμένος έλεγχος της επανάληψης l

$$u^{(l)}(k) = u^{(l-1)}(k) + \alpha(k) + \beta(k)\, \delta x(k) \qquad (6.27)$$

$$x^{(l)}(k+1) = f[x^{(l)}(k), u^{(l)}(k), k], \quad x(0) = x_0. \qquad (6.28)$$

Σε μερικές περιπτώσεις χρησιμοποιείται αντί της (6.27) η

$$u^{(l)}(k) = u^{(l-1)}(k) + \varepsilon\, [\alpha(k) + \beta(k)\, \delta x(k)] \qquad (6.29)$$

με $0 < \varepsilon \leq 1$ για να αποφευχθεί απόκλιση του αλγόριθμου. Και με το μέτρο αυτό όμως μπορεί να μην επιτυγχάνεται σύγκλιση σε μερικές περιπτώσεις οπότε χρειάζεται η λήψη επί πλέον μέτρων.

Η μέθοδος του ΔΙΑΔΠ μπορεί να επεκταθεί σε προβλήματα με γραμμικούς περιορισμούς ανισότητας. Στην περίπτωση αυτή επιλύεται σε κάθε βαθμίδα κάθε επανάληψης ένα πρόβλημα Τετραγωνικού Προγραμματισμού, πράγμα που αυξάνει τον υπολογιστικό φόρτο, διατηρείται όμως η αύξηση του υπολογιστικού φόρτου ως προς τις διαστάσεις του προβλήματος κατά τρόπο πολυωνυμικό και όχι εκθετικό.

Τα μειονεκτήματα του ΔΙΑΔΠ σε σχέση με τον ΔΔΠ αφορούν:
- Την δυνατότητα **μη σύγκλισης**
- Την δυνατότητα σύγκλισης προς **σχετικό ελάχιστο**
- Την δυσκολία να ληφθούν υπόψη γενικοί περιορισμοί ανισότητας ή διακριτά επιτρεπτά πεδία τιμών.

Σημειώνουμε ότι ο ΔΙΑΔΠ επιλύει το αρχικό ΠΒΕ χωρίς την χρήση διακριτοποίησης, αφού το πρόβλημα τετραγωνικής προσέγγισης κάθε επανάληψης επιλύεται αναλυτικά. Τέλος, ο ΔΙΑΔΠ παρέχει μέσω της (6.27) μία βέλτιστη τροχιά και όχι έναν κανόνα ελέγχου.

6.5 Συγκριτικό Παράδειγμα

Θα εξετάσουμε εδώ, βάσει ενός απλού προβλήματος, την δυνατότητα μείωσης του υπολογιστικού φόρτου που προσφέρουν διάφορες απλοποιήσεις ή παραλλαγές του Δυναμικού Προγραμματισμού. Θεωρούμε το ΠΒΕ:

Ελαχιστοποίηση

$$J = \sum_{k=0}^{9} [x_1(k)^2 + 0.5x_2(k)^2 + u_1(k)^2 + u_2(k)^2] \qquad (6.30)$$

υπό την καταστατική εξίσωση

$$x_1(k+1) = x_1(k) + 0.6321\, x_2(k) + 0.3679\, u_1(k) + u_2(k) \qquad (6.31)$$

$$x_2(k+1) = 0.3679\, x_2(k) + 0.6321\, u_1(k) \qquad (6.32)$$

και τους περιορισμούς

$$0 \le x_1(k) \le 2, \quad 0 \le x_2(k) \le 2 \qquad (6.33)$$

$$|u_1(k)| \le 1, \quad |u_2(k)| \le 1 \qquad (6.34)$$

με αρχική κατάσταση

$\mathbf{x}(0) = [1 \;\; 1]^T$

και τελικό στόχο

$\mathbf{x}(10) = [1 \;\; 0]^T$.

Σημειώνουμε ότι η γραμμική καταστατική εξίσωση (6.31), (6.32) είναι εύκολα αντιστρέψιμη ως προς $\mathbf{u}(k)$ κατά τις (3.18) ή (6.3), πράγμα που επιτρέπει την άμεση εφαρμογή του ΑΔΠ του Κεφαλαίου 6.3.

Θα επιχειρήσουμε καταρχήν την εφαρμογή του αλγορίθμου του Διακριτού Δυναμικού Προγραμματισμού (Κεφ. 3.3), χωρίς απλοποιήσεις ή παραλλαγές, με διακριτοποιήσεις Δx_i=0.5 και Δx_i=0.1, i=1,2. Διακριτοποίηση της επιτρεπτής περιοχής ελέγχου και γραμμική παρεμβολή

δεν είναι αναγκαίες λόγω αντιστρεψιμότητας της καταστατικής εξίσωσης κατά την (3.18). Ο Πίνακας 6.2 συνοψίζει τα αποτελέσματα λύσης και υπολογιστικού φόρτου για τις δύο διακριτοποιήσεις. Ο αριθμός α κατά το Κεφάλαιο 6.3γ αναφέρεται στον αριθμό επιτρεπτών διακριτών σημείων ανά μεταβλητή κατάστασης, ενώ α^2 είναι εδώ ο αριθμός διακριτών σημείων ανά βαθμίδα (γιατί;). Ο συνολικός αριθμός μεταβάσεων (χωρίς να λαμβάνεται υπόψη η προσβασιμότητα σημείων κατάστασης λόγω των περιορισμών ελέγχου (6.34)) είναι κατά την (3.20) $K\cdot\alpha^2\alpha^2 = K\cdot\alpha^4$. Όπως φαίνεται στον Πίνακα 6.2, η μικρότερη διακριτοποίηση οδηγεί σε καλύτερο κριτήριο κόστους, αλλά ο υπολογιστικός φόρτος (αριθμός μεταβάσεων) είναι 311 φορές μεγαλύτερος.

Σε μια πρώτη επιχείρηση λύσης με μειωμένο υπολογιστικό φόρτο θεωρούμε μόνον σημεία $\mathbf{x}(k)$, $k=1,\ldots,K-1$, που, υπό επιτρεπτό έλεγχο $\mathbf{u}(k)$, $k=0,\ldots,K-1$,
(i) είναι προσβάσιμα από την αρχική κατάσταση $\mathbf{x}(0)$ του προβλήματος και
(ii) μπορούν να οδηγήσουν στην τελική κατάσταση $\mathbf{x}(K)$ του προβλήματος.
Υπό τις συνθήκες αυτές ο συνολικός αριθμός μεταβάσεων, για διακριτοποίηση $\Delta x_i=0.1$, μειώνεται σε 419,694, δηλαδή στο 21.6% εκείνου του Πίνακα 6.2.

Θα επιλύσουμε τώρα το πρόβλημα με ΔΔΔΠ κατά το Κεφάλαιο 6.2, εφαρμόζοντας καταρχήν διακριτοποίηση $\Delta x_i=0.5$. Αυτό οδηγεί συνολικά σε 5,050 μεταβάσεις (αν ληφθούν υπόψη μόνον τα προσβάσιμα σημεία κατάστασης). Στη συνέχεια εφαρμόζουμε διακριτοποίηση $\Delta x_i=0.1$, σε ένα διάδρομο πλάτους 0.5 περί την πρώτη λύση, πράγμα που οδηγεί σε ένα νέο πρόβλημα με ένα σύνολο 12,660 μεταβάσεων προς αξιολόγηση. Η λύση της δεύτερης αυτής επανάληψης οδηγεί σε J=3.1566 με συνολικό αριθμό μεταβάσεων των δύο επαναλήψεων ίσο προς 17,710, δηλαδή μόνον 0,9% αυτού που απαιτήθηκε κατά τον Πίνακα 6.2 για $\Delta x_i=0.1$. Οι επαναλήψεις μπορούν να συνεχισθούν αν απαιτείται υψηλότερη ακρίβεια λύσης.

Θα εφαρμόσουμε τέλος ΑΔΠ (Κεφάλαιο 6.3) για την επίλυση του προβλήματος. Τα υπολογιζόμενα αποτελέσματα της τρίτης και τέταρτης επανάληψης είναι ταυτόσημα, πράγμα που σημαίνει ότι η μέθοδος έχει συγκλίνει. Ο απαιτούμενος συνολικός αριθμός μεταβάσεων είναι 1,000 και 26,460 για αντίστοιχες διακριτοποιήσεις $\Delta x_i=0.5$ και $\Delta x_i=0.1$, αντιστοιχεί δηλαδή στο 16% και 1.4% των αντίστοιχων αριθμών του Πίνακα 6.2. Η διακριτοποίηση $\Delta x_i=0.1$ οδηγεί σε μεγαλύτερο ποσοστό μείωσης του υπολογιστικού φόρτου, ο οποίος όμως παραμένει αισθητά υψηλότερος σε απόλυτες τιμές σχετικά με την διακριτοποίηση $\Delta x_i=0.5$.

Δx_i	α (βλ. (3.20))	α^2	αριθμός μεταβάσεων	J
0.5	5	25	6,250	3.4081
0.1	21	441	1,944,810	3.1566

Πίνακας 6.2: Αποτελέσματα Διακριτού Δυναμικού Προγραμματισμού.

Προβλήματα

6.1 Ζητούμε να προσδιορίσουμε τον έλεγχο που μεγιστοποιεί το κριτήριο

$$J = \sum_{k=0}^{K-1} \sqrt{u_1(k) + u_2(k)}$$

λαμβάνοντας υπόψη τις καταστατικές εξισώσεις

$$x_1(k+1) = x_1(k) - u_1(k) + u_2(k)$$
$$x_2(k+1) = x_2(k) - u_2(k) + z(k)$$

τους περιορισμούς $(i = 1, 2)$

$$0 \le x_i(k) \le 3$$
$$0 \le u_i(k) \le 4$$

και τις συνοριακές τιμές κατάστασης $\mathbf{x}(0) = \mathbf{x}(K) = [1 \;\; 1]^T$.

Θεωρώντας ότι K = 3 και ότι το z(k) δίνεται από τον πίνακα

k	0	1	2
z(k)	0	4	1

θα επιχειρηθεί επίλυση του προβλήματος με χρήση Αυξάνοντος Δυναμικού Προγραμματισμού (ΑΔΠ) με τροχιά εκκίνησης ελέγχου $u_1^0(k) = u_2^0(k) = z(k)$ για k=0,1,2. Συγκεκριμένα, ζητείται να εκτελεσθεί μόνον η $1^η$ επανάληψη της μεθόδου κατά την οποία θα γίνει αλλαγή της μεταβλητής κατάστασης x_1.

(a) Εκφράστε (για την ως άνω $1^η$ επανάληψη) τις μεταβλητές κατάστασης **x**(k) και ελέγχου **u**(k) του προβλήματος σαν συναρτήσεις των βαθμωτών μεταβλητών α(k) κατάστασης και c(k) ελέγχου, προσδιορίστε το κριτήριο κόστους σαν συνάρτηση των α(k) και c(k) και καθορίστε τους περιορισμούς των βαθμωτών μεταβλητών α(k) και c(k).

(b) Επιλύστε γραφικά το προκύπτον μονοδιάστατο πρόβλημα μεγιστοποίησης.

6.2 Ζητούμε να προσδιορίσουμε τον έλεγχο που μεγιστοποιεί το κριτήριο κόστους

$$J = \sum_{k=0}^{K-1} \sqrt{u_1(k) + u_2(k)}$$

λαμβάνοντας υπόψη τις καταστατικές εξισώσεις

$$x_1(k+1) = x_1(k) - u_1(k) + u_2(k)$$
$$x_2(k+1) = x_2(k) - u_2(k) + z(k)$$

τους περιορισμούς $(i = 1, 2)$

$$0 \le x_i(k) \le 3$$
$$0 \le u_i(k) \le 4$$

και τις συνοριακές τιμές κατάστασης $\mathbf{x}(0) = \mathbf{x}(K) = [1 \;\; 1]^T$.

Θεωρώντας ότι K = 3 και ότι το z(k) δίνεται από τον πίνακα

k	0	1	2
z(k)	0	4	1

επιχειρείται η επίλυση του προβλήματος με χρήση Αυξάνοντος Δυναμικού Προγραμματισμού (ΑΔΠ). Συγκεκριμένα, έχει εκτελεστεί η πρώτη επανάληψη του αλγορίθμου με $\Delta x_i=\Delta u_i=1$, i=1, 2 (κατά την οποία έγινε αλλαγή της μεταβλητής κατάστασης $x_1(k)$) και έχει υπολογισθεί η βέλτιστη τροχιά ελέγχου

$$u_1^{(1)}(0)=1,\ \ u_1^{(1)}(1)=1,\ \ u_1^{(1)}(2)=3$$
$$u_2^{(1)}(0)=1,\ \ u_2^{(1)}(1)=4,\ \ u_2^{(1)}(2)=1.$$

Ζητείται να εκτελεστεί μόνον η 2η επανάληψη της μεθόδου στην οποία θα γίνει αλλαγή της μεταβλητής κατάστασης $x_2(k)$.

(a) Υπολογίστε για την πρώτη επανάληψη του αλγορίθμου τις τιμές των μεταβλητών κατάστασης $x_1(k)$, $x_2(k)$, k=0,1,2,3, καθώς και τη βέλτιστη τιμή του αντικειμενικού κριτηρίου.

(b) Εκφράστε για τη 2η επανάληψη τις μεταβλητές κατάστασης x(k) και ελέγχου u(k) του προβλήματος σαν συναρτήσεις της βαθμωτής μεταβλητής κατάστασης α(k) και ελέγχου c(k). Προσδιορίστε το κριτήριο κόστους σαν συνάρτηση των α(k) και c(k) και καθορίστε τους περιορισμούς των βαθμωτών μεταβλητών α(k) και c(k).

(c) Επιλύστε γραφικά το προκύπτον μονοδιάστατο πρόβλημα μεγιστοποίησης και στη συνέχεια συμπληρώστε τον παρακάτω πίνακα.

k	0	1	2	3
c(k)				
α(k)				
$x_1^{(2)}(k)$				
$x_2^{(2)}(k)$				
$u_1^{(2)}(k)$				
$u_2^{(2)}(k)$				

6.3 Θεωρούμε το δυναμικό σύστημα

$$x(k+1)=x(k)+u(k),\ \ x(0)=3$$

με τους περιορισμούς

$$u(k)\in U=\{1,0,-1\}$$

$$x(k)\in X=\{0,1,2,3,4,5\}.$$

Ζητείται η βέλτιστη τροχιά ελέγχου για τη μεταφορά της αρχικής κατάστασης σε ελάχιστο χρόνο στην τελική τροχιά που δίδεται από τα σημεία

$$[x(K),K]\in\{(5,3),(4,4),(3,5),(2,5),(1,5)\}.$$

(a) Διατυπώστε το πρόβλημα ελάχιστου χρόνου με το κριτήριο κόστους, την καταστατική εξίσωση και όλους τους περιορισμούς.

(b) Προσδιορίστε όλες τις επιτρεπτές μεταβάσεις καθώς και το βέλτιστο κανόνα ελέγχου του προβλήματος, μέσω του αλγορίθμου Δυναμικού Προγραμματισμού.

(c) Θεωρήστε τώρα την επίλυση του προβλήματος μέσω Διακριτού Διαφορικού Δυναμικού Προγραμματισμού (ΔΔΔΠ). Λαμβάνοντας υπόψη την τροχιά εκκίνησης $u^{(0)}(0)=0,\ u^{(0)}(1)=0,\ u^{(0)}(2)=0,\ u^{(0)}(3)=0,\ u^{(0)}(4)=0$ και εύρος διαδρόμου ίσο με 1, εκτελέστε 2 επαναλήψεις (l =1,2) του αλγορίθμου ΔΔΔΠ. (**Σημείωση:** σε περίπτωση εύρεσης πολλαπλών βέλτιστων διαδρομών σε κάποια από τις επαναλήψεις του αλγορίθμου, επιλέξτε αυτή που περιλαμβάνει μεγαλύτερες τιμές της μεταβλητής κατάστασης.)

7. ΣΤΟΧΑΣΤΙΚΟΣ ΔΥΝΑΜΙΚΟΣ ΠΡΟΓΡΑΜΜΑΤΙΣΜΟΣ

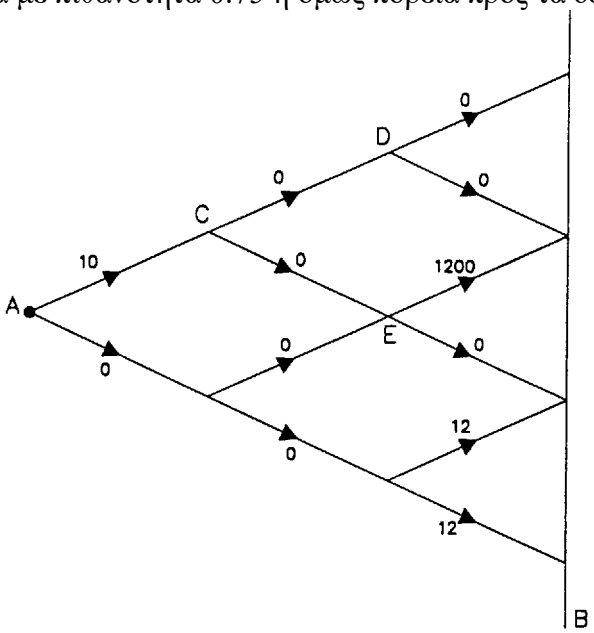
7.1 Εισαγωγή

Μέχρι το παρόν σημείο ασχοληθήκαμε με προβλήματα βελτιστοποίησης, στα οποία όλα τα εμφανιζόμενα μεγέθη, εκτός των ζητούμενων βέλτιστων τροχιών, θεωρούνταν γνωστά. Σε πολλές πρακτικές εφαρμογές (ιδιαίτερα μελλοντικού σχεδιασμού) παρουσιάζεται όμως μια αβεβαιότητα σε μερικά μεγέθη (π.χ. μελλοντική ζήτηση ενός προϊόντος ή μελλοντικές εισροές μιας δεξαμενής ύδατος κλπ.) που συχνά μπορεί να εκφρασθεί με **στοχαστικές** ή **τυχαίες μεταβλητές**. Οι ακριβείς (π.χ. μελλοντικές) τιμές των στοχαστικών μεταβλητών είναι μεν άγνωστες κατά την επίλυση του προβλήματος βελτιστοποίησης, μπορούν όμως να γίνουν ρεαλιστικές παραδοχές ως προς την **πιθανολογική κατανομή** των τιμών αυτών. Το πρόβλημα βελτιστοποίησης έγκειται τότε στον προσδιορισμό βέλτιστου ελέγχου για την ελαχιστοποίηση της **αναμενόμενης τιμής του κριτηρίου κόστους** λαμβάνοντας υπόψη τα συνήθη δεδομένα του προβλήματος και την πιθανολογική κατανομή των τυχαίων μεταβλητών.

Η επίλυση στοχαστικών προβλημάτων βελτιστοποίησης οδηγεί εν μέρει σε αποτελέσματα που δεν έχουν αντιστοιχία στην αιτιοκρατική βελτιστοποίηση. Το ακόλουθο παράδειγμα καταδεικνύει μερικές από αυτές τις ιδιαιτερότητες.

Παράδειγμα 7.1: Θεωρούμε το πρόβλημα αποφάσεων πολλαπλών βαθμίδων του Σχ. 7.1. Ο Δυναμικός Προγραμματισμός μπορεί ως γνωστόν να εφαρμοσθεί για τον υπολογισμό της βέλτιστης τροχιάς από το σημείο Α στον τελικό στόχο Β. Η λύση δίδει όχι απλώς την βέλτιστη τροχιά από Α προς Β, αλλά τον βέλτιστο κανόνα ελέγχου που εμπεριέχει την βέλτιστη ακολουθητέα κατεύθυνση (**Α**ριστερά ή **Δ**εξιά) σε κάθε κόμβο του δικτύου. Εν απουσία μη αναμενόμενων διαταραχών οδηγούν, ως γνωστόν, τόσο ο βέλτιστος έλεγχος ανοιχτού βρόχου όσο και ο βέλτιστος κανόνας ελέγχου στην ίδια ελάχιστη τιμή του κριτηρίου κόστους.

Θα μεταλλάξουμε τώρα το πρόβλημα αυτό εισάγοντας αβεβαιότητα ως εξής. Όταν σε κάποιο κόμβο του δικτύου λαμβάνεται η απόφαση **Α** (αριστερή κατεύθυνση) τότε υπάρχει όντως πορεία προς τα αριστερά με πιθανότητα 0.75 ή όμως πορεία προς τα δεξιά με πιθανότητα 0.25



Σχήμα 7.1: Πρόβλημα αποφάσεων πολλαπλών βαθμίδων.

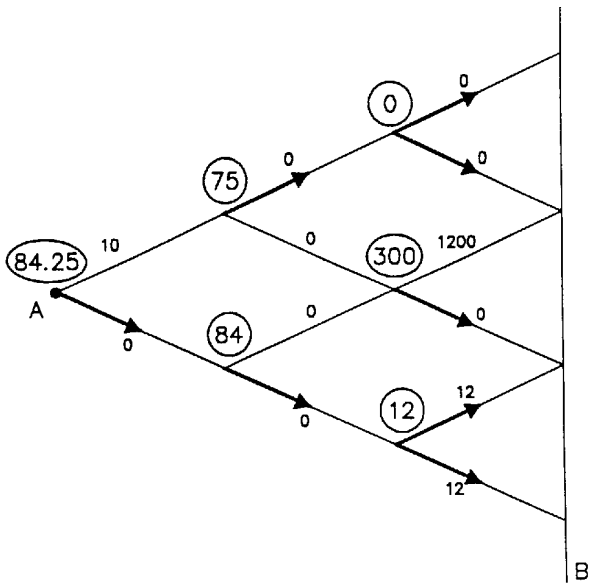
(π.χ. επειδή ο οδηγός του αντίστοιχου οχήματος είναι αφηρημένος). Αν πάλι λαμβάνεται η απόφαση **Δ** (δεξιά κατεύθυνση) τότε έχουμε πιθανότητα δεξιάς πορείας 0.75 και αριστερής πορείας 0.25. Υπό τις συνθήκες αυτές μπορούμε βέβαια να επηρεάσουμε την τροχιά του οχήματος από Α προς Β καθώς και το αντίστοιχο κόστος, δεν μπορούμε όμως να τα καθορίσουμε με βεβαιότητα. Πράγματι, κάθε τροχιά από Α σε Β είναι δυνατή υπό οποιονδήποτε έλεγχο, με διαφορετικές όμως πιθανότητες που εξαρτώνται από τις τιμές του ελέγχου. Μια πολιτική στο στοχαστικό (αβέβαιο) αυτό περιβάλλον μπορεί να θεωρηθεί βέλτιστη αν ελαχιστοποιεί την **αναμενόμενη τιμή** του κόστους μεταφοράς από Α προς Β λαμβάνοντας υπόψη τις παραπάνω πιθανότητες.

Για τον υπολογισμό της **βέλτιστης τροχιάς ελέγχου (ανοιχτού βρόχου)** μπορούμε να θεωρήσουμε τις 8 δυνατές τροχιές (πολιτικές) από Α προς Β και να επιλέξουμε εκείνη που ελαχιστοποιεί το αναμενόμενο κόστος μεταφοράς. Για παράδειγμα, η πολιτική **Α-Α-Α** οδηγεί με πιθανότητα 27/64 στην διαδρομή **Α-Α-Α** με κόστος 10, με πιθανότητα 9/64 στην διαδρομή **Δ-Α-Α** με κόστος 1200 κλπ. Η αναμενόμενη τιμή του κόστους μεταφοράς για την πολιτική **Α-Α-Α** υπολογίζεται έτσι

$$J_{AAA} = \frac{27}{64}10 + \frac{9}{64}(1200 + 1210 + 10) + \frac{3}{64}(10 + 0 + 12) + \frac{1}{64}12 = 345.75.$$

Εκτελώντας τους αντίστοιχους υπολογισμούς για κάθε δυνατή πολιτική, καταλήγουμε στην βέλτιστη τροχιά ελέγχου **Α-Α-Δ** με το ελάχιστο αναμενόμενο κόστος 120.75.

Για τον προσδιορισμό ενός **βέλτιστου κανόνα ελέγχου** θα εφαρμόσουμε την μέθοδο του Κεφαλαίου 2, αρχίζοντας δηλαδή από τον στόχο Β και επιλέγοντας, προχωρώντας όπισθεν βαθμίδα προς βαθμίδα, την βέλτιστη κατεύθυνση σε κάθε κόμβο. Ας υποθέσουμε για παράδειγμα ότι βρισκόμαστε στο σημείο C (Σχ. 7.1) της βαθμίδας k και ότι γνωρίζουμε τα ελάχιστα αναμενόμενα κόστη V_D, V_E των δύο επόμενων κόμβων D, E της βαθμίδας k+1. Το αναμενόμενο κόστος V_C στο C με απόφαση **Α** είναι τότε 0.75 V_D + 0.25 V_E ενώ με απόφαση **Δ** είναι 0.25 V_D + 0.75 V_E. Η μικρότερη των δύο αυτών αναμενόμενων τιμών δίδει το V_C και η αντίστοιχη απόφαση είναι η βέλτιστη στον κόμβο C. Τα αποτελέσματα των υπολογισμών αυτών με τις βέλτιστες κατευθύνσεις και τις ελάχιστες αναμενόμενες τιμές κόστους (σε κύκλο) δείχνει το Σχ. 7.2.



Σχήμα 7.2: Βέλτιστος κανόνας ελέγχου.

Το ελάχιστο αναμενόμενο κόστος μεταφοράς από Α προς Β είναι επομένως για τον βέλτιστο κανόνα ελέγχου $V_A = 84.25$, είναι δηλαδή μικρότερο από το αναμενόμενο κόστος της βέλτιστης πολιτικής. Το χειρότερο αποτέλεσμα της βέλτιστης τροχιάς ελέγχου οφείλεται στο γεγονός ότι η σειρά αποφάσεων (τροχιά ελέγχου), που προσδιορίστηκε στη βάση των δεδομένων πιθανοτήτων, εφαρμόζεται απαραβίαστα, χωρίς δηλαδή να ληφθεί υπόψη η **πραγματική** τροχιά της κατάστασης. Εν αντιθέσει, ο βέλτιστος κανόνας ελέγχου είναι εύκαμπτος καθότι εξαρτά την απόφαση κάθε βαθμίδας από την **πραγματική** κατάσταση του συστήματος.

Η διαπίστωση αυτή μας οδηγεί στην ακόλουθη, τρίτη, δυνατή μεθοδολογία λύσης. Υπολογίζουμε καταρχήν στο σημείο Α μια βέλτιστη τροχιά ελέγχου, **Α-Α-Δ**. Στην συνέχεια εφαρμόζουμε την πρώτη απόφαση, δηλαδή **Α**, και αναμένουμε την μετάβαση στην επόμενη βαθμίδα. Από τον κόμβο (κατάσταση) που βρισκόμαστε εκεί, υπολογίζουμε μια βέλτιστη τροχιά ελέγχου προς Β και ούτω καθεξής. Η μεθοδολογία αυτή είναι προφανώς μεικτή. Υπολογίζει δηλαδή βέλτιστη τροχιά ελέγχου, όχι όμως μια μόνο φορά (στο σημείο Α), αλλά κατ΄ επανάληψη, σε κάθε ενδιάμεσο κόμβο μετάβασης. Ονομάζουμε την μεθοδολογία αυτή **βέλτιστη τροχιά ελέγχου κλειστού βρόχου**. Το αναμενόμενο κόστος στο παρόν Παράδειγμα εφαρμόζοντας την μέθοδο αυτή είναι 84.75, είναι δηλαδή μικρότερο από το κόστος της βέλτιστης τροχιάς ελέγχου αλλά μεγαλύτερο από το κόστος του βέλτιστου κανόνα ελέγχου.

■

Το παραπάνω παράδειγμα κατέδειξε τρεις δυνατές μεθόδους για την επίλυση στοχαστικών προβλημάτων αποφάσεων πολλαπλών βαθμίδων:

(a) **Βέλτιστη τροχιά ελέγχου (ΒΤΕ)** που εφαρμόζεται απαρέγκλιτα, χωρίς χρήση πραγματικών μετρήσεων κατάστασης.

(b) **Βέλτιστος κανόνας ελέγχου (ΒΚΕ)** που βασίζεται στην χρήση πραγματικών μετρήσεων κατάστασης.

(c) **Βέλτιστη τροχιά ελέγχου κλειστού βρόχου (ΒΤΕΚΒ)** που κάνει μεν χρήση πραγματικών μετρήσεων κατάστασης, χωρίς όμως το γεγονός αυτό να περιλαμβάνεται στην διατύπωση και λύση του μαθηματικού προβλήματος.

Τα αποτελέσματα είναι κάλλιστα για ΒΚΕ και ακολουθεί ΒΤΕΚΒ και ΒΤΕ. Η ΒΤΕΚΒ οδηγεί προφανώς σε καλύτερα αποτελέσματα από την ΒΤΕ λόγω της χρήσης μετρήσεων σε πραγματικό χρόνο και αντίστοιχης βελτίωσης της τροχιάς ελέγχου. Γιατί όμως είναι η ΒΤΕΚΒ χειρότερη από τον ΒΚΕ; Αυτό οφείλεται στο γεγονός ότι και οι δύο μέθοδοι κάνουν μεν χρήση πραγματικών μετρήσεων, στον ΒΚΕ όμως το γεγονός αυτό περιλαμβάνεται στην μαθηματική διατύπωση του προβλήματος και λαμβάνεται έτσι υπόψη κατά τη λύση. Να σημειωθεί ότι ΒΚΕ και ΒΤΕΚΒ οδηγούν σε ταυτόσημα αποτελέσματα σε μερικές ειδικές κατηγορίες προβλημάτων.

7.2 Το Πρόβλημα Στοχαστικού Βέλτιστου Ελέγχου (ΠΣΒΕ)

Στο κεφάλαιο αυτό θα διατυπώσουμε ένα γενικό **Πρόβλημα Στοχαστικού Βέλτιστου Ελέγχου (ΠΣΒΕ)** περιοριζόμενοι στην περίπτωση διακριτού χρόνου. Ζητούμενο του προβλήματος θα είναι ένας (στοχαστικός) ΒΚΕ αφού έτσι επιτυγχάνουμε τα καλύτερα αποτελέσματα σε στοχαστικό περιβάλλον.

Θεωρούμε το δυναμικό σύστημα διακριτού χρόνου

$$\mathbf{x}(k+1) = \mathbf{f}[\mathbf{x}(k), \mathbf{u}(k), \mathbf{z}(k), k], \quad k = 0,...,K-1 \qquad (7.1)$$

με διανύσματα κατάστασης $\mathbf{x}(k)\in\mathcal{X}(k)\subset R^n$, ελέγχου $\mathbf{u}(k)\in\mathcal{U}[\mathbf{x}(k),k]\subset R^m$ και στοχαστικών διαταραχών $\mathbf{z}(k)\in\mathcal{Z}(k)\subset R^p$. Θεωρούμε ότι το διάνυσμα διαταραχών $\mathbf{z}(k)$ παίρνει τιμές από την περιοχή $\mathcal{Z}(k)$ σύμφωνα με μια γνωστή συνάρτηση πιθανολογικής κατανομής $P[\mathbf{z}\,|\,\mathbf{x}(k), \mathbf{u}(k), k]$ που μπορεί να εξαρτάται από την κατάσταση, τον έλεγχο και το χρόνο. Η τιμή $\mathbf{z}(k)$ θεωρείται ανεξάρτητη από προηγούμενες τιμές $\mathbf{z}(k-1)$, $\mathbf{z}(k-2)$,...

Ένας **κανόνας ελέγχου** του συστήματος (7.1)

$$\mathbf{u}(k) = \mathbf{R}[\mathbf{x}(k), k], \quad k = 0,...,K-1 \qquad (7.2)$$

λέγεται **επιτρεπτός**, αν δίδει για κάθε $\mathbf{x}(k)\in\mathcal{X}(k)$ μιαν επιτρεπτή τιμή ελέγχου $\mathbf{u}(k)\in\mathcal{U}[\mathbf{x}(k), k]$. Παρατηρούμε επομένως ότι στο παρόν πρόβλημα η $\mathcal{X}(k)$ θεωρείται απλά σαν περιοχή τιμών που απορρέει από την φύση της κατάστασης $\mathbf{x}(k)$ και όχι σαν περιορισμός κατάστασης όπως στην αιτιοκρατική περίπτωση. Πράγματι, στην στοχαστική περίπτωση δεν είναι δυνατό να εγγυηθούμε μιαν επιτρεπτή περιοχή του $\mathbf{x}$, αφού οι στοχαστικές διαταραχές $\mathbf{z}$ οδηγούν, μέσω της καταστατικής εξίσωσης (7.1), σε στοχαστικές τιμές του $\mathbf{x}$, που είναι αδύνατο να ελεγχθούν απόλυτα.

Θεωρούμε τέλος την αναμενόμενη τιμή ως προς $\mathbf{z}(k)$, $k = 0,...,K-1$, ενός κριτηρίου κόστους διακριτού χρόνου

$$J = E\{\vartheta[\mathbf{x}(K)] + \sum_{k=0}^{K-1} \varphi[\mathbf{x}(k), \mathbf{u}(k), \mathbf{z}(k), k]\} \qquad (7.3)$$

με δεδομένο τελικό χρόνο K. Σημειώνουμε εδώ την απουσία τελικού στόχου, αφού, για λόγους που αναπτύχθηκαν παραπάνω, οι στοχαστικές τιμές της κατάστασης $\mathbf{x}(K)$ δεν είναι δυνατόν να ελεγχθούν πλήρως. Το βασικό **πρόβλημα στοχαστικού βέλτιστου ελέγχου** έγκειται στον προσδιορισμό ενός επιτρεπτού κανόνα ελέγχου (7.2) που ελαχιστοποιεί το κριτήριο κόστους (7.3) λαμβάνοντας υπόψη την καταστατική εξίσωση (7.1).

Ας σημειωθεί ότι:

- Το ΠΣΒΕ διατυπώθηκε για ένα σύνολο αρχικών καταστάσεων $\mathbf{x}(0) \in \mathcal{X}(0)$.
- Οι συναρτήσεις του προβλήματος ϑ, φ, $\mathbf{f}$ μπορούν να υφίστανται σε μορφή πίνακος.
- Η λύση του προβλήματος είναι όχι μια τροχιά ελέγχου, αλλά μια συνάρτηση $\mathbf{R}[\mathbf{x}(k), k]$, ο βέλτιστος κανόνας ελέγχου, που ενδεχόμενα παρέχεται σε μορφή πίνακα. Προφανώς η λύση αυτή προϋποθέτει την διαθεσιμότητα πραγματικών μετρήσεων κατάστασης $\mathbf{x}(k)$ κατά την εφαρμογή του ελέγχου.
- Το πρόβλημα προσδιορισμού ΒΤΕ θα ετίθετο αν ζητούσαμε $\mathbf{u}(k) = \mathbf{R}(k)$ αντί της (7.2).
- Ο βέλτιστος κανόνας ελέγχου (7.2) δεν εγγυάται αποτελέσματα ισότιμα εκείνων που θα επιτυγχάνονταν αν οι ακριβείς τιμές της $\mathbf{z}(k)$, $k = 0,...,K-1$, ήταν εκ των προτέρων γνωστές. Ο κανόνας είναι βέλτιστος δεδομένης της πιθανολογικής κατανομής της $\mathbf{z}(k)$ και εφ΄ όσον οι πραγματικές (μελλοντικές) τιμές της είναι άγνωστες την στιγμή της επίλυσης του προβλήματος.

Παράδειγμα 7.2: Το στοχαστικό πρόβλημα του Παραδ. 7.1 μπορεί να εκφρασθεί σαν ΠΣΒΕ αν θεωρήσουμε στο σημείο Α την αρχή των αξόνων (x = 0, k = 0). Έχουμε τότε

$$x(k+1) = x(k) + u(k) + z(k) \ , \ k = 0, 1, 2$$

με $\mathcal{X}(0) = \{0\}$, $\mathcal{X}(1) = \{1, -1\}$, $\mathcal{X}(2) = \{2, 0, -2\}$, $\mathcal{X}(3) = \{3, 1, -1, -3\}$ και $\mathcal{U} = \{1, -1\}$, $\mathcal{Z} = \{2, 0, -2\}$. Η πιθανολογική κατανομή της διαταραχής εξαρτάται από τις τιμές ελέγχου και χαρακτηρίζεται από τις ακόλουθες διακριτές τιμές πιθανοτήτων

$$p(z = 2 \mid u = 1) = 0, \qquad p(z = 2 \mid u = -1) = \frac{1}{4}$$
$$p(z = 0 \mid u = 1) = \frac{3}{4}, \qquad p(z = 0 \mid u = -1) = \frac{3}{4}$$
$$p(z = -2 \mid u = 1) = \frac{1}{4}, \qquad p(z = -2 \mid u = -1) = 0.$$

Για το κριτήριο κόστους έχουμε $\vartheta[x(3)] = 0$ και

$$\varphi[x = 0, U = 1, k = 0] = 10, \qquad \varphi[x = 0, U = -1, k = 0] = 0$$
$$\varphi[x(1), U(1), k = 1] = 0$$
$$\varphi[x = 2, U(2), k = 2] = 0, \qquad \varphi[x = -2, U(2), k = 2] = 12$$
$$\varphi[x = 0, U = 1, k = 2] = 1200, \qquad \varphi[x = 0, U = -1, k = 2] = 0$$

με τον ορισμό U = u+z. Ο βέλτιστος κανόνας ελέγχου για το πρόβλημα αυτό έχει μορφή πίνακα που απορρέει από το Σχ. 7.2.

■

Παράδειγμα 7.3: Η δυναμική εξέλιξη του περιεχομένου μιας αποθήκης περιγράφεται με

$$x(k+1) = x(k) + u(k) - z(k) \tag{7.4}$$

όπου x(k) είναι η ποσότητα κάποιου αποθηκευμένου προϊόντος, u(k) είναι η παραγγελλόμενη και άμεσα καταφθάνουσα ποσότητα και z(k) είναι η ζήτηση (πωλήσεις) με γνωστή πιθανολογική κατανομή και ανεξάρτητες τιμές z(0),...,z(K–1). Αν η ζήτηση z(k) υπερβαίνει το άθροισμα x(k)+u(k) της αποθηκευμένης και παραγγελθείσας ποσότητας, τότε έχουμε x(k+1) αρνητικό και η επιπλέον ζήτηση καλύπτεται στα επόμενα χρονικά διαστήματα. Το προς ελαχιστοποίηση κριτήριο κόστους είναι

$$J = E\{\sum_{k=0}^{K-1} [c\, u(k) + p \max\{0, z(k) - x(k) - u(k)\} + h \max\{0, x(k) + u(k) - z(k)\}]\} \tag{7.5}$$

όπου c είναι η τιμή αγοράς ανά μονάδα του προϊόντος, p είναι το κόστος ανά μονάδα μη καλυφθείσας ζήτησης και h είναι το κόστος ανά μονάδα αποθηκευμένης ποσότητας. Το κριτήριο κόστους (7.5) εκφράζει επομένως το αναμενόμενο ολικό κόστος για την κάλυψη της στοχαστικής ζήτησης. Ο έλεγχος για ελαχιστοποίηση δίδεται βέβαια από τις παραγγελόμενες ποσότητες u(k).

■

Παράδειγμα 7.4 (Πρόβλημα προθεσμίας): Ένα άτομο κατέχει ένα αγαθό (π.χ. ένα ακίνητο), για την πώληση του οποίου δέχεται σε τακτικά χρονικά διαστήματα προσφορές. Οι τιμές των προσφορών $z(k) \in \mathcal{Z}$, $k = 0,...,K-1$, είναι τυχαίες, ανεξάρτητες μεταξύ τους και με γνωστή, μη μεταβλητή πιθανολογική κατανομή. Αν ο ιδιοκτήτης δεχθεί μια δεδομένη προσφορά, μπορεί να επενδύσει το αντίστοιχο ποσό με τόκο $r>0$ ανά χρονικό διάστημα. Αν την απορρίψει, πρέπει να περιμένει μια νέα προσφορά στο επόμενο χρονικό διάστημα. Απορριφθείσες προσφορές εκπίπτουν στα επόμενα χρονικά διαστήματα. Ο ιδιοκτήτης πρέπει όμως να πωλήσει το αγαθό το αργότερο κατά το χρονικό διάστημα K. Το πρόβλημα έγκειται στον προσδιορισμό μιας στρατηγικής αποδοχής ή απόρριψης, που μεγιστοποιεί το τελικό έσοδο κατά το χρονικό διάστημα K.

Οι προσφορές $z(k) \in \mathcal{Z} \subset R^+$ είναι προφανώς οι στοχαστικές διαταραχές του προβλήματος. Ο έλεγχος $u(k)$ παίρνει δύο τιμές, δηλαδή 1 για πώληση και 0 για απόρριψη της προσφοράς $z(k-1)$. Ως επιτρεπτή περιοχή κατάστασης θέτουμε $x(k) \in \mathcal{Z} \cup \{0\}$. Αν $x(k) = 0$, τότε το αγαθό έχει ήδη πωληθεί. Αν $x(k) \neq 0$, τότε το αγαθό δεν έχει ακόμη πωληθεί και η τιμή του $x(k)$ εκφράζει την τελευταία προσφορά. Έχουμε έτσι την ακόλουθη εξίσωση κατάστασης

$$x(k+1) = \begin{cases} 0 & \text{αν } u(k) = 1 \text{ ή } x(k) = 0 \\ z(k) & \text{άλλως.} \end{cases}$$

Το υπό μεγιστοποίηση κριτήριο κόστους (7.3) ορίζεται μέσω των ακόλουθων συναρτήσεων (εξηγήστε τις λεπτομερώς βάσει των ανωτέρω)

$$\vartheta[x(K)] = x(K)$$

$$\varphi[x(k), u(k), k] = (1+r)^{K-k} x(k)\, u(k)$$

■

7.3 Μέθοδος Στοχαστικού Δυναμικού Προγραμματισμού

Η μεθοδολογία του Δυναμικού Προγραμματισμού εφαρμόσθηκε ήδη στην στοχαστική της εκδοχή στο Παραδ. 7.1. Στο κεφάλαιο αυτό θα αναπτυχθεί η γενική μορφή του **Στοχαστικού Δυναμικού Προγραμματισμού** και η εφαρμογή της στο ΠΣΒΕ του Κεφαλαίου 7.2.

Όπως και στο Κεφάλαιο 3.2 ορίζουμε καταρχήν το υπολειπόμενο αναμενόμενο κόστος J_k ως εξής

$$J_k = E\{\vartheta[\mathbf{x}(K)] + \sum_{\kappa=k}^{K-1} \varphi[\mathbf{x}(\kappa), \mathbf{u}(\kappa), \mathbf{z}(\kappa), \kappa]\} \qquad (7.6)$$

όπου η αναμενόμενη τιμή λαμβάνεται ως προς $\mathbf{z}(\kappa)$, $\kappa=k,...,K-1$. Για ένα δεδομένο πρόβλημα, το ελάχιστο αναμενόμενο κόστος $J_k^* = \min J_k$ (λαμβάνοντας υπόψη όλους τους περιορισμούς) εξαρτάται αποκλειστικά από $\mathbf{x}(k)$ και k. Ονομάζουμε το ελάχιστο κόστος $V[\mathbf{x}(k), k]$ και έχουμε

$$V[\mathbf{x}(k), k] = \min J_k = \min\{E\{\varphi[\mathbf{x}(k), \mathbf{u}(k), \mathbf{z}(k), k] + J_{k+1}\}\} \qquad (7.7)$$

όπου το ελάχιστο νοείται ως προς όλους τους **u**(κ), κ=k,..,K−1, που εκπληρούν τους περιορισμούς. Η αναμενόμενη τιμή στην (7.7) λαμβάνεται μόνον ως προς **z**(k) καθότι οι **z**(k+1),..., **z**(K−1) είναι ανεξάρτητες των **x**(k), **u**(k) και **z**(k). Δι΄ εφαρμογής της Αρχής του Βέλτιστου στην (7.7) έχουμε

$$V[\mathbf{x}(k),k] = \min\{ \underset{\mathbf{z}(k)}{E} \{\varphi[\mathbf{x}(k), \mathbf{u}(k), \mathbf{z}(k), k] + V[\mathbf{x}(k+1), k+1]\}\} \qquad (7.8)$$

και δι΄ αντικαταστάσεως της (7.1)

$$V[\mathbf{x}(k),k] = \min\{ \underset{\mathbf{z}(k)}{E} \{\varphi[\mathbf{x}(k), \mathbf{u}(k), \mathbf{z}(k), k] + V[f[\mathbf{x}(k), \mathbf{u}(k), \mathbf{z}(k), k], k+1]\}\} \quad (7.9)$$

όπου βέβαια το ελάχιστο νοείται πλέον μόνον ως προς τους $\mathbf{u}(k)\in\mathcal{U}[\mathbf{x}(k),k]$ του χρονικού διαστήματος k. Η (7.9) είναι η **στοχαστική μορφή του αναδρομικού τύπου του Bellman** η οποία κατανέμει το στοχαστικό πρόβλημα αποφάσεων πολλαπλών βαθμίδων σε K μονοβαθμιαία στοχαστικά προβλήματα αποφάσεων και η οποία οδηγεί στον προσδιορισμό ενός στοχαστικού βέλτιστου κανόνα ελέγχου. Η συνοριακή συνθήκη για την (7.9) είναι προφανώς

$$V[\mathbf{x}(K), K] = \vartheta[\mathbf{x}(K)]. \qquad (7.10)$$

Η επίλυση του αναδρομικού τύπου (7.9) πραγματοποιείται αρχίζοντας από τον τελικό χρόνο K με την (7.10) και προχωρώντας όπισθεν, όπως και στην αιτιοκρατική περίπτωση. Στην στοχαστική μορφή απαιτείται όμως για τον υπολογισμό του κριτηρίου κόστους κατά τη διαδικασία της ελαχιστοποίησης ο υπολογισμός αντίστοιχων αναμενόμενων τιμών ως προς **z**(k), όπως είδαμε και στο Παραδ. 7.1. Ο Στοχαστικός Δυναμικός Προγραμματισμός οδηγεί σε απόλυτο ελάχιστο αν οι μονοβαθμιαίες ελαχιστοποιήσεις είναι απόλυτες.

Παράδειγμα 7.5: Μια ηλεκτρική οικιακή συσκευή παρουσιάζει βλάβη. Έχουμε K προσφορές επισκευής της από αντίστοιχους τεχνίτες με κόστη h_k και πιθανότητες επιτυχούς επισκευής p_k, k=0,…,K−1, όπου το κόστος h_k χρεώνεται ανεξάρτητα από την επιτυχή ή μη έκβαση της επισκευής, και όπου κάθε πιθανότητα p_k θεωρείται ανεξάρτητη από ενδεχόμενη προηγούμενη αποτυχία επισκευής από άλλον τεχνίτη. Εναλλακτική λύση είναι η αγορά καινούργιας συσκευής με κόστος A. Τι μας συμφέρει να κάνουμε;

Θεωρούμε μια **δεδομένη διαδοχή** τεχνιτών k=0,…,K−1 (το ερώτημα του προσδιορισμού της **βέλτιστης** διαδοχής θα τεθεί αργότερα) και θέτουμε για κάθε k το ερώτημα αν θα πρέπει να εμπιστευθούμε την συσκευή στον αντίστοιχο τεχνίτη ή όχι. Αν ναι (u(k) = 1), πληρώνουμε h_k και έχουμε p_k πιθανότητα επιτυχούς επισκευής. Αν αποφασίσουμε να μην την εμπιστευθούμε (u(k) = 0) ή αν ο k τεχνίτης αποτύχει, θέτουμε το ίδιο ερώτημα για τον τεχνίτη k+1, κ.ο.κ. Αν η συσκευή δεν έχει επισκευασθεί μέχρι και την βαθμίδα K−1, τότε αγοράζουμε καινούργια με κόστος A.

Για την μοντελοποίηση του προβλήματος θεωρούμε ως μεταβλητή κατάστασης $x(k)\in\{0,1\}$, όπου 0 ή 1 σημαίνει ότι η συσκευή έχει επισκευασθεί ή όχι, αντίστοιχα, σε μια προηγούμενη βαθμίδα 0,…,k−1. Έχουμε έτσι την καταστατική εξίσωση (εξηγήστε την)

$$x(k+1) = [1-u(k)z(k)]x(k)$$

όπου $z(k)\in\{0,1\}$ είναι στοχαστική μεταβλητή που παίρνει τιμές 1 με πιθανότητα p_k (επιτυχής επισκευή) και 0 με πιθανότητα $1-p_k$ (αποτυχία επισκευής). Έχουμε προφανώς $x(0) = 1$ (γιατί;).

Το κριτήριο κόστους (προς ελαχιστοποίηση) σύμφωνα με τα παραπάνω δεδομένα έχει ως εξής

$$J = E\{Ax(K) + \sum_{k=0}^{K-1} h_k u(k)\}.$$

Για την επίλυση του προβλήματος έχουμε καταρχήν από την (7.10)

$$V[x(K), K] = Ax(K) \tag{7.11}$$

και από την (7.9)

$$\begin{aligned} V[x(k),k] &= \min_{u(k)} \{h_k u(k) + \underset{z(k)}{E} \{V[x(k+1), k+1]\}\} \\ &= \min_{u(k)} \{h_k u(k) + p_k V[[1-u(k)]x(k), k+1] + (1-p_k)V[x(k), k+1]\} \\ &= \min\{V[x(k), k+1],\ h_k + p_k V(0, k+1) + (1-p_k)V[x(k),k+1]\}. \end{aligned}$$

Εξ αυτής συνεπάγονται

$$u(k) = \begin{cases} 1 & \alpha\nu\ h_k < p_k[V[x(k), k+1] - V(0, k+1)] \\ 0 & \text{άλλως} \end{cases}$$

$$V[x(k), k] = \begin{cases} h_k + p_k V(0, k+1) + (1-p_k) V[x(k), k+1] & \alpha\nu\ h_k < p_k[V[x(k), k+1] - V(0, k+1)] \\ V[x(k), k+1] & \text{άλλως}. \end{cases} \tag{7.12}$$

Για $k = K-1$ η (7.12) δίδει (κάνοντας χρήση της (7.11)) μετά από μερικές πράξεις

$$V[x(K-1), K-1] = \begin{cases} B_{K-1} & \alpha\nu\ \ x(K-1) = 1 \\ 0 & \alpha\nu\ \ x(K-1) = 0 \end{cases} \tag{7.13}$$

όπου η σταθερά $B_{K-1}=\min\{h_{K-1}+(1-p_{K-1})A, A\}$ εξαρτάται από τα δεδομένα του προβλήματος.

Από τις (7.13), (7.12) εύκολα εξάγεται ότι, αν $x(k)=0$, τότε $u(k)=0$ και $V(0,k)=0$. Στρέφουμε τώρα την προσοχή μας στην περίπτωση $x(k)=1$. Εφαρμόζοντας την (7.13) στην (7.12) για $k = K-2, K-3,\ldots$ έχουμε

$$V(1, k) = B_k = \min\{h_k + (1-p_k)B_{k+1}, B_{k+1}\} \tag{7.14}$$

άρα έχουμε για τον βέλτιστο έλεγχο

$$u(k) = \begin{cases} 1 & \alpha\nu\ \ h_k / p_k < B_{k+1}\ \ \ \kappa\alpha\iota\ \ x(k) = 1 \\ 0 & \text{άλλως}. \end{cases} \tag{7.15}$$

Συμπερασματικά, για ένα δεδομένο πρόβλημα, θα πρέπει να υπολογισθούν καταρχήν μέσω της

(7.14) με B_K=A τα B_k, k=K−1,…,1, και να προσδιορισθεί στη συνέχεια ο βέλτιστος κανόνας ελέγχου (7.15).

Η λύση του προβλήματος μπορεί να ερμηνευθεί και γραφικά κατά το Σχ. 7.3. Τα τελικά κόστη είναι A για x(K) = 1 και 0 για x(K) = 0 (γιατί;). Για x(k) = 0, k=1,…,K−1, η συσκευή έχει ήδη επιδιορθωθεί και άρα η μοναδική μετάβαση είναι προς x(k+1)=0 με μηδενικό κόστος. Για x(k)=1, η συσκευή δεν έχει ακόμη επιδιορθωθεί και άρα υπάρχουν δύο δυνατές επιλογές που αντανακλώνται στον κανόνα ελέγχου (7.15):

- Είτε u(k)=0, οπότε η συσκευή δεν παραδίδεται στον k τεχνίτη για επιδιόρθωση, άρα έχουμε μετάβαση μηδενικού κόστους προς x(k+1)=1˙ για την επιλογή αυτή έχουμε προφανώς κόστος κόμβου k ίσο προς B_{k+1}
- Είτε u(k)=1, δηλαδή η συσκευή παραδίδεται στον τεχνίτη k, οπότε έχουμε μετάβαση κόστους h_k είτε προς x(k+1)=0 με πιθανότητα p_k, είτε προς x(k+1)=1 με πιθανότητα $(1-p_k)$˙ για την επιλογή αυτή έχουμε επομένως κόστος κόμβου k ίσο προς $h_k+(1-p_k)B_{k+1}$.

Η βέλτιστη επιλογή προσδιορίζεται λοιπόν από την σύγκριση τιμών των δύο ως άνω εναλλακτικών επιλογών, δηλαδή $h_k+(1-p_k)B_{k+1} < B_{k+1} \Leftrightarrow h_k/p_k < B_{k+1}$ όπως και στον (7.15).

Σημειώνουμε εδώ δύο ιδιότητες της βέλτιστης λύσης που θα χρησιμεύσουν παρακάτω:

(a) $B_0 \le B_1 \le \ldots \le B_K = A$.

(b) Αν $h_k/p_k < B_{k+1}$ τότε $h_k/p_k \le B_k$.

Η ιδιότητα (a) απορρέει άμεσα από την (7.14). Για την απόδειξη της (b) έχουμε καταρχήν u(k)=1, δηλαδή $B_k = h_k+(1-p_k)B_{k+1}$, και άρα προς απόδειξη

$$h_k/p_k \le B_k = h_k +(1-p_k)\, B_{k+1} \Leftrightarrow -\, h_k(1-p_k) + (1-p_k)\, B_{k+1}P_k \ge 0$$

$$\Leftrightarrow\ (1-p_k)(B_{k+1}p_k-h_k) \ge 0$$

όπου η τελευταία ανισότητα ισχύει προφανώς αφού $p_k \in [0,1]$.

Το βέλτιστο αναμενόμενο κόστος του προβλήματος είναι προφανώς B_0 (Σχ. 7.3). Το κόστος αυτό εξαρτάται από την συγκεκριμένη διαδοχή k=0,…,K−1 των τεχνιτών, άρα τίθεται το ερώτημα της βέλτιστης διαδοχής. Θα αποδείξουμε ότι για την βέλτιστη διαδοχή ισχύει

$$h_0/p_0 \le h_1/p_1 \le \ldots \le h_{K-1}/p_{K-1} \qquad (7.16)$$

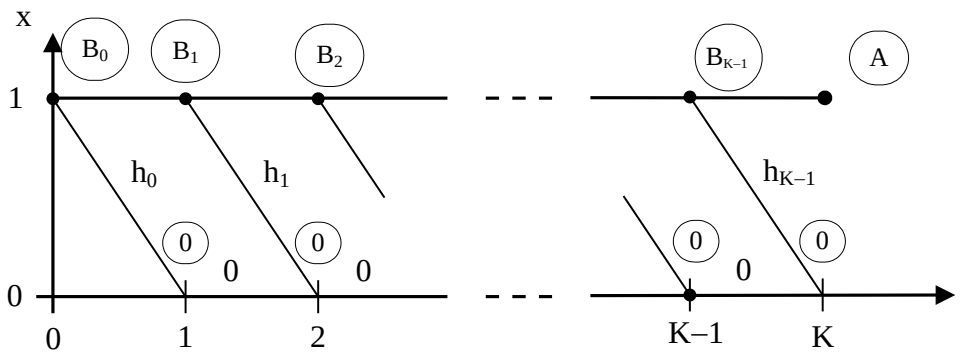
ότι δηλαδή θα πρέπει στη διαδοχή των τεχνιτών να προπορεύονται οι "φτηνοί (δηλαδή μικρό h) και καλοί (δηλαδή μεγάλη p)".

Για την απόδειξη της (7.16) θα θεωρήσουμε καταρχήν δύο τεχνίτες α και β, όπου $h_\alpha/p_\alpha < h_\beta/p_\beta$, και θα θέσουμε το εξής ερώτημα: Έστω ότι οι τεχνίτες α και β πρέπει να καταλάβουν τις θέσεις k και k+1. Ποιός θα έπρεπε να προηγηθεί, έτσι ώστε για δεδομένο B_{k+2} να έχουμε το μικρότερο B_k; Διακρίνουμε 3 περιπτώσεις:

(i) $B_{k+2} \le h_\alpha/p_\alpha < h_\beta/p_\beta$

(ii) $h_\alpha/p_\alpha \le B_{k+2} \le h_\beta/p_\beta$

(iii) $h_\alpha/p_\alpha < h_\beta/p_\beta \le B_{k+2}$.



Σχήμα 7.3: Παραστατική λύση του Παραδ. 7.5.

Στην περίπτωση (i) είναι προφανές από την (7.15) ότι $u(k)=u(k+1)=0$, και άρα $B_k=B_{k+1}=B_{k+2}$, ανεξαρτήτως σειράς διαδοχής των α, β.

Στην περίπτωση (ii) με σειρά διαδοχής α, β, έχουμε $u(k+1) = 0$, και άρα $B_{k+1}^{\alpha\beta} = B_{k+2}$, και $u(k)=1$, και άρα $B_k^{\alpha\beta} = h_\alpha+(1-p_\alpha)B_{k+2}$. Με σειρά διαδοχής β, α έχουμε $u(k+1)=1$ και $B_{k+1}^{\beta\alpha} = h_\alpha+(1-p_\alpha)B_{k+2}$. Λόγω της ιδιότητας (a) ισχύει $B_{k+1}^{\beta\alpha} \le B_{k+2} \le h_\beta/p_\beta$ και άρα $u(k) = 0$ και $B_k^{\beta\alpha} = B_{k+1}^{\beta\alpha}$. Έτσι $B_k^{\alpha\beta} = B_k^{\beta\alpha}$ και άρα η τιμή του B_k δεν επηρεάζεται ούτε στην περίπτωση (ii) από την σειρά διαδοχής των α, β.

Στην περίπτωση (iii) με σειρά διαδοχής α, β, έχουμε $B_{k+1}^{\alpha\beta} = h_\beta + (1-p_\beta) B_{k+2}$ και, επειδή $h_\alpha/p_\alpha < h_\beta/p_\beta \le B_{k+1}^{\alpha\beta}$ (λόγω της ιδιότητας (b)), έχουμε

$$B_k^{\alpha\beta} = h_\alpha + (1-p_\alpha)[h_\beta + (1-p_\beta) B_{k+2}].$$

Με σειρά διαδοχής β, α έχουμε

$$B_k^{\beta\alpha} = \min\{B_{k+1}^{\beta\alpha}, h_\beta + (1-p_\beta) B_{k+1}^{\beta\alpha}\}$$

όπου $B_{k+1}^{\beta\alpha} = h_\alpha+(1-p_\alpha) B_{k+2}$. Από την σύγκριση των δύο όρων εξάγεται ότι $B_k^{\alpha\beta} < B_k^{\beta\alpha}$ και άρα ο α πρέπει να προηγείται.

Συμπερασματικά, επειδή για δεδομένη διαδοχή 0,…,k−1 η μείωση του B_k σίγουρα δεν αυξάνει το B_0, η διαδοχή α, β είναι προτιμητέα της β, α, και αν εφαρμόσουμε την συλλογιστική αυτή σε κάθε ζεύγος της συνολικής σειράς των τεχνιτών, καταλήγουμε τελικά στην υπό απόδειξη (7.16).

■

Παράδειγμα 7.6: Θα εφαρμόσουμε την μέθοδο του Στοχαστικού Δυναμικού Προγραμματισμού για την επίλυση του προβλήματος προθεσμίας του Παραδ. 7.4. Με τα δεδομένα του προβλήματος εξάγουμε από τις (7.9), (7.10)

$$V[x(K), K] = x(K) \qquad (7.17)$$

$$V[x(k),k] = \max_{u(k)}\{(1+r)^{K-k}x(k)u(k) + \underset{z(k)}{E}\{V[x(k+1),k+1]\}\}$$

$$= \begin{cases} \max\{(1+r)^{K-k}x(k), E\{V[z(k),k+1]\}\} & \text{αν } x(k) \neq 0 \\ V(0,k+1) & \text{αν } x(k) = 0. \end{cases} \quad (7.18)$$

Στην περίπτωση x(k)=0, είναι προφανές ότι έχουμε V(0,k) = 0. Για x(k) ≠ 0 έχουμε από την (7.18) την εξής βέλτιστη πολιτική:

Αποδοχή της προσφοράς z(k−1) = x(k), αν x(k) > α(k)

Απόρριψη της προσφοράς z(k−1) = x(k), αν x(k) < α(k)

όπου

$$\alpha(k) = \frac{E\{V[z(k), k+1]\}}{(1+r)^{K-k}}. \quad (7.19)$$

Για x(k)=α(k) είναι και οι δύο αποφάσεις βέλτιστες. Το πρόβλημα ανάγεται έτσι στον υπολογισμό της χρονικής σειράς τιμών α(k), η οποία αποτελεί ένα χρονικά μεταβλητό κατώφλι αποφάσεων (βλ. Σχ. 7.4).

Για απλούστευση των υπολογισμών ορίζουμε το άτοκο βέλτιστο αναμενόμενο έσοδο

$$L[x(k),k] = \frac{V[x(k),k]}{(1+r)^{K-k}} \quad \forall x(k) \neq 0 \quad (7.20)$$

και έχουμε από την αναδρομική εξίσωση (7.17), (7.18)

$$L[x(k),k] = \begin{cases} x(k) & \text{αν} \quad x(k) > \alpha(k) \\ \alpha(k) & \text{αν} \quad x(k) < \alpha(k). \end{cases} \quad (7.21)$$

Από τους ορισμούς (7.19), (7.20) συνεπάγεται

$$\alpha(k) = \frac{1}{1+r} E\{L[z(k), k+1]\}. \quad (7.22)$$

Λαμβάνοντας υπόψη τον γενικό τύπο

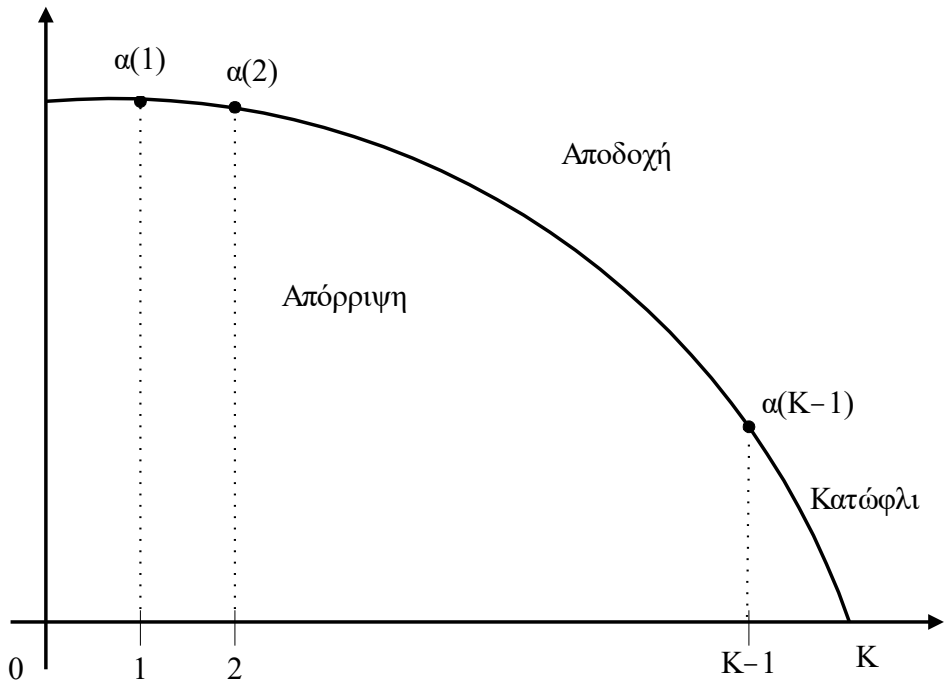
$$E\{f(z)\} = \int_{-\infty}^{\infty} f(z)\cdot dP(z) \quad (7.23)$$

όπου P(z) είναι η συνάρτηση κατανομής της τυχαίας μεταβλητής z, έχουμε από τις (7.21), (7.22)

$$\alpha(k) = \frac{1}{1+r}\{\int_{0}^{\alpha(k+1)} \alpha(k+1)\, dP[z(k)] + \int_{\alpha(k+1)}^{\infty} z(k)\, dP[z(k)]\}$$

και μετά από επεξεργασία του πρώτου όρου

$$\alpha(k) = \frac{P[\alpha(k+1)]}{1+r}\alpha(k+1) + \frac{1}{1+r}\int_{\alpha(k+1)}^{\infty} z(k) dP[z(k)]. \quad (7.24)$$



Σχήμα 7.4: Βέλτιστο, χρονικά μεταβλητό κατώφλι αποφάσεων.

Η εξίσωση αυτή είναι, για δεδομένη συνάρτηση P(z), μια αιτιοκρατική δυναμική εξίσωση διακριτού χρόνου που μπορεί να επιλυθεί αναδρομικά εκκινώντας από την τελική τιμή α(K) = 0 για τον υπολογισμό της σειράς α(K−1),...,α(0). Η τελική τιμή α(K)=0 είναι εύλογη (αφού η τελευταία προσφορά γίνεται οπωσδήποτε δεκτή), μπορεί όμως και να αποδειχθεί γράφοντας την (7.19) για $k = K-1$

$$\alpha(K-1)= E\{V[x(K),K]\}/(1+r) = E\{x(K)\}/(1+r) = E\{z(K-1)\}/(1+r)$$

και διαπιστώνοντας ότι η τελευταία εξίσωση ικανοποιεί όντως την (7.24) για k=K−1 αν α(K)=0.

Η συνάρτηση πιθανολογικής κατανομής P(z) είναι γενικά μια μονοτονικά αύξουσα συνάρτηση. Κάνοντας χρήση της ιδιότητας αυτής, μπορεί να αποδειχθεί ότι η σειρά των α(k) που απορρέει από την (7.24) είναι μονοτονικά φθίνουσα, δηλαδή α(K) ≤ α(K−1) ≤ ... ≤ α(0) (βλ. Σχ. 7.4), όπως άλλωστε αναμένονταν για το δεδομένο πρόβλημα (γιατί;), και ότι συγκλίνει προς μια στάσιμη τιμή $\bar{\alpha}$ που ικανοποιεί την ακόλουθη στάσιμη εκδοχή της (7.24)

$$\bar{\alpha}(1+r) = P(\bar{\alpha})\bar{\alpha} + \int_{\bar{\alpha}}^{\infty} z\,dP(z) \qquad (7.25)$$

Άρα, όταν η προθεσμία πώλησης είναι αρκούντως μακρινή (K→ ∞), έχουμε την ακόλουθη βέλτιστη πολιτική αποφάσεων:

Αποδοχή αν $x(k) > \bar{\alpha}$
Απόρριψη αν $x(k) < \bar{\alpha}$

όπου $\bar{\alpha}$ είναι η λύση της εξίσωσης (7.25) (βλ. Προβλ. 7.1).

Το πρόβλημα προθεσμίας επιλύεται κατά παρόμοιο τρόπο στην ειδική περίπτωση διακριτής πιθανολογικής κατανομής των προσφορών, δηλαδή αν $z(k)\in\mathcal{Z}=\{Z_1, Z_2, \ldots, Z_n\}$ (όπου $Z_1 \le Z_2 \le \ldots \le Z_n$) και $z(k)=Z_i$ με γνωστή πιθανότητα p_i, όπου βέβαια

$$\sum_{i=1}^{n} p_i = 1.$$

Στην περίπτωση αυτή, ο γενικός τύπος (7.23) παίρνει την μορφή

$$E\{f(z)\} = \sum_{i=1}^{n} p_i f(Z_i) \qquad (7.23a)$$

και άρα έχουμε με τις (7.21), (7.22)

$$\alpha(k) = \frac{1}{1+r}\{\sum_{i=1}^{j(k)} p_i\, \alpha(k+1) + \sum_{i=j(k)+1}^{n} p_i Z_i\}$$

όπου j(k) είναι ο δείκτης-όριο, έτσι ώστε $Z_{j(k)} \leq \alpha(k+1)$ και $Z_{j(k)+1} \geq \alpha(k+1)$. Καταλήγουμε έτσι, αντί της (7.24), στην αιτιοκρατική δυναμική εξίσωση διακριτού χρόνου

$$\alpha(k) = [\sum_{i=1}^{j(k)} p_i/(1+r)]\alpha(k+1) + \sum_{i=j(k)+1}^{n} p_i Z_i/(1+r) \qquad (7.24a)$$

που μπορεί να επιλυθεί για γνωστά p_i, Z_i, i=1,…,n, εκκινώντας από την ίδια συνοριακή συνθήκη όπως παραπάνω (βλ. Πρόβλημα 7.6).

7.4 Διακριτός Στοχαστικός Δυναμικός Προγραμματισμός

Για την εφαρμογή του Στοχαστικού Δυναμικού Προγραμματισμού στην αριθμητική επίλυση ΠΣΒΕ με συνεχείς περιοχές τιμών $\mathscr{X}$, $\mathscr{U}$, $\mathscr{Z}$ πρέπει, όπως και στην αιτιοκρατική εκδοχή του Κεφαλαίου 3.3, να προβούμε σε διακριτοποίηση των περιοχών τιμών με την δημιουργία διακριτού πλέγματος. Επιπλέον, πρέπει κάθε διακριτή τιμή $\mathbf{z}^i(k)$ διαταραχών να συνοδεύεται από αντίστοιχες τιμές δεσμευμένων πιθανοτήτων $p[\mathbf{z}^i(k) \mid \mathbf{x}^j(k), \mathbf{u}^l(k), k]$, έχουμε δηλαδή διακριτοποίηση της συνάρτησης πιθανολογικής πυκνότητας.

Μετά την διακριτοποίηση των περιοχών τιμών και της συνάρτησης πιθανολογικής πυκνότητας μπορούμε να επιλύσουμε τον αναδρομικό τύπο (7.9) (με την συνοριακή συνθήκη (7.10)) αριθμητικά, όπως και στο Κεφάλαιο 3.3. Εδώ απαιτείται όμως και ο αριθμητικός υπολογισμός της αναμενόμενης τιμής που εμπεριέχεται στην (7.9), όπως είδαμε και στο Παραδ. 7.1. Πρέπει να τονισθεί ότι, σε αντίθεση με την αιτιοκρατική περίπτωση και την ειδική περίπτωση του Παραδ. 7.1, το γενικό ΠΣΒΕ δυσκολεύει την γραφική επίλυση μονοδιάστατων προβλημάτων στο διάγραμμα (x,k). Αυτό οφείλεται στο γεγονός ότι μία και η αυτή μετάβαση από κάποιο διακριτό $x^i(k)$ σε κάποιο διακριτό $x^j(k+1)$ είναι δυνατή με περισσότερους του ενός συνδυασμούς τιμών [u(k), z(k)] (βλ. για παράδειγμα το Σχ. 7.3). Επί παραδείγματι, για την καταστατική εξίσωση (7.4) έχουμε μετάβαση από x(0)=0 σε x(1)=1 μέσω [u(0)=1, z(0)=0] ή [u(0)=0, z(0)=1] κλπ. Κατά συνέπεια, μια συγκεκριμένη μετάβαση μπορεί να συνοδεύεται από διαφορετικές τιμές κόστους ανάλογα με την υφή της. Βέβαια, ανάλογη περίπτωση μπορεί να εμφανισθεί και σε αιτιοκρατικά προβλήματα, αν π.χ. έχουμε βαθμωτή κατάσταση x, αλλά δύο ή περισσότερες μεταβλητές ελέγχου u_1, u_2. Σε απλές περιπτώσεις όμως, με λίγες διακριτές τιμές u(k) και z(k), μπορεί να επιχειρηθεί γραφική επίλυση, π.χ. με την εισαγωγή ενδιάμεσων κόμβων όπως στο Παράδειγμα 7.8.

Σημειώνουμε ότι, λόγω των υπολογισμών αναμενόμενων τιμών, ο αναγκαίος χρόνος υπολογισμού (3.20) πολλαπλασιάζεται στην περίπτωση στοχαστικών προβλημάτων με γ^p, όπου γ είναι ο αριθμός διακριτών σημείων ανά διάσταση του διανύσματος διαταραχών και p= dim(**z**).

Τα πλεονεκτήματα και μειονεκτήματα του Διακριτού Στοχαστικού Δυναμικού Προγραμματισμού είναι παρόμοια με αυτά που αναφέρθηκαν στην αιτιοκρατική περίπτωση (Κεφάλαιο 3.3ε). Στην στοχαστική περίπτωση θα πρέπει όμως επί πλέον να τονισθεί ότι δεν υπάρχουν πρακτικά εφαρμόσιμες εναλλακτικές μέθοδοι λύσης ΠΣΒΕ πέραν του Στοχαστικού Δυναμικού Προγραμματισμού.

Παράδειγμα 7.7: Θεωρούμε μια παραλλαγμένη και διακριτοποιημένη εκδοχή του προβλήματος αποθήκευσης στο Παραδ. 7.3, για την οποία οι αποθηκευμένες ποσότητες x(k) και οι παραγγελόμενες ποσότητες u(k) είναι μη αρνητικοί ακέραιοι αριθμοί. Επίσης θεωρούμε μια ανώτατη επιτρεπτή τιμή X για το άθροισμα x(k)+u(k). Τέλος, υποθέτουμε ότι η υπερζήτηση z(k)−x(k)−u(k)>0 τιμωρείται μεν κατά το (7.5), αλλά εκπίπτει (δεν καλύπτεται στα επόμενα χρονικά διαστήματα). Έχουμε επομένως αντί της (7.4)

$$x(k+1) = \max[0, x(k) + u(k) - z(k)].$$

Θα επιλύσουμε το πρόβλημα αυτό με τις ακόλουθες αριθμητικές τιμές: X=2, K=3, h=c=1, p=3, και μη μεταβλητή πιθανολογική κατανομή της ζήτησης z(k) ως εξής

$$p(z = 0) = \frac{1}{10},\ p(z = 1) = \frac{7}{10},\ p(z = 2) = \frac{2}{10}.$$

Για την αρχική κατάσταση θεωρούμε αποκλειστικά x(0) = 0. Η επίλυση του προβλήματος μέσω Διακριτού Στοχαστικού Δυναμικού Προγραμματισμού έχει ως εξής.

Βαθμίδα 2: Για κάθε $x(2)\in\{0,1,2\}$ πρέπει να προσδιορισθεί ο αντίστοιχος u(2)=R[x(2),2] που ελαχιστοποιεί

$$J_2 = \underset{z(2)}{E}\{u(2) + \max[0, x(2)+u(2)-z(2)] + 3\max[0, z(2)-x(2)-u(2)]\}$$

λαμβάνοντας υπόψη $0\le u(2)\le 2-x(k)$. Στην παραπάνω εξίσωση χρησιμοποιήθηκε η V[x, 3]=0 που απορρέει από την έλλειψη τελικού κόστους. Επιχειρούμε παραδειγματικά την ελαχιστοποίηση για x(2)=1 ως εξής

$$\begin{aligned} J_2 &= \underset{z(2)}{E}\{u(2) + \max[0, 1+u(2)-z(2)] + 3\max[0, z(2)-1-u(2)]\} \\ &= u(2) + \frac{1}{10}\{\max[0, 1+u(2)] + 3\max[0, -1-u(2)]\} \\ &\quad + \frac{7}{10}\{\max[0, u(2)] + 3\max[0, -u(2)]\} \\ &\quad + \frac{2}{10}\{\max[0, u(2)-1] + 3\max[0, 1-u(2)]\}. \end{aligned}$$

Οι περιορισμοί δίδουν στην περίπτωση αυτή $u(2)\in\{0,1\}$ και από την παραπάνω εξίσωση έχουμε για u(2)=0 J_2=0.7 και για u(2)=1 J_2 =1.9, οπότε R[x(2)=1, 2]= 0 και V[x(2)=1, 2] = 0.7 και κατά παρόμοιο τρόπο

R[x(2)=0, 2] =1, V[x(2)=0, 2] =1.7

R[x(2)=2, 2]=0, V[x(2)=2, 2]=0.9.

Βαθμίδα 1: Για κάθε $x(1)\in\{0, 1, 2\}$ πρέπει να προσδιορισθεί ο αντίστοιχος u(1) = R[x(1), 1] που ελαχιστοποιεί

$$\tilde{J}_1 = \underset{z(1)}{E}\{u(1) + \max[0, x(1)+u(1)-z(1) + 3\max[0, z(1)-x(1)-u(1)] + V[\max[0, x(1)+u(1)-z(1)], 2]\}$$

λαμβάνοντας υπόψη $0 \le u(1) \le 2-x(1)$. Για $x(1)=1$ έχουμε

$$\begin{aligned}\tilde{J}_1 &= \underset{z(1)}{E}\{u(1) + \max[0, 1+u(1)-z(1) + 3\max[0, z(1)-x(1)-u(1)] \\ &\quad + V[\max[0, x(1)+u(1)-z(1)], 2]\} \\ &= u(1) + \frac{1}{10}\{\max[0, 1+u(1)] + 3\max[0, -1-u(1)] + V[\max[0, 1+u(1)], 2]\} \\ &\quad + \frac{7}{10}\{\max[0, u(1)]+3\max[0, -u(1)] + V[\max[0, u(1)], 2]\} \\ &\quad + \frac{2}{10}\{\max[0, u(1)-1] + 3\max[0, 1- u(1)] + V[\max[0, u(1)-1], 2]\}.\end{aligned}$$

Οι περιορισμοί δίδουν $u(1)\in\{0,1\}$ και έτσι έχουμε για $u(1)=0$ και $u(1)=1$ τα αναμενόμενα κόστη $J_1=2.3$ και $J_1=2.82$, αντίστοιχα, οπότε $R[x(1) = 1, 1]=0$ και $V[x(1) = 1, 1]=2.3$ και κατά παρόμοιο τρόπο

$$R[x(1)=0, 1] = 1, \qquad V[x(1)=0, 1] = 3.3$$
$$R[x(1)=2, 1] = 0, \qquad V[x(1)=2, 1] = 1.82.$$

Βαθμίδα 0: Για $x(0)=0$ πρέπει να προσδιορισθεί ο $u(0)=R[x(0),0]$ που ελαχιστοποιεί

$$\tilde{J}_0 = \underset{z(0)}{E}\{u(0) + \max[0, u(0)-z(0)] + 3\max[0, z(0)-u(0)] + V[\max[0, u(0)-z(0)], 1]\}$$

λαμβάνοντας υπόψη $0 \le u(0) \le 2$. Κατά παρόμοιο τρόπο όπως παραπάνω έχουμε

$$R[x(0)=0, 0] = 1, \qquad V[x(0)=0, 0] = 4.9.$$

Συμπερασματικά ο βέλτιστος κανόνας ελέγχου των τριών βαθμίδων μπορεί να συμπτυχθεί ως εξής

$$R[x(k),k] = \begin{cases} 1 & \text{αν } x(k) = 0 \\ 0 & \text{άλλως} \end{cases}$$

για $k=0, 1, 2$.

■

Θα παρουσιάσουμε τώρα ένα παράδειγμα γραφικής επίλυσης ενός ΠΣΒΕ με λίγες διακριτές τιμές για $u(k)$ και $z(k)$. Πρόκειται για ένα στοχαστικό παίγνιο, για την επίλυση του οποίου γενικεύεται η διαδικασία min-max του Κεφαλαίου 2.2 ώστε να ληφθούν υπόψη οι εμπλεκόμενες στοχαστικές μεταβλητές.

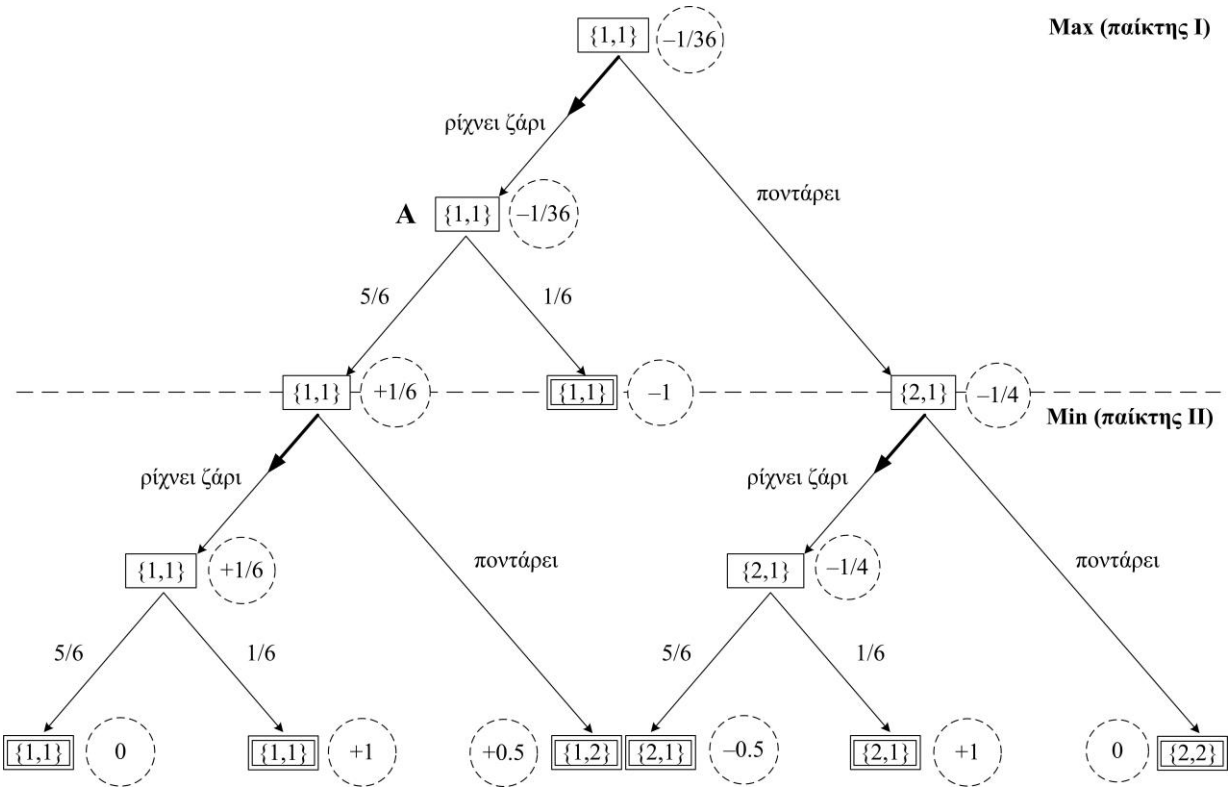
Παράδειγμα 7.8: Δύο παίκτες ρίχνουν ζάρια με τον εξής τρόπο. Αρχικά κάθε παίκτης ποντάρει από ένα ευρώ. Κατόπιν ο παίκτης Ι ρίχνει το ζάρι. Αν το αποτέλεσμα της ρίψης είναι 1, ο παίκτης Ι χάνει (δηλαδή παίρνει τα χρήματα ο παίκτης ΙΙ) και το παιχνίδι τελειώνει. Αν το αποτέλεσμα είναι 2, 3, 4, 5 ή 6, τότε έρχεται η σειρά του παίκτη ΙΙ (λέμε σε αυτή την

περίπτωση ότι ο παίκτης Ι 'περνάει'). Προσθέτουμε τον κανόνα ότι ο παίκτης Ι μπορεί να αποφύγει να ρίξει το ζάρι αν προσθέσει ένα ακόμη ευρώ στο τραπέζι, οπότε αυτόματα 'περνάει'. Στη συνέχεια ο παίκτης ΙΙ μπορεί να κάνει ό,τι ακριβώς και ο παίκτης Ι, δηλαδή μπορεί είτε να ρίξει ζάρι είτε να ποντάρει και αυτός επί πλέον ένα ευρώ και να 'περάσει'. Τα αποτελέσματα της ρίψης του ζαριού από τον παίκτη ΙΙ είναι όμοια (αν ο ΙΙ φέρει 1, τότε παίρνει ο Ι όλα τα χρήματα που είναι στο τραπέζι, διαφορετικά ο ΙΙ 'περνάει'). Το παιχνίδι τελειώνει αν κάποιος παίκτης χάσει ή αν 'περάσουν' και οι δύο παίκτες. Στην τελευταία περίπτωση τα χρήματα που έχουν ποντάρει μέχρι τότε οι παίκτες, μοιράζονται ανάμεσά τους.

Το Σχ. 7.5 αναπαριστά την πλήρη ανάπτυξη και επίλυση του παιγνίου, ειδωμένου από την σκοπιά του παίκτη που παίζει πρώτος. Οι κόμβοι-ορθογώνια δείχνουν την τρέχουσα κατάσταση του παιγνίου η οποία χαρακτηρίζεται από το συνολικό ποσό που έχει ποντάρει κάθε παίκτης στο τραπέζι, δηλαδή {ποσό παίκτη Ι, ποσό παίκτη ΙΙ}. Το κέρδος των τελικών κόμβων (που σημειώνονται στο Σχ. 7.5 με διπλό πλαίσιο) χαρακτηρίζεται από το ποσό που κερδίζει ή χάνει ο παίκτης Ι. Για παράδειγμα, αν οι παίκτες βρίσκονται στην κατάσταση {1,2} (3 ευρώ στο τραπέζι) και έχουν 'περάσει' και οι δύο παίκτες, το παιχνίδι τελειώνει και τα χρήματα μοιράζονται, επομένως ο παίκτης Ι κερδίζει 0.5 ευρώ.

Προκειμένου να ληφθεί υπόψη η στοχαστικότητα του προβλήματος, εισάγονται (όπου απαιτείται) **ενδιάμεσοι κόμβοι**. Για παράδειγμα, κατά την εκκίνηση του παιγνίου, έχει ο παίκτης Ι τις εξής επιλογές:

- Είτε να ποντάρει, οπότε μεταβαίνουμε με βεβαιότητα στην επόμενη βαθμίδα.
- Είτε να ρίξει ζάρι, οπότε η μετάβαση στην επόμενη βαθμίδα εξαρτάται από το αποτέλεσμα της ζαριάς. Η στοχαστικότητα αυτή αναπαριστάται με έναν ενδιάμεσο κόμβο (Α στο Σχ. 7.5), από τον οποίο εξέρχονται δύο δυνατές μεταβάσεις με (σημειούμενες) πιθανότητες 5/6 και 1/6, αντίστοιχα, σύμφωνα με τους κανόνες του παιγνίου.



Σχήμα 7.5: Επίλυση στοχαστικού παιγνίου.

Μετά την πλήρη ανάπτυξη του στοχαστικού παιγνίου, μπορεί να γίνει η επίλυση, αρχίζοντας από την τελευταία βαθμίδα, όπως και στο Κεφάλαιο 2.2. Η πρώτη διαφορά από τα αιτιοκρατικά παίγνια (ή προβλήματα, γενικότερα) έγκειται βέβαια στο γεγονός ότι τα προς προσδιορισμό βέλτιστα κέρδη κάθε κόμβου (εκτός των τελικών κόμβων) αντιστοιχούν σε αναμενόμενες τιμές λόγω στοχαστικότητας. Η δεύτερη διαφορά έγκειται στην ανάγκη υπολογισμού αναμενόμενων τιμών σε κάθε ενδιάμεσο κόμβο, βάσει των αντίστοιχων πιθανοτήτων. Για παράδειγμα, στην κατάσταση {2,1} όπου παίζει ο παίκτης ΙΙ, έχει να επιλέξει μεταξύ της σίγουρης ισοπαλίας (μηδέν κέρδος για τον παίκτη Ι) αν ποντάρει, και κάποιου αναμενόμενου κέρδους αν ρίξει το ζάρι. Το αναμενόμενο κέρδος στον αντίστοιχο ενδιάμεσο κόμβο υπολογίζεται ως εξής:

$$1\cdot\frac{1}{6}+(-0.5)\cdot\frac{5}{6}=-\frac{1}{4}.$$

Έτσι ο παίκτης ΙΙ έχει συνολικά να επιλέξει μεταξύ αναμενόμενου κέρδους $-1/4$ (αν ρίξει ζάρι) και 0 (αν ποντάρει), και επειδή ελαχιστοποιεί, η βέλτιστη απόφαση σε αυτήν την κατάσταση είναι να επιλέξει να ρίξει το ζάρι. Αντίστοιχα υπολογίζονται τα αναμενόμενα κέρδη για τις υπόλοιπες καταστάσεις του παιγνίου, μέχρι την κατάσταση εκκίνησης, δηλαδή την τελική επίλυση.

■

7.5 Επεκτάσεις του Προβλήματος

Στο κεφάλαιο αυτό θα εξετάσουμε μερικές επεκτάσεις του προβλήματος που ανάγονται τελικά στο βασικό ΠΣΒΕ του Κεφαλαίου 7.2.

α) Προβλήματα με Νεκρούς Χρόνους

Μπορούμε να επεκτείνουμε το είδος των εξεταζόμενων συστημάτων θεωρώντας αντί της (7.1) την ακόλουθη καταστατική εξίσωση που περιλαμβάνει νεκρούς χρόνους

$$\mathbf{x}(k+1) = \mathbf{f}[\mathbf{x}(k),\mathbf{x}(k-1), \mathbf{u}(k), \mathbf{u}(k-1), \mathbf{z}(k),k] \ , \ \ k = 0,\ldots, K-1$$

$$\mathbf{x}(1) = \mathbf{f}_0[\mathbf{x}(0),\mathbf{u}(0),\mathbf{z}(0)]. \qquad (7.26)$$

Αυτή η δυναμική εξίσωση μπορεί να αναχθεί στην μορφή (7.1) με τον ορισμό νέων μεταβλητών κατάστασης $\tilde{\mathbf{x}}(k+1) = \mathbf{x}(k)$ και $\bar{\mathbf{x}}(k+1) = \mathbf{u}(k)$ και αντικατάστασή τους στην (7.26), οπότε έχουμε βέβαια ένα επαυξημένο διάνυσμα κατάστασης $[\mathbf{x}^T\ \tilde{\mathbf{x}}^T\ \bar{\mathbf{x}}^T]^T$. Κατά τον ίδιο τρόπο μπορούν να αντιμετωπισθούν νεκροί χρόνοι στο κριτήριο κόστους ή στους περιορισμούς $\mathbf{u}(k)\in\mathcal{U}[\mathbf{x}(k), \mathbf{x}(k-1), k]$. Η επεξεργασία μεγαλύτερων νεκρών χρόνων $\mathbf{x}(k-2)$, $\mathbf{u}(k-2)$, $\mathbf{x}(k-3)$, $\mathbf{u}(k-3)$,... μπορεί προφανώς να γίνει κατά ανάλογο τρόπο μέσω περαιτέρω επεκτάσεων του διανύσματος κατάστασης. Η επέκταση αυτή ισχύει βέβαια και για την αιτιοκρατική περίπτωση του Κεφαλαίου 3.

β) Συσχετιζόμενες, Μετρήσιμες Διαταραχές

Αν οι διαταραχές $\mathbf{z}(k)$ εξαρτώνται όχι από $\mathbf{x}(k)$, $\mathbf{u}(k)$ αλλά από προηγούμενες τιμές $\mathbf{z}(k-1)$, $\mathbf{z}(k-2)$,..., υφίσταται συχνά η δυνατότητα σχηματισμού ενός **προτύπου διαταραχών**

$$\mathbf{z}(k+1) = \mathbf{F}(k)\mathbf{z}(k) + \xi(k), \quad k = 0,1,..., K-1 \qquad (7.27)$$

όπου $\xi(k)$ είναι χρονικά ανεξάρτητες διαταραχές με γνωστή πιθανολογική κατανομή. Προσδένοντας την (7.27) στην (7.1) έχουμε ένα επαυξημένο πρόβλημα με διάνυσμα κατάστασης $[\mathbf{x}^T\ \mathbf{z}^T]^T$ και διαταραχές ξ στη μορφή του βασικού ΠΣΒΕ. Η επέκταση αυτή προϋποθέτει βέβαια ότι οι καθεαυτό διαταραχές $\mathbf{z}(k)$, που μετατράπηκαν σε μεταβλητές κατάστασης του επαυξημένου προβλήματος, είναι μετρήσιμες σε πραγματικό χρόνο καθότι πρέπει να συμπεριληφθούν στον (επαυξημένο) βέλτιστο κανόνα ελέγχου.

γ) Χρήση Προγνώσεων

Ο βέλτιστος κανόνας ελέγχου (7.2) κάποιου συγκεκριμένου προβλήματος εξαρτάται ως γνωστόν από τις γνωστές πιθανολογικές κατανομές των εμπλεκόμενων στοχαστικών μεταβλητών. Σε κάποιες εφαρμογές έχουμε όμως σε πραγματικό χρόνο νεότερες πληροφορίες για τις διαταραχές $\mathbf{z}(k)$, π.χ. νεότερες πιθανολογικές κατανομές. Στις περιπτώσεις αυτές υφίσταται βέβαια η επιθυμία χρησιμοποίησης των νεότερων πληροφοριών με στόχο την βελτίωση του ελέγχου.

Θα θεωρήσουμε την περίπτωση διαταραχών $\mathbf{z}(k)$ που είναι ανεξάρτητες των $\mathbf{x}(k)$, $\mathbf{u}(k)$. Επιπλέον υποθέτουμε ότι η πιθανολογική κατανομή της $\mathbf{z}(k)$ στο χρονικό διάστημα k προέρχεται από ένα πεπερασμένο σύνολο $\{P_1(\mathbf{z}|k),..., P_q(\mathbf{z}|k)\}$ γνωστών εναλλακτικών κατανομών. Κατά την εφαρμογή του κανόνα ελέγχου σε πραγματικό χρόνο, καταφθάνει στην αρχή του χρονικού διαστήματος k η πληροφορία για την πιθανολογική κατανομή της διαταραχής στο διάστημα αυτό. Η εκ των προτέρων πιθανότητα ότι η κατανομή στο διάστημα k θα είναι $P_i(\mathbf{z}|k)$ είναι P_i^k και είναι γνωστή. Με τα δεδομένα αυτά μπορούμε να περιγράψουμε την διαδικασία πρόγνωσης μέσω μιας νέας μεταβλητής κατάστασης

$$\tilde{x}(k+1) = \xi(k)$$

όπου $\tilde{x}(k+1)$ παίρνει τιμές 1, 2,..., q και $\xi(k)$ είναι μια τυχαία μεταβλητή που παίρνει τιμές 1, 2,..., q με αντίστοιχες πιθανότητες $P_1^k, P_2^k,..., P_q^k$. Η έννοια αυτού του ορισμού είναι ότι, όταν η $\xi(k)$ παίρνει την τιμή i, τότε η $\mathbf{z}(k+1)$ έχει την κατανομή $P_i(\mathbf{z}|k+1)$. Έχουμε λοιπόν ένα επαυξημένο σύστημα

$$\begin{bmatrix} \mathbf{x}(k+1) \\ \tilde{x}(k+1) \end{bmatrix} = \begin{bmatrix} \mathbf{f}[x(k), \mathbf{u}(k), \mathbf{z}(k), k] \\ \xi(k) \end{bmatrix}$$

και η πιθανολογική κατανομή της $\mathbf{z}(k)$ εξαρτάται από την κατάσταση, δηλαδή από $\tilde{x}(k)$, όπως περιγράφηκε παραπάνω, έχουμε δηλαδή $P(z|\tilde{x}(k),k) = P_{\tilde{x}(k)}(z|k)$. Ο αντίστοιχος επαυξημένος κανόνας ελέγχου εξαρτάται τότε από τις πραγματικές μετρήσεις της καθεαυτό κατάστασης $\mathbf{x}(k)$, αλλά και από τις νεότερες πληροφορίες $\tilde{x}(k)$ περί της ισχύουσας κατανομής των διαταραχών $\mathbf{z}(k)$.

δ) Μη Μεταβλητός Κανόνας Ελέγχου

Υπό ορισμένες παραδοχές, η λύση του ΠΣΒΕ οδηγεί σε **μη μεταβλητό βέλτιστο κανόνα ελέγχου**. Η ακριβής διατύπωση των παραδοχών αυτών απαιτεί διεξοδικότερη ανάλυση που ξεπερνά τα όρια του παρόντος συγγράμματος. Στην πλειοψηφία των πρακτικά ενδιαφερόντων προβλημάτων μπορούμε να αναμένουμε μη μεταβλητό βέλτιστο κανόνα ελέγχου αν εκπληρούνται οι ακόλουθες παραδοχές:

- Οι συναρτήσεις φ, **f** του προβλήματος είναι μη μεταβλητές.
- Η συνάρτηση φ είναι θετικά ημιορισμένη.
- Οι περιορισμοί του προβλήματος και η πιθανολογική κατανομή των διαταραχών είναι μη μεταβλητά.
- Ο χρονικός ορίζοντας K είναι άπειρος.

Τα αντίστοιχα ΠΣΒΕ έχουν καταρχήν βέβαια νόημα αν το κριτήριο κόστους, παρά τον άπειρο χρονικό ορίζοντα, παίρνει πεπερασμένες τιμές τουλάχιστον για μερικές αρχικές καταστάσεις $\mathbf{x}(0)$. Σε πολλά πρακτικά ενδιαφέροντα προβλήματα έχει επίσης νόημα η ελαχιστοποίηση της μέσης τιμής (ανά βαθμίδα) του κριτηρίου κόστους, η οποία μπορεί να παίρνει πεπερασμένες τιμές ενώ το συνολικό κριτήριο κόστους είναι άπειρο.

Ένας τρόπος υπολογισμού μη μεταβλητού κανόνα ελέγχου είναι η εφαρμογή του αλγόριθμου Στοχαστικού Δυναμικού Προγραμματισμού εκκινώντας με κάποια τυχαία συνοριακή τιμή (7.10) και προχωρώντας όπισθεν, υπολογίζοντας $\mathbf{R}[\mathbf{x}(K-1),K-1]$, $\mathbf{R}[\mathbf{x}(K-2),K-2]$,... μέχρι τη σύγκλιση προς ένα στάσιμο κανόνα ελέγχου $\mathbf{R}[\mathbf{x}(k)]$ που αντιστοιχεί στο ζητούμενο μη μεταβλητό έλεγχο. Ένας αναλυτικός υπολογισμός του μη μεταβλητού κανόνα ελέγχου εξετάσθηκε ήδη στο Παράδ. 7.6 για $K\to\infty$.

Παράδειγμα 7.9: Θεωρούμε το πρόβλημα αυτόνομης κίνησης ενός οχήματος σε δυσδιάστατο χώρο, έτσι ώστε να αποφεύγονται συγκρούσεις με ενδεχόμενα κινούμενα εμπόδια. Το όχημα πρέπει να κινείται κατά το δυνατό με μέγιστη ταχύτητα $u_1(k) \approx u_{1,max}$ στην κατεύθυνση της συντεταγμένης αρ. 1 και κατά το δυνατό χωρίς αποκλίσεις, δηλαδή $u_2(k) \approx 0$. Αν συναντήσει στο δρόμο του ένα κινούμενο εμπόδιο, το όχημα μπορεί να μειώσει ταχύτητα ή και να οπισθοδρομήσει ($u_1(k) < u_{1,max}$) ή να αποκλίνει ($u_2(k) \neq 0$) για να αποφύγει σύγκρουση με το εμπόδιο. Όχημα και εμπόδια θεωρούνται σχήματος ορθογωνίου παραλληλόγραμμου και κινούνται χωρίς να αλλάζουν τον προσανατολισμό τους, κινούνται δηλαδή έχοντας τις πλευρές τους παράλληλες προς τους άξονες συντεταγμένων (ή θεωρούνται κυκλικά, οπότε δεν έχει καμία επιρροή ο προσανατολισμός κίνησης). Υπό τις παραδοχές αυτές μπορούμε να θεωρήσουμε το όχημα σαν ένα σημείο, προσθέτοντας τις διαστάσεις του στις διαστάσεις του εμποδίου. Θεωρούμε $\mathbf{x}(k) \in \mathscr{X} \subset R^2$ την μετρήσιμη απόσταση του σημείου-οχήματος από το κέντρο του εμποδίου. Η καταστατική εξίσωση κίνησης έχει (για διακριτό χρονικό διάστημα T=1) ως εξής

$$\mathbf{x}(k+1) = \mathbf{x}(k) - \mathbf{u}(k) + \mathbf{z}(k) \qquad (7.28)$$

με περιορισμούς ταχύτητας $\mathbf{u}(k) \in R^2$ ως εξής (βλ. Σχ. 7.6)

$$u_1(k) \in \mathscr{U}_1 = \{0.8, 0.4, 0, -0.4\}, \quad u_2(k) \in \mathscr{U}_2 = \{0.4, 0, -0.4\}. \qquad (7.29a)$$

Η ταχύτητα $\mathbf{z}(k)$ του εμποδίου θεωρείται στοχαστική διεργασία με

$$z_i(k) \in \{0.8, 0.4, 0, -0.4, -0.8\}, \; i = 1, 2 \qquad (7.29b)$$

και ισοκατανομή πιθανοτήτων γι΄ αυτές τις τιμές. Οι τιμές **z**(k) είναι ανεξάρτητες των **x**(k), **u**(k) καθώς και των **z**(k–1), **z**(k–2),.... Για την εκπλήρωση της περιγραφείσας αποστολής του οχήματος ορίζουμε το ακόλουθο κριτήριο κόστους προς ελαχιστοποίηση

$$J = E\{ \sum_{k=0}^{\infty} [\alpha_1(0.8 - u_1(k))^2 + \alpha_2 u_2(k)^2 + \psi(\mathbf{x}(k))]\} \qquad (7.30)$$

με

$$\psi(\mathbf{x}) = \begin{cases} 1 & \text{αν} \quad |x_1| \le \frac{l_1}{2} \quad \text{και} \ |x_2| \le \frac{l_2}{2} \\ 0 & \text{άλλως} \end{cases}$$

όπου έχουμε $l_1 = 1.6$ και $l_2 = 4$ τις γνωστές διαστάσεις του εμποδίου (επαυξημένες με τις διαστάσεις του οχήματος) και τις παραμέτρους βάρους $\alpha_1 = 0.05$ και $\alpha_2 = 0.1$. Το κριτήριο κόστους (7.30) τιμωρεί, μέσω των πρώτων δύο όρων, αποκλίσεις από την επιθυμητή πορεία του οχήματος και, μέσω του τρίτου όρου, μια πιθανή σύγκρουση με το εμπόδιο. Επειδή οι δύο αυτές προδιαγραφές είναι μερικά ανταγωνιστικές (γιατί;), πρέπει η επιλογή των βαρών α_1, α_2 να γίνει κατάλληλα, έτσι ώστε να επιτευχθεί ένας καλός συμβιβασμός. Αν οι τιμές των α_1, α_2 επιλεγούν υπέρ το δέον υψηλές, τότε το όχημα θα προχωρεί γρήγορα και χωρίς αποκλίσεις προς τα εμπρός αδιαφορώντας για ενδεχόμενες συγκρούσεις, ενώ αν επιλεγούν υπέρ το δέον χαμηλές, το όχημα θα κινείται διστακτικά ή και θα υποχωρεί στην θέα και μόνον ενός εμποδίου. Ο χρονικός ορίζοντας επελέγει άπειρος, καθότι είναι άγνωστη η διάρκεια των χειρισμών για την αποφυγή μιας σύγκρουσης. Με τα δεδομένα αυτά έχουμε συγκεντρωτικά το ακόλουθο πρόβλημα στοχαστικού βέλτιστου ελέγχου:

Να προσδιορισθεί ένας βέλτιστος κανόνας ελέγχου

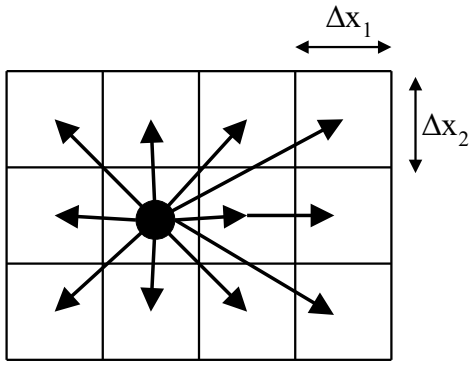
$$\mathbf{u}(k) = \mathbf{R}[\mathbf{x}(k)] \qquad (7.31)$$

που ελαχιστοποιεί το κριτήριο κόστους (7.30) λαμβάνοντας υπόψη τους περιορισμούς (7.28) και (7.29a) και την πιθανολογική κατανομή της ταχύτητας εμποδίου.

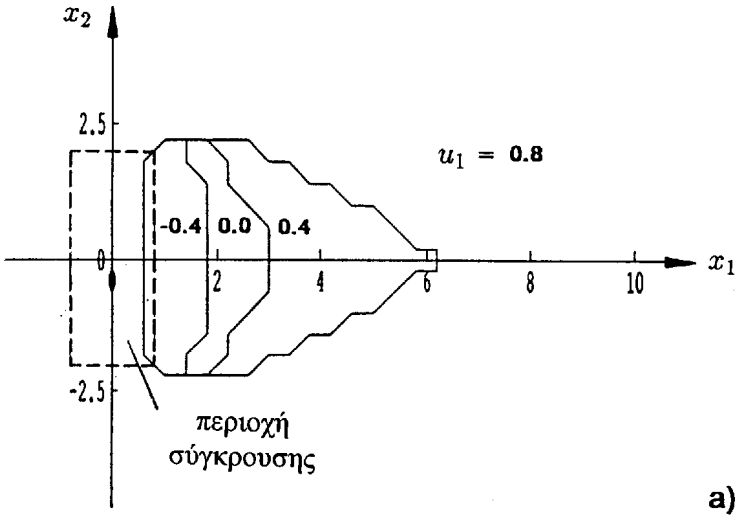
Οι παραδοχές του παρόντος κεφαλαίου εκπληρούνται, και άρα αναμένουμε ένα μη μεταβλητό βέλτιστο κανόνα ελέγχου. Για την εφαρμογή του Διακριτού Στοχαστικού Δυναμικού Προγραμματισμού προβαίνουμε σε διακριτοποίηση της περιοχής τιμών $\mathcal{X}$ με διακριτά διαστήματα $\Delta x_1 = \Delta x_2 = 0.4$, βλ. Σχ. 7.6. Η αριθμητική επίλυση του προβλήματος μέχρι τη σύγκλιση προς ένα στάσιμο κανόνα ελέγχου απαίτησε τον υπολογισμό K = 24 βαθμίδων, για τον οποίο χρειάστηκαν 25min 27s υπολογιστικού χρόνου σε έναν (σήμερα απαρχαιωμένο) σταθμό εργασίας VAX 2000.

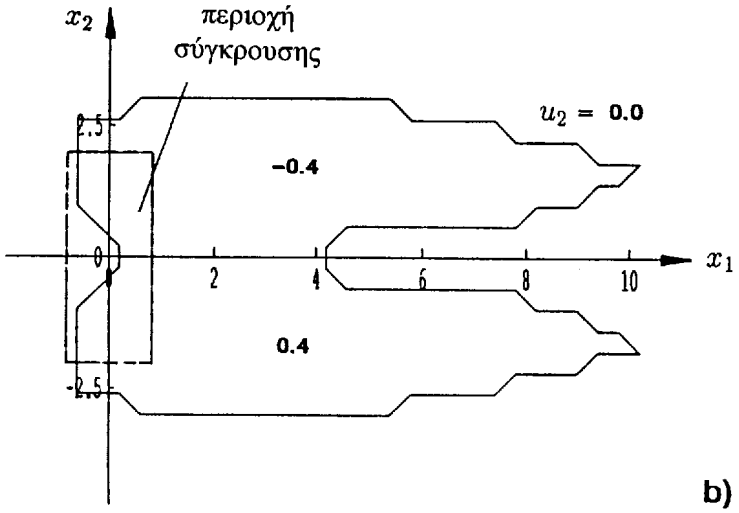
Το Σχ. 7.7 αναπαριστά την σχηματική μορφή του υπολογισθέντος κανόνα ελέγχου (7.31), όπου η αρχή των αξόνων αντιστοιχεί στην εκάστοτε θέση του οχήματος, ενώ το διακεκομμένο ορθογώνιο παραλληλόγραμμο καθορίζει την περιοχή σύγκρουσης. Η περιοχή $\mathcal{X}$ περί το όχημα χωρίζεται στο Σχ. 7.7 σε τομείς με αντίστοιχες τιμές ταχύτητας για το όχημα με την εξής έννοια: Αν το κέντρο του εμποδίου σε κάποια χρονική στιγμή k βρίσκεται σε κάποιον τομέα, τότε το όχημα πρέπει να κινηθεί με τις αντίστοιχες αναγραφόμενες βέλτιστες ταχύτητες $u_1(k)$, $u_2(k)$ που, άρα, είναι συναρτήσεις του **x**(k) όπως απαιτεί η (7.31).

Το Σχ. 7.7a δείχνει ότι το όχημα μειώνει την ταχύτητα $u_1(k)$ ή σταματά ή οπισθοδρομεί, αν το εμπόδιο βρίσκεται μπροστά του στους αντίστοιχους τομείς. Αποκλίσεις $u_2(k)$ πραγματοποιούνται σύμφωνα με το Σχ. 7.7b αν η παρούσα θέση του εμποδίου θεωρείται

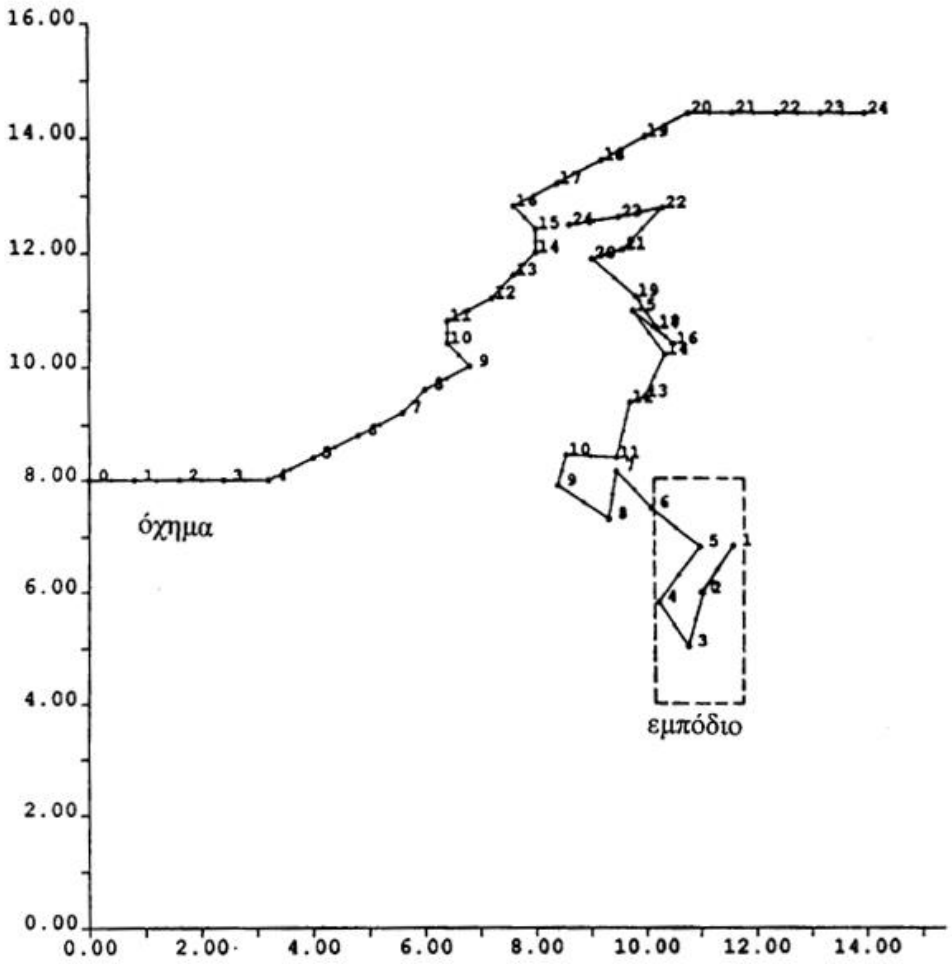


Σχήμα 7.6: Δυνατότητες κίνησης του οχήματος σε κάθε χρονικό διάστημα.





Σχήμα 7.7: Βέλτιστος κανόνας ελέγχου για την κίνηση αυτόνομου οχήματος: (a) $u_1(\mathbf{x})$ και (b) $u_2(\mathbf{x})$.



Σχήμα 7.8: Σενάριο αποφυγής εμποδίου από αυτόνομο όχημα.

επικίνδυνη για σύγκρουση. Είναι ενδιαφέρον να παρατηρήσουμε ότι το όχημα αργεί να αποκλίνει αν το εμπόδιο βρίσκεται ακριβώς μπροστά του, κοντά στον άξονα x_1. Ο λόγος αυτής της φλεγματικής συμπεριφοράς έγκειται στο ότι θεωρήσαμε το εμπόδιο ως στοχαστικά κινούμενο, και επομένως το όχημα, για να αποφύγει ενδεχόμενα ζιγκ-ζαγκ που κοστίζουν, περιμένει να διαπιστώσει την κατεύθυνση ενδεχόμενης απόκλισης του εμποδίου έτσι ώστε το ίδιο να αποκλίνει προς την αντίθετη κατεύθυνση.

Το Σχ. 7.8 δείχνει ένα σενάριο εφαρμογής του κανόνα ελέγχου του Σχ. 7.7 με ένα εμπόδιο, του οποίου οι ταχύτητες $z_i(k)$, $i = 1, 2$, επελέγησαν με ισότιμη πιθανότητα από το πεδίο τιμών $[-1, 1]$, υπάρχει δηλαδή απόκλιση από την παραδοχή (7.29b) βάσει της οποίας προσδιορίσθηκε ο κανόνας ελέγχου. Οι δύο τροχιές του Σχ. 7.8 φέρουν τους διακριτούς χρονικούς δείκτες κατά την κίνηση του οχήματος και του εμποδίου, του οποίου οι διαστάσεις απεικονίζονται μόνο στην αρχική του θέση. Το Σχ. 7.8 δείχνει πως το όχημα επιχειρεί καταρχήν να αποκλίνει προς τα αριστερά, η γρήγορη προσέγγιση του εμποδίου το οδηγεί όμως σε οπισθοδρόμηση ($k = 9$). Μια δεύτερη επιχείρηση καταλήγει παρόμοια ($k=15$), αλλά η τρίτη επιχείρηση αποδεικνύεται επιτυχής.

■

ε) Προβλήματα με Ελλιπή Πληροφορία

Στην διατύπωση του ΠΣΒΕ στο Κεφάλαιο 7.2 θεωρήσαμε ότι η κατάσταση **x**(k) του υπό έλεγχο συστήματος είναι μετρήσιμη με ικανή ακρίβεια σε πραγματικό χρόνο. Σε πολλές πρακτικές εφαρμογές συμβαίνει όμως:

– Να μην είναι ακριβείς οι μετρήσεις κατάστασης ή
– Να μην είναι μετρήσιμες όλες οι μεταβλητές κατάστασης ή
– Να είναι υπέρ το δέον δαπανηρή η πλήρης μέτρηση κατάστασης, οπότε ενδιαφέρουν ενδεχόμενες εναλλακτικές λύσεις ή
– Να είναι η εκτίμηση κατάστασης ο στόχος και όχι το μέσον επίλυσης του προβλήματος βελτιστοποίησης.

Για τις περιπτώσεις αυτές θεωρούμε ότι οι διαθέσιμες μετρήσεις σε πραγματικό χρόνο είναι

$$\mathbf{y}(k) = \mathbf{c}[\mathbf{x}(k), \mathbf{u}(k-1), \mathbf{v}(k), k], \quad k = 1,\ldots,K-1 \qquad (7.32)$$

όπου $\mathbf{y}(k) \in \mathcal{Y} \subset R^q$ είναι οι μετρήσιμες **μεταβλητές εξόδου** του συστήματος. Η **διαταραχή μετρήσεως** $\mathbf{v}(k) \in \mathcal{V}(k)$ είναι μια τυχαία μεταβλητή με γνωστή πιθανολογική κατανομή που μπορεί να εξαρτάται από τον χρόνο, την κατάσταση και τον έλεγχο, είναι όμως ανεξάρτητη των $\mathbf{u}(k-1),\ldots,\mathbf{u}(0)$ και $\mathbf{z}(k),\ldots,\mathbf{z}(0)$. Επιπλέον, η μη μετρήσιμη αρχική κατάσταση $\mathbf{x}(0)$ θεωρείται τυχαία μεταβλητή με δεδομένη πιθανολογική κατανομή. Τέλος, η επιτρεπτή περιοχή ελέγχου $\mathcal{U}(k)$ θεωρείται, για το πρόβλημα με ελλιπή πληροφορία, ανεξάρτητη της κατάστασης καθότι η τελευταία δεν είναι πλέον γνωστή με ακρίβεια. Συγκεντρώνοντας την διαθέσιμη πληροφορία για κάποιο χρονικό διάστημα k στο **διάνυσμα πληροφορίας**

$$\mathbf{I}(k) = [\mathbf{y}(0)^T\ \mathbf{u}(0)^T\ \mathbf{y}(1)^T \ldots \mathbf{y}(k-1)^T\ \mathbf{u}(k-1)^T\ \mathbf{y}(k)^T]^T \qquad (7.33)$$

μπορούμε να διατυπώσουμε το ΠΣΒΕ με ελλιπή πληροφορία αντικαθιστώντας την (7.2) με την

$$\mathbf{u}(k) = \mathbf{R}[\mathbf{I}(k), k], \quad k = 0,\ldots, K-1. \qquad (7.34)$$

Η επίλυση του προβλήματος αυτού είναι καταρχήν δυνατή μέσω Στοχαστικού Δυναμικού Προγραμματισμού, ο απαιτούμενος υπολογιστικός φόρτος αυξάνει όμως (πλην ειδικών κατηγοριών προβλημάτων) ακόμη περισσότερο. Σε κάθε περίπτωση, η λεπτομερής παρουσίαση των αντίστοιχων επεκτάσεων της βασικής μεθοδολογίας βρίσκεται πέραν των ορίων του παρόντος συγγράμματος.

Προβλήματα

7.1 Αναπτύξτε την αναδρομική εξίσωση για τον υπολογισμό του βέλτιστου, χρονικά μεταβλητού κατωφλίου αποφάσεων α(k) για την επίλυση του προβλήματος προθεσμίας του Παραδ. 7.6 χρησιμοποιώντας την παραδοχή ότι οι προσφορές z(k) είναι ισοκατανεμημένες στην περιοχή $[z_{min}, z_{max}]$, δηλαδή

$$P(z) = \begin{cases} 1 & \text{αν} \quad z > z_{max} \\ (z - z_{min})/(z_{max} - z_{min}) & \text{αν} \quad z_{min} \le z \le z_{max} \\ 0 & \text{αν} \quad z < z_{min}. \end{cases}$$

Υπολογίστε την στάσιμη βέλτιστη πολιτική αποφάσεων για την περίπτωση $K \to \infty$ για τις αριθμητικές τιμές $z_{min} = 1$, $z_{max} = 2$, $r = 0.01$.

7.2 (**Προβλήματα Min-Max**) Στο βασικό ΠΣΒΕ του Κεφαλαίου 7.2 μπορούμε να θεωρήσουμε την εξής παραλλαγή. Για τις διαταραχές $\mathbf{z}(0),...,\mathbf{z}(K-1)$ δεν υπάρχουν πιθανολογικές κατανομές, παρά μόνο πεδία τιμών $\mathcal{Z}[\mathbf{x}(k), \mathbf{u}(k), k] \subset \mathcal{D}(k)$, $k=0,..., K-1$, που ενδεχομένως εξαρτώνται από $\mathbf{x}(k)$, $\mathbf{u}(k)$. Θεωρούμε το πρόβλημα προσδιορισμού ενός κανόνα ελέγχου $\mathbf{u}(k) = \mathbf{R}[\mathbf{x}(k),k] \in \mathcal{U}[\mathbf{x}(k), k]$ $\forall[\mathbf{x}(k), k]$ που ελαχιστοποιεί το κριτήριο κόστους

$$J = \max\{\vartheta[\mathbf{x}(K)] + \sum_{k=0}^{K-1} \varphi[\mathbf{x}(k), \mathbf{u}(k), \mathbf{z}(k), k]\}$$

όπου το ως άνω μέγιστο νοείται ως προς $\mathbf{z}(k) \in \mathcal{Z}[\mathbf{x}(k), \mathbf{u}(k), k]$, $k = 0,...,K-1$. Αναπτύξτε έναν αλγόριθμο Δυναμικού Προγραμματισμού για το πρόβλημα αυτό.

7.3 Μια βιομηχανική μονάδα παράγει 2 τύπους τετραδίων σε ποσότητες $u_1(k)$ και $u_2(k)$, αντίστοιχα, στην αρχή μιάς περιόδου k. Η μέγιστη δυνατότητα παραγωγής της βιομηχανίας είναι U_1 και U_2 για τα τετράδια τύπου 1 και τύπου 2, αντίστοιχα. Η βιομηχανία διατηρεί αποθήκη μέγιστης χωρητικότητας X, στην οποία διατηρεί τις τυχόν περισσευούμενες και τις άμεσα παραγόμενες ποσότητες τετραδίων, δημιουργώντας απόθεμα $x_1(k)$ και $x_2(k)$ αποθηκευμένων τετραδίων τύπου 1 και 2, αντίστοιχα. Η ζήτηση των τετραδίων $z_1(k)$ και $z_2(k)$ είναι στοχαστική με γνωστή πιθανολογική κατανομή $p_1(z_1(k)|k)$ και $p_2(z_2(k)|k)$, αντιστοίχως, για τα τετράδια τύπου 1 και 2, με $0 \le z_1(k) \le Z_1$ και $0 \le z_2(k) \le Z_2$. Η βιομηχανία δεν αφήνει εκκρεμείς παραγγελίες για το μέλλον, και έτσι, αν η ζήτηση $z_1(k)$ και $z_2(k)$ ξεπεράσει το άθροισμα $x_1(k)+u_1(k)$ και $x_2(k)+u_2(k)$ αντίστοιχα, η υπερζήτηση χάνεται και το απόθεμα μηδενίζεται (άρα $x_1(k) \ge 0$, $x_2(k) \ge 0$). Τα έσοδα πώλησης (τιμή πώλησης μείον κόστος παραγωγής) ανά μονάδα τετραδίων είναι α_1 και α_2, ενώ τα κόστη αποθήκευσης ανά μονάδα τετραδίων είναι b_1 και b_2 για τα τετράδια τύπου 1 και 2, αντίστοιχα. Βάση αυτών των δεδομένων η βιομηχανία σκοπεύει να αναπτύξει μια πολιτική για την παραγωγή τετραδίων για τις επόμενες K χρονικές περιόδους που να μεγιστοποιεί το αναμενόμενο ολικό κέρδος (έσοδα πώλησης μείον κόστη αποθήκευσης). Για την επίλυση του προβλήματος αυτού θεωρείται διακριτοποίηση $\Delta = \Delta x_i = \Delta u_i = \Delta z_i$ όπου X/Δ, U_i/Δ, Z_i/Δ, $i = 1,2$, ακέραιοι.

(a) Να προσδιορισθούν:
   (i) η καταστατική εξίσωση που εκφράζει τη δυναμική εξέλιξη του περιεχομένου της αποθήκης της βιομηχανίας,
   (ii) όλοι οι περιορισμοί του προβλήματος και
   (iii) το αντικειμενικό κριτήριο που εκφράζει το αναμενόμενο ολικό κέρδος.

(b) Να αναπτυχθεί ένας ψευδοκώδικας (ή διάγραμμα ροής) προγράμματος, το οποίο, δεχόμενο ως δεδομένα εισόδου:
   - τις τιμές Δ, K, X, U_1, U_2, Z_1, Z_2, a_1, a_2, b_1, b_2, και
   - τις πιθανολογικές κατανομές $p_i(j\Delta|k)$, $i=1,2$, $j=0,1,...,Z_i/\Delta$, $k=0,1,...,K-1$,

   υπολογίζει το βέλτιστο κανόνα ελέγχου υπό μορφή πίνακα $u_i = R_i(j\Delta, m\Delta, k)$, $i = 1,2$, $j = 0,1,..., X/\Delta$, $m = 0,1...,X/\Delta$, $j+m \le X/\Delta$, $k = 0,1,...,K-1$.

7.4 Ένα πρακτορείο τύπου θέλει να προγραμματίσει τις παραγγελίες που θα κάνει τις έξι πρώτες ημέρες της εβδομάδας, αρχίζοντας από τη Δευτέρα. Το πρακτορείο διαθέτει προς πώληση έναν αριθμό n εφημερίδων, με τιμή πώλησης κάθε φύλλου της εφημερίδας i ίση με μ_i ευρώ, i = 1,...,n. Το πρωί κάθε ημέρας k = 0,...,5 παραγγέλλεται μια ποσότητα $u_i(k)$ φύλλων της εφημερίδας i. Η παραγγελία εκτελείται αμέσως, δηλαδή ο χρόνος άφιξης των εφημερίδων που παραγγέλλονται κάθε πρωί είναι αμελητέος. Η τιμή αγοράς ενός φύλλου της εφημερίδας i είναι ρ_i ευρώ. Κάθε ημέρα k υπάρχει μια στοχαστική ζήτηση $d_i(k)$ για την εφημερίδα i. Τα $d_i(k)$ είναι ανεξάρτητα μεταξύ τους ως προς το χρόνο k και το είδος i της εφημερίδας. Αν κάποιος πελάτης ζητήσει μια εφημερίδα της οποίας τα φύλλα έχουν εξαντληθεί, τότε δεν αγοράζει καμία εφημερίδα από το εν λόγω πρακτορείο και πηγαίνει σε κάποιο άλλο.

Αφού γίνουν όλες οι πωλήσεις, τα φύλλα που έχουν περισσέψει αποθηκεύονται σε μια κοινή αποθήκη χωρητικότητας X φύλλων εφημερίδων οποιουδήποτε είδους. Αν η αποθήκη είναι γεμάτη, τότε τα φύλλα που περισσεύουν πετιούνται. Έστω $\alpha_i[\mathbf{x}(k), \mathbf{u}(k), \mathbf{d}(k)]$ ο αριθμός των φύλλων της εφημερίδας i που πετιούνται την ημέρα k και $\mathbf{x}(k)$ ένα διάνυσμα της μορφής $[x_1(k) \ldots x_n(k)]^T$ όπου $x_i(k)$ είναι ο αριθμός των φύλλων της εφημερίδας i που έχουν συγκεντρωθεί από την αρχή της εβδομάδας μέχρι και την ημέρα k στην αποθήκη.

Την έβδομη ημέρα της εβδομάδας δεν γίνονται παραγγελίες και πωλήσεις. Το μόνο που γίνεται είναι το άδειασμα της αποθήκης. Όλα τα φύλλα των εφημερίδων που βρίσκονται στην αποθήκη επιστρέφονται. Για κάθε φύλλο της εφημερίδας i που επιστρέφεται, δίδεται στο πρακτορείο ένα ποσό l_i ευρώ, όπου $l_i<\rho_i$.

Από τα παραπάνω γίνεται κατανοητό ότι η δυναμική του αποθέματος των εφημερίδων έχει ως εξής:

$$x_i(k+1) = x_i(k)+\max\{0, u_i(k)-d_i(k)\} - \alpha_i[\mathbf{x}(k), \mathbf{u}(k), \mathbf{d}(k)], \quad k = 0,\ldots,5,\ i = 1,\ldots,n.$$

Ο υπολογισμός των $\alpha_i[\mathbf{x}(k), \mathbf{u}(k), \mathbf{d}(k)]$ γίνεται με τέτοιο τρόπο ώστε να ελαχιστοποιείται η ζημία $\sum_{i=1}^{n}\alpha_i[\mathbf{x}(k), \mathbf{u}(k), \mathbf{d}(k)]\,(\rho_i-l_i)$.

Το κριτήριο κέρδους που πρέπει να μεγιστοποιηθεί έχει προφανώς ως εξής:

$$J = E\{\vartheta[\mathbf{x}(6)] + \sum_{k=0}^{5}\varphi[\mathbf{x}(k), \mathbf{u}(k), \mathbf{d}(k),k]\}$$

όπου

$$\vartheta[\mathbf{x}(6)] = \sum_{i=1}^{n}x_i(6)l_i$$

$$\varphi[\mathbf{x}(k), \mathbf{u}(k), \mathbf{d}(k), k] = \sum_{i=1}^{n}\{d_i(k)\mu_i-[u_i(k)+\alpha_i[\mathbf{x}(k), \mathbf{u}(k), \mathbf{d}(k)]]\rho_i\}.$$

(a) Κατασκευάστε έναν αλγόριθμο ή διάγραμμα ροής προγράμματος για τον προσδιορισμό των $\alpha_i[\mathbf{x}(k), \mathbf{u}(k), \mathbf{d}(k)]$, $i = 1,...,n$.

(b) Μορφοποιήστε το πρόβλημα των παραγγελιών του πρακτορείου σαν ένα πρόβλημα βέλτιστου ελέγχου αναγράφοντας:
   (i) Το κριτήριο κέρδους
   (ii) Την καταστατική εξίσωση
   (iii) Το πεδίο τιμών της μεταβλητής κατάστασης $\mathbf{x}(k)$
   (iv) Τους περιορισμούς ελέγχου.

(c) Κατασκευάστε το διάγραμμα ροής ενός προγράμματος που επιλύει το πρόβλημα. Σαν είσοδο θα πρέπει να δέχεται τα ρ_i, μ_i, l_i, X και τις κατανομές των $d_i(k)$. Η έξοδος θα πρέπει να είναι ο πίνακας βέλτιστης πολιτικής παραγγελιών (κλειστού βρόχου) και το αναμενόμενο μέγιστο κέρδος.

7.5 Θεωρούμε το στοχαστικό παίγνιο των τεσσάρων σπίρτων το οποίο έχει ως εξής: Τέσσερα σπίρτα τοποθετούνται πάνω σ' ένα τραπέζι και δύο παίκτες παίζουν εναλλάξ σύμφωνα με τους ακόλουθους κανόνες. Κάθε φορά που παίζει ένας παίκτης, πρέπει να αποφασίσει αν θα σηκώσει ένα ή δύο σπίρτα και να ανακοινώσει την απόφασή του στον αντίπαλο παίκτη. Στη συνέχεια ρίχνει ένα κέρμα και προσπαθεί να μαντέψει το αποτέλεσμα (κορώνα ή γράμματα). Τα δύο ενδεχόμενα (κορώνα, γράμματα) θεωρούνται ισοπίθανα. Αν μαντέψει σωστά το αποτέλεσμα, πράττει αυτό που έχει ανακοινώσει στον αντίπαλό του, αλλιώς χάνει τη σειρά του χωρίς να κάνει τίποτε και παίζει ο επόμενος παίκτης. Κερδίζει όποιος παίκτης σηκώσει από το τραπέζι το τελευταίο σπίρτο.

(a) Θεωρώντας ότι η τρέχουσα κατάσταση του παιγνίου χαρακτηρίζεται από τον αριθμό των σπίρτων που βρίσκονται πάνω στο τραπέζι, αναπτύξτε το παίγνιο από την κατάσταση εκκίνησης (4 σπίρτα) μέχρι βάθος 4 (δύο κινήσεις για κάθε παίκτη).
(<u>Υπόδειξη:</u> Κάθε κατάσταση του παιγνίου αρκεί να εμφανίζεται μία μόνο φορά σε κάθε βαθμίδα.)

(b) Υποθέτουμε ότι μετά από τις 4 κινήσεις (δύο για κάθε παίκτη) το παίγνιο τερματίζει. Αν πάνω στο τραπέζι υπάρχουν ακόμη σπίρτα, το αποτέλεσμα είναι ισοπαλία. Εφαρμόστε την διαδικασία min–max για την επίλυση του παιγνίου λαμβάνοντας υπόψη τα εξής:
   - Οι τιμές των κόμβων τελικής έκβασης του παιγνίου έχουν ως εξής:
     +1, αν κερδίζει ο παίκτης από τη σκοπιά του οποίου επιλύετε το παίγνιο
     –1, αν κερδίζει ο άλλος παίκτης
     0, σε κάθε άλλη περίπτωση (ισοπαλία).
   - Οι τιμές των ενδιάμεσων κόμβων αντιστοιχούν σε αναμενόμενες τιμές εξαιτίας του στοχαστικού χαρακτήρα του παιγνίου.

   Σημειώστε στο αντίστοιχο σχήμα τις βέλτιστες αποφάσεις των δύο παικτών (με βέλη), καθώς και τις αναμενόμενες τιμές των κόμβων.

(c) Εξηγήστε τη φυσική σημασία των αναμενόμενων τιμών των κόμβων.

7.6 Μια ναυτιλιακή εταιρία διαθέτει ένα άδειο κοντέινερ χωρητικότητας X (σε βάρος). Κάθε περίοδο k=0, 1,…, K−1 εμφανίζεται ένας πελάτης που προσφέρεται να πληρώσει h_k ώστε ένα αντικείμενο βάρους s_k (ακέραιος αριθμός) να φορτωθεί στο κοντέινερ. Οι προσφορές $[h_k\ s_k]^T=\mathbf{z}_k \in \mathbb{Z}$, k=0, 1,…, K−1, είναι τυχαίες, ανεξάρτητες μεταξύ τους και με γνωστή, μη μεταβλητή πιθανολογική κατανομή. Κάθε περίοδο k η εταιρία επιλέγει να αποδεχθεί (στην περίπτωση που υπάρχει ακόμη χώρος στο κοντέινερ) ή να απορρίψει την προσφορά. Αντικειμενικός στόχος της εταιρίας είναι να μεγιστοποιήσει τα αναμενόμενα έσοδα από την χρήση του κοντέινερ κατά το χρονικό διάστημα K.

(a) Διατυπώστε το πρόβλημα με το κριτήριο κόστους, την καταστατική εξίσωση και όλους τους περιορισμούς.

(b) Προσδιορίστε τον αναδρομικό τύπο του Bellman για το συγκεκριμένο πρόβλημα διακριτού στοχαστικού δυναμικού προγραμματισμού.

(c) Επιλύστε το πρόβλημα για K=2, X=3 και $\mathbf{z}_k \in \{[1\ 1]^T, [3\ 2]^T\}$, με

$P(\mathbf{z}_k=[1\ 1]^T)=0.6$
$P(\mathbf{z}_k=[3\ 2]^T)=0.4.$

7.7 Θεωρούμε το εξής πρόβλημα ρίψης δύο νομισμάτων (1 και 2) το οποίο εξελίσσεται σε δύο γύρους. Δύο κάλπικα νομίσματα είναι πάνω στο τραπέζι και σε κάθε γύρο στρίβουμε ένα από τα δύο νομίσματα (δηλαδή μπορούμε να στρίψουμε δύο φορές το ίδιο νόμισμα) και βλέπουμε το αποτέλεσμα Κορώνα (Κ) ή Γράμματα (Γ). Ανάλογα με τα αποτελέσματα στο τέλος των δύο γύρων κερδίζουμε ή χάνουμε κάποια χρήματα ως εξής:

(a) Αν και τις δύο φορές το αποτέλεσμα είναι Κ, κερδίζουμε 3.

(b) Αν και τις δύο φορές το αποτέλεσμα είναι Γ, κερδίζουμε 1.

(c) Αν την πρώτη φορά το αποτέλεσμα είναι Κ και την δεύτερη Γ (ή Γ και Κ, αντίστοιχα), χάνουμε 2.

Οι πιθανότητες του πρώτου νομίσματος να φέρει Κ ή Γ είναι 3/4 και 1/4, αντίστοιχα, ενώ οι ίδιες πιθανότητες για το δεύτερο νόμισμα είναι 1/3 και 2/3, αντίστοιχα. Προσδιορίστε την βέλτιστη στρατηγική του παίκτη (δηλαδή ποιο νόμισμα να στρίψει σε κάθε γύρο) ώστε να μεγιστοποιείται η αναμενόμενη αμοιβή του στο τέλος των δύο γύρων.

Υπόδειξη: Προσεγγίστε το πρόβλημα στοχαστικού δυναμικού προγραμματισμού κατά το πρόβλημα του απρόσεκτου οδηγού οδικού δικτύου του Παραδείγματος 7.1.

7.8 Οι Ολυμπιακές Αερογραμμές (ΟΑ) σχεδιάζουν μια πολιτική κρατήσεων για την πτήση Αθήνα-Χανιά η οποία εκτελείται με ένα αεροσκάφος χωρητικότητας X θέσεων και τιμή εισιτηρίου *v*. Η πολιτική κρατήσεων σχεδιάζεται για έναν ορίζοντα K ωρών πριν την ώρα πτήσης. Έστω x(k) ο αριθμός των κρατήσεων στην αρχή κάθε περιόδου $k \in [0,K]$. Οι ΟΑ λαμβάνοντας υπόψη το γεγονός ότι κάθε επιβάτης με κράτηση μπορεί να ακυρώσει την κράτησή του και μάλιστα ατελώς, ακολουθούν μια πολιτική υπερκρατήσεων για να αντισταθμίσουν τις πιθανές ακυρώσεις κρατήσεων και κατά συνέπεια την απώλεια εσόδων

(δηλαδή μπορεί x(k)>X και η τελική κατάσταση x(K) είναι ελεύθερη). Εάν ένας επιβάτης με κράτηση φθάσει την ώρα της πτήσης στο αεροδρόμιο και δεν υπάρχει διαθέσιμη θέση (λόγω υπερκράτησης), οι ΟΑ καταβάλουν στον επιβάτη μια χρηματική ρήτρα μ (πέραν της επιστροφής του αντιτίμου ν). Έτσι, τα έσοδα των ΟΑ από την πτήση εξαρτώνται από τον τελικό αριθμό επιβατών μείον τις ενδεχόμενα καταβληθείσες ρήτρες.

Έστω $z(k) \in [0,\text{int}(x(k)/2)]$ ο αριθμός ακυρωμένων κρατήσεων κατά την διάρκεια της περιόδου k και $d(k) \in [0,D]$ η ζήτηση νέων κρατήσεων κατά την διάρκεια της περιόδου k. Οι d(k), z(k) είναι στοχαστικές μεταβλητές με γνωστές πιθανολογικές κατανομές p[d(k)] και p[z(k)|x(k)], αντίστοιχα. Στην αρχή κάθε περιόδου k, δεδομένου του αριθμού προηγουμένων κρατήσεων x(k), η εταιρία αποφασίζει το μέγιστο όριο ($u(k) \in [0,D]$) αιτήσεων νέων κρατήσεων που θα γίνουν αποδεκτές κατά την διάρκεια της τρέχουσας περιόδου. Η εταιρεία θέτει ως ανώτατο όριο υπερκράτησης λX θέσεις (όπου λ ακέραιος) και για το σκοπό αυτό τίθεται ο περιορισμός $u(k) \le \lambda X - x(k)$. Ζητείται η βέλτιστη πολιτική u(k)=R[x(k),k] αποδοχής νέων κρατήσεων με στόχο την μεγιστοποίηση των αναμενόμενων εσόδων από την πτήση.

(a) Η καταστατική εξίσωση που περιγράφει την εξέλιξη των κρατήσεων είναι, σύμφωνα με τα παραπάνω, η εξής:

$$x(k+1) = x(k) - z(k) + \min\{u(k),d(k)\}$$

με επιτρεπτή περιοχή ελέγχου $0 \le u(k) \le \min\{D, \lambda X - x(k)\}$.

Επίσης, το αντικειμενικό κριτήριο εσόδων είναι:

$$J = \nu \cdot \min\{x(k),X\} - \mu \cdot \max\{0, x(k) - X\}.$$

Εξηγήστε τις εξισώσεις αυτές.

(b) Αναπτύξτε ένα πρόγραμμα που θα επιλύει το πρόβλημα των κρατήσεων. Σαν δεδομένα εισόδου το πρόγραμμα θα δέχεται τα ν, μ, λ, X, K, D. Η πιθανολογική κατανομή των d(k), z(k) θεωρείται ομοιόμορφη εντός των αντίστοιχων ορίων των μεταβλητών. Η επίλυση του προβλήματος να γίνει για $0 \le x(k) \le \lambda X$. Η έξοδος του προγράμματος θα πρέπει να είναι ο πίνακας R[x(k),k] βέλτιστης πολιτικής κρατήσεων (κλειστού βρόχου).

(c) Χρησιμοποιήστε το πρόγραμμα του ερωτήματος (b) για να βρείτε τη βέλτιστη πολιτική u(k)=R[x(k),k] αποδοχής νέων κρατήσεων, για ν=100, μ=50, λ=2, X=12, K=7, D=4 και ομοιόμορφες πιθανολογικές κατανομές των d(k), z(k).

7.9 Ένα κατάστημα σκαφών αναψυχής θέλει να προγραμματίσει τις παραγγελίες σε σκάφη που θα κάνει μέσα στους επόμενους δύο μήνες, K=2. Στην αρχή κάθε μήνα το κατάστημα παραγγέλνει από την αντιπροσωπεία έναν αριθμό σκαφών $u(k) \in \{0,1,2\}$ και η παραγγελία εκτελείται αμέσως, δηλαδή ο χρόνος άφιξης των σκαφών είναι αμελητέος. Η τιμή αγοράς ενός σκάφους από την αντιπροσωπεία κοστίζει 5000 €. Κάθε μήνα, k=0, 1, υπάρχει μία στοχαστική ζήτηση σκαφών z(k) από πελάτες, η οποία δεν είναι γνωστή εκ των προτέρων, αλλά είναι γνωστή η πιθανολογική κατανομή της όπως φαίνεται στον επόμενο πίνακα. Οι ποσότητες ζήτησης z(k) κάθε περιόδου *k* είναι ανεξάρτητες μεταξύ τους.

	p(z=0)	**p(z=1)**	**p(z=2)**
k=0	0.2	0.5	0.3
k=1	0.1	0.2	0.7

Το κατάστημα ενοικιάζει έναν αποθηκευτικό χώρο με μέγιστη χωρητικότητα 2 σκάφη και κόστος αποθήκευσης 1000 € ανά σκάφος. Σε περίπτωση που η ζήτηση z(k) υπερβεί το άθροισμα της αποθηκευμένης και παραγγελθείσας ποσότητας, η ζήτηση δεν καλύπτεται στα επόμενα χρονικά διαστήματα, δηλαδή δε νοείται αρνητικό απόθεμα. Το κόστος μη καλυφθείσας ζήτησης είναι 8000 €/σκάφος. Στην αρχή του χρονικού ορίζοντα που μελετάται, η αποθήκη έχει 1 σκάφος. Η μέγιστη ποσότητα παραγγελίας στην αρχή κάθε μήνα είναι ίση με το διαθέσιμο ελεύθερο χώρο στην αποθήκη. Ζητείται η βέλτιστη πολιτική παραγγελιών σκαφών που πρέπει να ακολουθήσει το κατάστημα προκειμένου να ελαχιστοποιήσει το αναμενόμενο κόστος.

(a) Διατυπώστε το παραπάνω πρόβλημα αποθήκευσης ως Πρόβλημα Στοχαστικού Βέλτιστου Ελέγχου με την αντικειμενική συνάρτηση, την καταστατική εξίσωση και όλους τους περιορισμούς.

(b) Επιλύστε αριθμητικά και γραφικά το πρόβλημα προσδιορίζοντας τη βέλτιστη πολιτική παραγγελιών σκαφών του καταστήματος, u(k), που ελαχιστοποιεί το ολικό αναμενόμενο κόστος λαμβάνοντας υπόψη τους παραπάνω περιορισμούς.

7.10 Δύο φίλοι σχεδιάζουν να πάνε διακοπές και να επισκεφτούν μέσα σε K = 2 ημέρες δύο διαφορετικές πόλεις στις οποίες και θα διανυκτερεύσουν. Από έρευνα στο διαδίκτυο έχουν βρει το οικονομικότερο ξενοδοχείο για κάθε πόλη, οι τιμές (ανά άτομο) των οποίων φαίνονται στον Πίνακα 1. Μία εναλλακτική επιλογή είναι, αντί να μείνουν σε ξενοδοχείο, να διανυκτερεύσουν σε σκηνή σε κάποιο κάμπινγκ με έξοδα διαμονής $\varphi_2(k)$=20 € ανά άτομο (τα ίδια και για τις δύο πόλεις).

Οι φίλοι επίσης ενημερώνονται για την πρόβλεψη του καιρού για το διήμερο των διακοπών τους. Σύμφωνα με την πρόβλεψη, η πιθανότητα βροχής σε κάθε μία από τις πόλεις που θα διανυκτερεύσουν φαίνεται στον Πίνακα 2. Εάν οι φίλοι αποφασίσουν να διανυκτερεύσουν σε κάποια πόλη σε σκηνή και τελικά βρέξει (z=0) τότε θα αναγκαστούν να πάνε σε ξενοδοχείο και επομένως θα πληρώσουν και τα έξοδα του κάμπινγκ αλλά και το ξενοδοχείο της αντίστοιχης πόλης, ενώ εάν δεν βρέξει (z=1) θα πληρώσουν μόνο τα έξοδα του κάμπινγκ. Έτσι το αναμενόμενο κόστος, εάν αποφασίσουν να μείνουν σε σκηνή, είναι $E\{\varphi_2(k)+\varphi_1(k)\cdot[1-z(k)]\}$.

Πίνακας 1: Τιμές ξενοδοχείων (ανά άτομο) ανά ημέρα.

k	0	1
φ_1(k) σε €	30	45

Πίνακας 2: Πιθανότητα βροχής ανά ημέρα.

k	0	1
p(z=0)	0.6	0.8

Θεωρώντας ότι η κατάσταση x(k) δείχνει τον αριθμό των διανυκτερεύσεων σε κάμπινγκ, εκκινώντας από x(0)=0 και θεωρώντας x(k+1)=x(k)+u(k)·z(k), προσδιορίστε τη βέλτιστη απόφαση διαμονής $u(k)\in\{0,1\}$ για κάθε πόλη (όπου u(k)=0 σημαίνει διαμονή σε ξενοδοχείο και u(k)=1 διαμονή σε σκηνή), προκειμένου να ελαχιστοποιηθεί το ολικό αναμενόμενο κόστος διαμονής ανά άτομο. Επιπλέον οι φίλοι επιθυμούν να μείνουν σε

κάμπινγκ ένα από τα βράδια των διακοπών τους, γι' αυτό το λόγο ορίζουν ως τελικό κόστος το $\theta\left[x(K)\right] = 10 \cdot \left|1 - x(K)\right|$.

(a) Διατυπώστε το παραπάνω πρόβλημα ως Πρόβλημα Στοχαστικού Βέλτιστου Ελέγχου με την αντικειμενική συνάρτηση, την καταστατική εξίσωση και όλους τους περιορισμούς.

(b) Επιλύστε αριθμητικά και γραφικά το πρόβλημα προσδιορίζοντας τη βέλτιστη πολιτική (κανόνα ελέγχου) διαμονής των φίλων σε κάθε πόλη, u(k). Ποιο είναι το ολικό αναμενόμενο κόστος διαμονής (ανά άτομο) των διακοπών τους;

8. ΕΦΑΡΜΟΓΕΣ ΣΤΟΧΑΣΤΙΚΟΥ ΔΥΝΑΜΙΚΟΥ ΠΡΟΓΡΑΜΜΑΤΙΣΜΟΥ

Οι εφαρμογές του Στοχαστικού Δυναμικού Προγραμματισμού μπορούν να ταξινομηθούν σε δύο γενικές κατηγορίες. Στην πρώτη κατηγορία ανήκουν εφαρμογές που μπορούν να θεωρηθούν ότι προέρχονται από τη γενίκευση αντίστοιχων αιτιοκρατικών προβλημάτων. Στην κατηγορία αυτή ανήκει για παράδειγμα το πρόβλημα του απρόσεκτου οδηγού (Παράδ. 7.1), το πρόβλημα αποθήκευσης (Παράδ. 7.3, 7.7) που θεωρεί στοχαστική ζήτηση αντί της αιτιοκρατικής του Κεφαλαίου 5.1, καθώς και τα στοχαστικά παίγνια (Παράδ. 7.8) θεωρώντας πιθανότητες μελλοντικών κινήσεων, π.χ. λόγω χρήσης ζαριών, όπως γίνεται στο τάβλι. Κατά παρόμοιο τρόπο μπορούν να γενικευθούν το πρόβλημα ελέγχου δεξαμενών ύδατος (Παράδ. 6.2) θεωρώντας στοχαστικές (αντί αιτιοκρατικών) μελλοντικές εισροές, ο γραμμικός-τετραγωνικός έλεγχος, καθώς και πολλά άλλα αιτιοκρατικά προβλήματα. Στην δεύτερη κατηγορία ανήκουν στοχαστικά προβλήματα χωρίς αιτιοκρατική αντιστοιχία. Επί παραδείγματι, το πρόβλημα προθεσμίας (Παράδ. 7.4, 7.6) έχει νόημα μόνο στην στοχαστική περίπτωση, διότι, αν οι μελλοντικές προσφορές είναι γνωστές, η λύση είναι προφανής. Σημειώνουμε όμως και το γεγονός ότι μερικά καταρχήν στοχαστικά προβλήματα μπορούν να μετασχηματισθούν και τελικά να επιλυθούν με αιτιοκρατικές μεθόδους, όπως, για παράδειγμα, το πρόβλημα αξιοπιστίας του Κεφαλαίου 5.3α).

8.1 Η Έννοια του Ρίσκου

Η ιδιομορφία στοχαστικών προβλημάτων βελτιστοποίησης έγκειται κυρίως σε δύο στοιχεία που δεν εμφανίζονται σε αιτιοκρατικά προβλήματα, και συγκεκριμένα στην ιδιαίτερη σημασία της **συλλογής πληροφορίας πραγματικού χρόνου** (ανάδραση), όπως είχαμε την ευκαιρία να διαπιστώσουμε στο Παράδ. 7.1, και στην παρουσία **αβεβαιότητας** (ρίσκου) κατά την επιλογή του κριτηρίου κόστους.

Για να κατανοήσουμε καλύτερα την έννοια του ρίσκου, ας θεωρήσουμε το πρόβλημα επένδυσης ενός κεφαλαίου κ σε δύο διαφορετικές επενδύσεις Α και Β. Ας υποθέσουμε ότι η Α δίδει με βεβαιότητα 1.49 € για κάθε επενδυθέν ευρώ ενώ η Β δίδει με βεβαιότητα 1.50 € για κάθε επενδυθέν ευρώ. Αν έχουμε σαν στόχο την μεγιστοποίηση του κέρδους, προφανώς θα επενδύσουμε όλο το κ στην Β. Ας υποθέσουμε όμως τώρα ότι η Β προσφέρει 1.50 € **κατά μέσο όρο**, αλλά όχι με βεβαιότητα. Για παράδειγμα, ας υποθέσουμε ότι υπάρχει 80% πιθανότητα να χάσουμε το επενδυθέν ποσό και 20% πιθανότητα απόδοσης 7.50 € για κάθε επενδυθέν ευρώ. Τότε, αν το αντικειμενικό κριτήριο είναι η μεγιστοποίηση του **αναμενόμενου** κέρδους, θα πρέπει και πάλι όλο το κ να επενδυθεί στην Β, και το ίδιο ισχύει προφανώς για κάθε πιθανολογική κατανομή της απόδοσης της Β που έχει αναμενόμενη τιμή 1.50 €. Ποιός όμως θα ήταν διατεθειμένος να επενδύσει, π.χ., όλη την περιουσία του κ σε μια επένδυση όπου θα χάσει τα πάντα με πιθανότητα 80%; Πολλοί θα προτιμούσαν να επενδύσουν αποκλειστικά στην σίγουρη επένδυση Α, ενώ άλλοι θα ήταν διατεθειμένοι να πάρουν κάποιο ρίσκο επενδύοντας ένα μικρότερο ή μεγαλύτερο μέρος του κ και στην επένδυση Β. Το παράδειγμα αυτό δείχνει, ότι μια διατύπωση του προβλήματος που επιχειρεί μεγιστοποίηση του αναμενόμενου κέρδους μπορεί να είναι ατυχής, ακριβώς επειδή το ρίσκο κάθε επένδυσης δεν βρίσκει έκφραση κατά ικανοποιητικό τρόπο μέσα από το αντικειμενικό κριτήριο και άρα δεν επηρεάζει κατάλληλα την βέλτιστη απόφαση. Αυτό όμως γεννά το ερώτημα, με ποιο τρόπο θα μπορούσε να διατυπωθεί κατάλληλα το πρόβλημα.

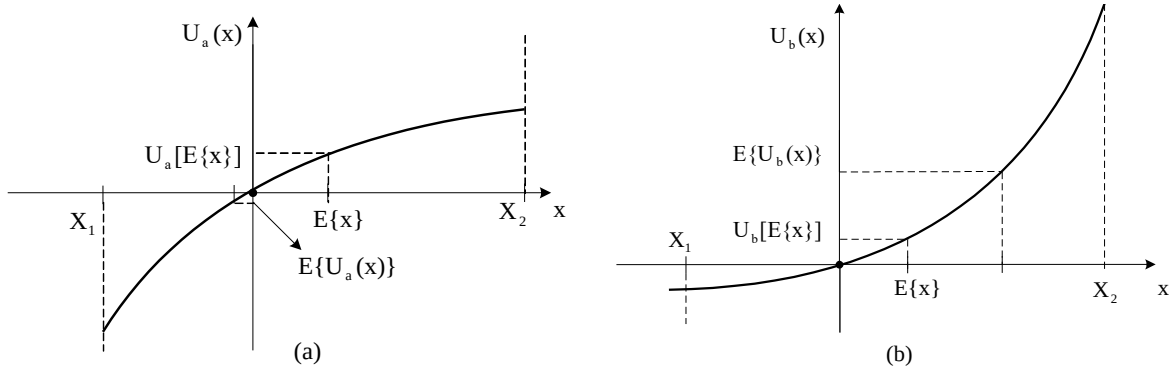
Μια προσέγγιση, που φέρει το όνομα **min-max** ή **max-min**, επιχειρεί μεγιστοποίηση του κέρδους βάσει της χειρότερης δυνατής απόδοσης κάθε επιλογής (βλ. π.χ. το Πρόβλημα 7.2). Στο προαναφερθέν παράδειγμα, η προσέγγιση min-max θα επιλέξει πάντοτε αποκλειστικά την σίγουρη επένδυση Α, πράγμα που θα μπορούσε να θεωρηθεί αποδεκτό. Η προσέγγιση min-max θα επιλέξει όμως αποκλειστικά την Α ακόμη και με σίγουρη απόδοση 0.01 € ανά επενδυθέν ευρώ, εφόσον υπάρχει έστω και ελάχιστη πιθανότητα για την Β να έχει μηδενική απόδοση, πράγμα που δείχνει ότι η προσέγγιση αυτή είναι άκρως απαισιόδοξη και τελικά, για πολλά προβλήματα, ακατάλληλη.

Είναι σαφές ότι διάφοροι επενδυτές (ή και ο ίδιος επενδυτής σε διαφορετικές καταστάσεις) θα επιλέξουν διάφορους καταμερισμούς επένδυσης μεταξύ Α και Β. Αυτές οι επιλογές μπορούν εμμέσως να εκφρασθούν μέσω μιας **συνάρτησης ωφέλειας** U(x), όπου x το κέρδος που απορρέει από τη μία ή την άλλη επιλογή. Το πρόβλημα στοχαστικής βελτιστοποίησης μπορεί τότε να επιχειρήσει την μεγιστοποίηση της αναμενόμενης τιμής της συνάρτησης ωφέλειας, δηλαδή $E\{U(x)\}\rightarrow\max$ αντί του $E\{x\}\rightarrow\max$. Είναι βέβαια προφανές ότι ως συνάρτηση ωφέλειας U(x) επιλέγεται μια μονοτονικά αύξουσα συνάρτηση.

Θεωρούμε έναν λήπτη αποφάσεων (π.χ. επενδυτή) με συνάρτηση ωφέλειας U(x), $x\in X$. Ο λήπτης αποφάσεων, και κατά συνέπεια η συνάρτηση ωφέλειας, λέγονται **εχθρικοί στο ρίσκο** αν

$$\underset{P}{E}\{U(x)\} \le U[\underset{P}{E}\{x\}] \qquad (8.1)$$

για κάθε πιθανολογική κατανομή P στο X με πεπερασμένες αναμενόμενες τιμές. Μπορεί να αποδειχθεί ότι η εχθρότητα στο ρίσκο είναι μαθηματικά ισοδύναμη με το κοίλον της συνάρτησης ωφέλειας. Αντιθέτως, ένας λήπτης αποφάσεων λέγεται **φιλικός προς το ρίσκο** αν ισχύει η αντίθετη ανισότητα στην (8.1), πράγμα που συμβαίνει αν η συνάρτηση ωφέλειας είναι κυρτή. Τέλος, ένας λήπτης αποφάσεων λέγεται **ουδέτερος στο ρίσκο** αν έχει συνάρτηση ωφέλειας γραμμική, οπότε ισχύει στην (8.1) η ισότητα. Το Σχ. 8.1. δίδει παραδείγματα συναρτήσεων ωφέλειας εχθρικής (a) και φιλικής (b) προς το ρίσκο. Γίνεται εμφανές ότι μια κοίλη συνάρτηση ωφέλειας (Σχ. 8.1a) αποδίδει μικρότερο βάρος σε ενδεχόμενα κέρδη (δηλαδή x>0) παρά σε αντίστοιχες ζημιές (x<0) και είναι, για το λόγο αυτό, εχθρική προς το ρίσκο ενδεχόμενων ζημιών, ενώ ισχύει ακριβώς το αντίθετο για κυρτές συναρτήσεις ωφέλειας (Σχ. 8.1b). Πράγματι, υποθέτοντας, π.χ., ότι το κέρδος x είναι ισοκατανεμημένο στο διάστημα $[X_1, X_2]$, η E{x} βρίσκεται στο μεσοδιάστημα των X_1 και X_2, και άρα η συνάρτηση ωφέλειας $U_a(x)$ του Σχ. 8.1a επιβεβαιώνει την (8.1) (εχθρική προς το ρίσκο) ενώ η $U_b(x)$ του Σχ. 8.1b επιβεβαιώνει την αντίθετη ανισότητα στην (8.1) (φιλική προς το ρίσκο). Έτσι, αν επί παραδείγματι η ζητούμενη απόφαση είναι επένδυση ή μη επένδυση (με ισοκατανεμημένη πιθανολογική κατανομή του κέρδους x), τότε η $U_a(x)$ οδηγεί στην απόφαση μη επένδυσης, αφού $E\{U_a(x)\}<0$, ενώ η $U_b(x)$ σε απόφαση επένδυσης, αφού $E\{U_b(x)\}>0$.



Σχήμα 8.1.: Παραδείγματα συναρτήσεων ωφέλειας εχθρικής (a) και φιλικής (b) προς το ρίσκο.

Για παράδειγμα, ένας παίκτης ρουλέτας ή λόττο (που δεν θεωρεί στοιχείο ωφέλειας το παιγνίδι καθεαυτό) είναι τυπικό παράδειγμα φιλικής στο ρίσκο πολιτικής, αφού, έχοντας τις επιλογές να παίξει ή να μην παίξει, επιλέγει να παίξει, παρότι $E\{x\} < 0$ (γιατί;), όπου x το κέρδος του παίκτη σε κάθε γύρισμα της ρουλέτας, παρότι δηλαδή κατά την αναμενόμενη τιμή του κέρδους δεν έχει παρά να χάσει.

Παράδειγμα 8.1: Έχετε την ευκαιρία, αν το αποφασίσετε, να παίξετε σε ένα παιγνίδι, όπου με πιθανότητα 0.9 θα χάσετε 100 € και με πιθανότητα 0.1 θα κερδίσετε 800 €. Ένας φίλος σας (και φίλος του ρίσκου) έχει μόνον 50 € διαθέσιμα και σας προτείνει να παίξετε συνεταιρικά. Για να σας δελεάσει, προτείνει να καταβάλει 50 € σε περίπτωση που χάσετε και να εισπράξει μόνο 300 € σε περίπτωση που κερδίσετε. Έχετε έτσι τρεις επιλογές: Να μην παίξετε καθόλου ή να παίξετε μόνος σας ή να παίξετε συνεταιρικά με τον φίλο σας. Τι θα "αποφάσιζε" ένα πρόβλημα στοχαστικής βελτιστοποίησης; Η απάντηση εξαρτάται από την συνάρτηση ωφέλειας που θα χρησιμοποιηθεί. Θα θεωρήσουμε τρεις επιλογές: Μια γραμμική συνάρτηση $J_1 = E\{x\}$, μια κοίλη συνάρτηση $J_2 = E\{\ln(200+x)\}$ και μια κυρτή συνάρτηση $J_3 = E\{e^{x/100}\}$. Αν δεν παίξετε, έχουμε με βεβαιότητα $x = 0$ και άρα $J_1 = 0$, $J_2 = \ln 200 = 5.3$, $J_3 = e^0 = 1$. Αν παίξετε μόνος, έχουμε

$$E\{x\} = -100 \cdot 0.9 + 800 \cdot 0.1 = -10$$

$$E\{\ln(200+x)\} = 0.9 \ln 100 + 0.1 \ln 1000 = 4.84$$

$$E\{e^{x/100}\} = 0.9\, e^{-1} + 0.1\, e^{8} = 298.$$

Αν παίξετε συνεταιρικά, έχουμε

$$E\{x\} = -50 \cdot 0.9 + 500 \cdot 0.1 = 5$$

$$E\{\ln(200+x)\} = 0.9 \ln 150 + 0.1 \ln 700 = 5.16$$

$$E\{e^{x/100}\} = 0.9 e^{-1/2} + 0.1\, e^{5} = 15.$$

Συμπερασματικά, η κοίλη συνάρτηση J_2 (εχθρική προς το ρίσκο) μεγιστοποιείται αν δεν παίξετε καθόλου, ενώ η γραμμική συνάρτηση J_1 (ουδέτερη στο ρίσκο) μεγιστοποιείται αν παίξετε συνεταιρικά, αφού μόνο τότε έχετε θετικό αναμενόμενο κέρδος. Η κυρτή συνάρτηση J_3 (φιλική προς το ρίσκο) μεγιστοποιείται αν παίξετε μόνος σας, οπότε έχετε μεν αρνητικό αναμενόμενο κέρδος ($J_1 = -10$) αλλά συγχρόνως έχετε την ευκαιρία (αν είστε τυχερός) του μέγιστου εν δυνάμει κέρδους.

■

Η έννοια του ρίσκου, εκφραζόμενη μέσω της συνάρτησης ωφέλειας, είναι λοιπόν σημαντική, αφού εμπεριέχει και χαρακτηρίζει (μοντελοποιεί) την συμπεριφορά του λήπτη αποφάσεων. Ως τοπικό (περί το x) μέτρον εχθρότητας στο ρίσκο θεωρείται το μέγεθος

$$r(x) = -U''(x)/U'(x) \qquad (8.2)$$

(όπου U', U'' είναι η πρώτη και δεύτερη παράγωγος αντίστοιχα και $U'(x) \neq 0$) που φέρει το όνομα **δείκτης απόλυτης εχθρότητας στο ρίσκο**. Επειδή μια συνάρτηση ωφέλειας $U(x)$ είναι εκ φύσεως μονοτονικά αύξουσα, έχουμε $U'(x) > 0$. Αν η $U(x)$ είναι κοίλη ή κυρτή, τότε $U''(x) < 0$ ή $U''(x) > 0$, και άρα $r(x) > 0$ ή $r(x) < 0$, αντίστοιχα. Αν η $U(x)$ είναι γραμμική, τότε έχουμε $U''(x) = r(x) = 0$. Ο δείκτης $r(x)$ παίζει συχνά ένα σημαντικό ρόλο στην ανάλυση της συμπεριφοράς των

ληπτών αποφάσεων, αποτελεί δε γενική παραδοχή ότι, για τους περισσότερους λήπτες αποφάσεων, ο r(x) είναι μια φθίνουσα ή τουλάχιστον μη αύξουσα συνάρτηση του x, ότι δηλαδή ο λήπτης αποφάσεων αποδέχεται ευκολότερα το ρίσκο όσο αυξάνει ο πλούτος του.

8.2 Το Πρόβλημα του Εισοδηματία

Θεωρούμε το πρόβλημα του εισοδηματία των Παραδ. Α.2, Β.2, 3.8, γενικευμένο κατά την υπόθεση ότι η απόδοση α του κεφαλαίου είναι μια στοχαστική μεταβλητή α = z(k), που προφανώς εξαρτάται από μελλοντικές γενικές οικονομικές εξελίξεις, έτσι ώστε να έχουμε την καταστατική εξίσωση

$$x(k+1) = z(k)\, u(k). \tag{8.3}$$

Θεωρούμε ότι η πιθανολογική κατανομή της z(k) δίδεται μέσω της συνάρτησης πυκνότητας p(z), ανεξάρτητης του χρόνου, και ότι η αναμενόμενη τιμή E{z} είναι μεγαλύτερη του 1. Στόχος είναι η μεγιστοποίηση της αναμενόμενης τιμής της συνάρτησης απόλαυσης για δεδομένη περίοδο K διαστημάτων, π.χ.

$$J = E\{\sum_{k=0}^{K-1} \sqrt{x(k) - u(k)}\} \tag{8.4}$$

λαμβάνοντας υπόψη

$$0 \le u(k) \le x(k). \tag{8.5}$$

Η έλλειψη τελικού κόστους $\vartheta[x(K)]$ στην (8.4) υποδηλώνει ότι ο εισοδηματίας δεν ενδιαφέρεται να διατηρήσει κάποιο κεφάλαιο στο τέλος του ορίζοντα K.

Από την (7.8) έχουμε χρησιμοποιώντας την (7.23)

$$V[x(k),k] = \max\{\sqrt{x(k) - u(k)} + \int_{-\infty}^{\infty} V[z \cdot u(k), k+1]\, p(z)\, dz\}. \tag{8.6}$$

Για k=K–1 και V[x(K),K] = 0 (λόγω (7.10)) έχουμε από την (8.6)

$$u(K-1) = 0, \quad V[x(K-1), K-1] = \sqrt{x(K-1)}. \tag{8.7}$$

Το αποτέλεσμα αυτό υποδηλώνει ότι ο εισοδηματίας θα καταναλώσει (ελλείψει τελικού κόστους) όλο το διαθέσιμο κεφάλαιο x(K–1) κατά την τελευταία χρονική περίοδο K–1, πράγμα αναμενόμενο.

Για k= K–2 έχουμε τότε

$$V[x(K-2), K-2] = \max\{\sqrt{x(K-2) - u(K-2)} + \int_{-\infty}^{\infty} \sqrt{z \cdot u(K-2)}\, p(z) dz\}. \tag{8.8}$$

Ορίζοντας, βάσει της γνωστής συνάρτησης p(z)

$$\sqrt{\alpha} = \int_{-\infty}^{\infty} \sqrt{z}\, p(z)dz \tag{8.9}$$

έχουμε από την (8.8)

$$V[x(K-2),K-2] = \max\{\sqrt{x(K-2) - u(K-2)} + \sqrt{\alpha u(K-2)}\}. \tag{8.10}$$

Η μεγιστοποίηση ως προς u(K–2) υπό (8.5) δίδει

$$u(K-2) = \frac{\alpha}{1+\alpha} x(K-2)\,. \tag{8.11}$$

$$V[x(K-2), K-2] = \sqrt{(1+\alpha)\, x(K-2)}\,. \tag{8.12}$$

Δεν είναι δύσκολο, συνεχίζοντας την όπισθεν πορεία, να αποδείξουμε μέσω τέλειας επαγωγής ότι ισχύει η γενική λύση

$$u(k) = \frac{\alpha + \alpha^2 + \ldots + \alpha^{K-k-1}}{1 + \alpha + \ldots + \alpha^{K-k-1}} x(k). \tag{8.13}$$

$$V[x(k),k] = \sqrt{(1 + \alpha + \ldots + \alpha^{K-k-1})\, x(k)}\,. \tag{8.14}$$

Παρατηρούμε ότι η λύση αυτή είναι πανομοιότυπη με την λύση του αιτιοκρατικού προβλήματος (Παράδ. 3.8) με τον κατάλληλο όμως ορισμό (8.9). Έτσι, αν ο εισοδηματίας επιλύσει το αιτιοκρατικό πρόβλημα του Παραδ. 3.8 χρησιμοποιώντας σαν τιμή του α την μέση τιμή της z(k) (και όχι τον ορισμό (8.9)), θα καταλήξει σε μη βέλτιστη λύση του αντίστοιχου στοχαστικού προβλήματος.

8.3 Το Πρόβλημα Προγραμματισμού Επενδύσεων

Το πρόβλημα προγραμματισμού επενδύσεων αφορά την κατανομή υπάρχοντος κεφαλαίου σε διαθέσιμες επενδύσεις με ή χωρίς ρίσκο. Θα θεωρήσουμε καταρχήν την επίλυση του προβλήματος για ένα και μόνο χρονικό διάστημα και θα γενικεύσουμε στην συνέχεια τα αποτελέσματα για την καθεαυτό δυναμική περίπτωση K διαστημάτων.

Έστω x_0 το αρχικό διαθέσιμο κεφάλαιο και n αβέβαιες επενδύσεις με στοχαστικές αποδόσεις z_1, $z_2,\ldots,z_n$. Έστω επίσης μια σίγουρη επένδυση με απόδοση α. Σημειώνουμε με u_i, i=1,...,n, το επενδυτέο ποσό σε κάθε αβέβαιη επένδυση και έχουμε έτσι τελικό κεφάλαιο (μετά από ένα χρονικό διάστημα)

$$x_1 = \alpha(x_0 - u_1 - \ldots - u_n) + \sum_{i=1}^{n} z_i u_i = \alpha x_0 + \sum_{i=1}^{n} (z_i - \alpha) u_i. \tag{8.15}$$

Το κριτήριο κόστους είναι η μεγιστοποίηση του

$$J = E[\,\vartheta\,(x_1)] \tag{8.16}$$

όπου ϑ είναι μια κοίλη, διπλά παραγωγίσιμη συνάρτηση ωφέλειας του κεφαλαιούχου.

Κατά την επίλυση του προβλήματος θα θεωρήσουμε ότι η συνάρτηση ωφέλειας ικανοποιεί

$$-\vartheta'(x_1)/\vartheta''(x_1) = s+bx_1 \quad \forall x_1 \tag{8.17}$$

όπου ϑ' και ϑ'' είναι η πρώτη και δεύτερη παράγωγος της ϑ, αντίστοιχα, και όπου s, b είναι σταθερές. Η παραδοχή (8.17) ικανοποιείται από τις συναρτήσεις

$$\begin{aligned} \vartheta(x) &= -e^{-x/s} && \text{για } b=0 \\ \vartheta(x) &= \ln(x+s) && \text{για } b=1 \\ \vartheta(x) &= [1/(b-1)](s+bx)^{1-(1/b)} && \text{άλλως.} \end{aligned} \tag{8.18}$$

Η (8.17) υποδηλώνει ότι ο δείκτης απόλυτης εχθρότητας στο ρίσκο $r(x_1)$ (βλ. (8.2)) του προβλήματος είναι $(s+bx)^{-1}$ και άρα ο δείκτης αυτός είναι μονοτονικά φθίνων με το x για $b>0$ και σταθερός για $b=0$.

Έστω $u_i = R_i(x_0)$, $i=1,...,n$, η βέλτιστη πολιτική. Ονομάζουμε την πολιτική **μερικά χωρισμένη** αν έχει την μορφή

$$R_i(x_0) = \gamma_i\, r(x_0) \tag{8.19}$$

όπου γ_i, $i=1,...,n$, είναι σταθερές και $r(x_0)$ είναι μια συνάρτηση του x_0, πανομοιότυπη για κάθε x_0. Υπό συνθήκες μερικού χωρισμού έχουμε προφανώς

$$R_i(x_0)/R_j(x_0) = \gamma_i/\gamma_j,\ 1 \le i, j \le n,\ \ \gamma_j \ne 0.$$

Στις περιπτώσεις που εξετάζουμε εδώ, προκύπτει, όπως θα αποδείξουμε, η γραμμική πολιτική

$$R_i(x_0) = \gamma_i \cdot (s+b\alpha x_0) \tag{8.20}$$

όπου s, b είναι οι σταθερές της επιλεχθείσας συνάρτησης ωφέλειας ϑ κατά την (8.18). Η πολιτική (8.20) είναι προφανώς μερικά χωρισμένη. Στην ειδική περίπτωση $s=0$ λέμε ότι η βέλτιστη πολιτική είναι και **ολικά χωρισμένη** με την έννοια ότι το ποσοστό του αρχικού κεφαλαίου που επενδύεται σε κάθε αβέβαιη επένδυση είναι σταθερό και ανεξάρτητο του ύψους του αρχικού κεφαλαίου x_0.

Θα δείξουμε τώρα ότι, υπό την παραδοχή (8.17), η βέλτιστη πολιτική δίδεται όντως από την (8.20). Θα δείξουμε επίσης ότι, υπό την ίδια παραδοχή, ορίζοντας $V(x_0) = \max J$, έχουμε

$$-V'(x_0)/V''(x_0) = (s/\alpha) + bx_0 \quad \forall x_0. \tag{8.21}$$

Προς απόδειξη των (8.20), (8.21) υπό την παραδοχή (8.17), θα υποθέσουμε καταρχήν ότι η βέλτιστη πολιτική έχει τη μορφή

$$R_i(x_0) = \gamma_i(x_0) \cdot (s+b\alpha x_0)$$

όπου $\gamma_i(x_0)$, i=1,...,n, είναι κάποιες παραγωγίσιμες συναρτήσεις, και θα αποδείξουμε ότι $\gamma_i'(x_0) = 0$, δηλαδή ότι οι γ_i είναι σταθερές. Πράγματι έχουμε με τις (8.15), (8.16) για την βέλτιστη πολιτική

$$dE[\vartheta(x_1)]/du_i = E\{\vartheta'[\alpha x_0 + \sum_{j=1}^{n}(z_j-\alpha)\gamma_j(x_0)\cdot(s+b\alpha x_0)](z_i-\alpha)\} = 0 \qquad (8.22)$$

για i=1,...,n. Παραγωγίζοντας την (8.22) ως προς x_0 έχουμε

$$E\{(z_i-\alpha)\ \vartheta''(x_1)\ [\alpha + \alpha b\sum_{j=1}^{n}(z_j-\alpha)\gamma_j(x_0) + (s+b\alpha x_0)\sum_{j=1}^{n}(z_j-\alpha)\gamma_j'(x_0)]\} = 0$$

εξού η ισότητα

$$E\{\vartheta''(x_1)\ (s+b\alpha x_0)\sum_{j=1}^{n}(z_j-\alpha)\ (z_i-\alpha)\ \gamma_j'(x_0)\} = -\ E\{\vartheta''(x_1)\ (z_i-\alpha)\ \alpha\ [1+\sum_{j=1}^{n}(z_j-\alpha)\gamma_j(x_0)b]\}.$$

Λαμβάνοντας την εξίσωση αυτή n φορές, δηλαδή για i = 1,...,n, έχουμε

$$E\{\begin{bmatrix}(z_1-\alpha)^2 & \dots & (z_1-\alpha)(z_n-\alpha)\\ M & & M\\ (z_n-\alpha)(z_1-\alpha) & \dots & (z_n-\alpha)^2\end{bmatrix}\vartheta''(x_1)(s+b\alpha x_0)\}\begin{bmatrix}\gamma_1'(x_0)\\ M\\ \gamma_n'(x_0)\end{bmatrix} =$$

$$= -\begin{bmatrix}E\{\vartheta''(x_1)(z_1-\alpha)\,\alpha\,[1+\sum_{j=1}^{n}(z_j-\alpha)\gamma_j(x_0)b]\}\\ M\\ E\{\vartheta''(x_1)(z_n-\alpha)\,\alpha\,[1+\sum_{j=1}^{n}(z_j-\alpha)\gamma_j(x_0)b]\}\end{bmatrix}. \qquad (8.23)$$

Από την (8.17) έχουμε

$$\vartheta''(x_1) = -\ \frac{\vartheta'(x_1)}{s+b[\alpha x_0+\sum(z_j-\alpha)\gamma_j(x_0)(s+b\alpha x_0)]} =$$

$$= -\ \frac{\vartheta'(x_1)}{(s+b\alpha x_0)[1+\sum(z_j-\alpha)\gamma_j(x_0)b]}. \qquad (8.24)$$

Αντικαθιστώντας την (8.24) στην (8.23) και κάνοντας χρήση της (8.22), διαπιστώνουμε ότι η δεξιά πλευρά της (8.23) είναι μηδενικό διάνυσμα. Καθότι ο πίνακας της αριστερής πλευράς της (8.23) είναι, σε μη εκφυλισμένα προβλήματα, πλήρους βαθμού, συμπεραίνουμε ότι $\gamma_i'(x_0) = 0$, i=1,...,n, άρα αποδείξαμε την (8.20) (υπό την παραδοχή ότι η συνάρτηση ωφέλειας ικανοποιεί την (8.17)).

Στρεφόμαστε τώρα στην απόδειξη της (8.21). Έχουμε

$$V(x_0) = \max E\{\vartheta(x_1)\} = E\{\vartheta[\alpha[1+\sum_{i=1}^{n}(z_i-\alpha)\gamma_i b]x_0+\sum_{i=1}^{n}(z_i-\alpha)\gamma_i s]\}$$

και άρα

$$V'(x_0) = E\{\vartheta'(x_1)\alpha[1+\sum_{i=1}^{n}(z_i-\alpha)\gamma_i b]\} \tag{8.25}$$

$$V''(x_0) = E\{\vartheta''(x_1)\alpha^2[1+\sum_{i=1}^{n}(z_i-\alpha)\gamma_i b]^2\}.$$

Η τελευταία ισότητα, μετά από μερικές πράξεις και χρήση της (8.24), δίδει

$$V''(x_0) = -E\{\vartheta'(x_1)\,\alpha\,[1+\sum_{i=1}^{n}(z_i-\alpha)\gamma_i\, b]\}\alpha/(s+b\alpha x_0) \tag{8.26}$$

και συνδυάζοντας τις (8.25), (8.26) έχουμε την υπό απόδειξη (8.21).

Το πρόβλημα μπορεί τώρα εύκολα να γενικευθεί στην περίπτωση K διαδοχικών χρονικών διαστημάτων ορίζοντας:

x_k το κεφάλαιο στην αρχή του διαστήματος k

u_i^k το επενδυθέν ποσό στην επένδυση i για το διάστημα k

z_i^k την απόδοση της επένδυσης i για το διάστημα k

α_k την απόδοση της σίγουρης επένδυσης για το διάστημα k.

Έχουμε τότε την καταστατική εξίσωση (βλ. (8.15))

$$x_{k+1} = \alpha_k x_k+\sum_{i=1}^{n}(z_i^k-\alpha_k)u_i^k. \tag{8.27}$$

Θεωρούμε τις $\mathbf{z}^k$ ανεξάρτητες των $\mathbf{z}^{k-1}$, $\mathbf{z}^{k-2}$,... και με γνωστή πιθανολογική κατανομή. Ζητείται η μεγιστοποίηση του κριτηρίου

$$J = E\{\vartheta(x_K)\}$$

κάνοντας την παραδοχή ότι η ϑ ικανοποιεί την (8.17).

Εφαρμόζοντας Στοχαστικό Δυναμικό Προγραμματισμό έχουμε

$$V(x_K,K) = \vartheta(x_K) \tag{8.28}$$

$$V(x_k,k) = \max_{u_i^k} E\{V[\alpha_k x_k + \sum_{i=1}^{n}(z_i^k-\alpha_k)u_i^k, k+1]\}. \tag{8.29}$$

Από τη λύση (8.20) του μονοβαθμιαίου προβλήματος ξέρουμε ότι για το διάστημα K–1 ισχύει η βέλτιστη πολιτική (βλ. (8.20))

$$\mathbf{R}(x_{K-1}, K-1) = \boldsymbol{\gamma}^{K-1}(s + b\alpha_{K-1}\, x_{K-1}) \tag{8.30}$$

όπου $\boldsymbol{\gamma}^{K-1}$ είναι κάποιο (ζητούμενο) διάνυσμα. Επίσης ισχύει

$$-V'(x_{K-1},K-1)/V''(x_{K-1},K-1) = (s/\alpha_{K-1})+bx_{K-1}. \qquad (8.31)$$

Άρα η παραδοχή (8.17) ισχύει και για την μεγιστοποίηση (8.29) του διαστήματος K–2, οπότε έχουμε από τη λύση του μονοβαθμιαίου προβλήματος την βέλτιστη πολιτική

$$\mathbf{R}(x_{K-2},K-2) = \boldsymbol{\gamma}^{K-2} \cdot [(s/\alpha_{K-1})+b\alpha_{K-2}x_{K-2}]$$

και κατά παρόμοιο τρόπο

$$\mathbf{R}(x_k,k) = \boldsymbol{\gamma}^{k} \cdot [(s/\alpha_{k-1}...\alpha_{K-1}) + b\alpha_k x_k] \qquad (8.32)$$

όπου τα διανύσματα $\boldsymbol{\gamma}^{k}$, k=0,..., K–1, εξαρτώνται από τις πιθανολογικές κατανομές των z_i^k και μπορούν να υπολογισθούν από την μεγιστοποίηση στον αναδρομικό τύπο (8.29). Οι συναρτήσεις V ικανοποιούν την συνθήκη

$$-V'(x_k,k)/V''(x_k,k) = (s/\alpha_{K-1}...\alpha_k) + bx_k. \qquad (8.33)$$

Συγκρίνοντας την πολυβαθμιαία βέλτιστη πολιτική (8.32) (με την ιδιότητα (8.33) του βέλτιστου αντικειμενικού κριτηρίου) με την μονοβαθμιαία πολιτική (8.20) συμπεραίνουμε τα ακόλουθα:

- Αν s=0 στην (8.18), δηλαδή $\vartheta(x) = \ln x$ για b=1 ή $\vartheta(x) = [1/(b-1)](bx)^{1-(1/b)}$ για $b \neq 0$ και $b\neq 1$, τότε η βέλτιστη πολιτική για κάθε διάστημα του πολυβαθμιαίου προβλήματος είναι πανομοιότυπη με την μονοβαθμιαία βέλτιστη πολιτική που επιχειρεί σε κάθε διάστημα k (ανεξάρτητα από τα μελλοντικά διαστήματα) να μεγιστοποιήσει $E\{\vartheta(x_{k+1})\}$. Όταν (όπως στην περίπτωση αυτή) η δυνατότητα επανεπένδυσης σε μελλοντικά διαστήματα δεν επηρεάζει την βέλτιστη πολιτική ενός χρονικού διαστήματος, αναφερόμαστε σε **μυωπική** (βέλτιστη) **πολιτική.**
- Μυωπική πολιτική έχουμε επίσης αν $\alpha_k=1$, k=0,...,K–1, (π.χ. αν ο πληθωρισμός είναι ίσος με τον τραπεζικό τόκο).
- Μπορεί να αποδειχθεί ότι μυωπική πολιτική έχουμε στην περίπτωση s=0 ακόμη και αν α_k, k=0,...,K–1, είναι ανεξάρτητες τυχαίες μεταβλητές.
- Στην περίπτωση $s \neq 0$, $\alpha_k > 1$, $K \to \infty$, έχουμε επίσης μη μεταβλητή μυωπική πολιτική (γιατί;).

Παράδειγμα 8.2: Θεωρούμε ένα πρόβλημα προγραμματισμού επενδύσεων με ποσοστό σίγουρης επένδυσης α. Θεωρούμε μια αβέβαιη επένδυση με πιθανότητα p ποσοστού απόδοσης A, όπου $0 \leq A < \alpha$, και πιθανότητα 1–p ποσοστού απόδοσης B, όπου $B > \alpha$. Σαν συνάρτηση ωφέλειας έχουμε $\vartheta(x_K) = \ln x_K$ (δηλαδή s=0 και b=1) και ζητούμε μεγιστοποίηση του $J=E[\vartheta(x_K)]$. Η καταστατική εξίσωση (8.27) δίδει

$$x_{k+1} = \alpha x_k + (z_k-\alpha)u_k. \qquad (8.34)$$

Από την (8.28) έχουμε $V(x_K,K) = \ln x_K$ και από τις (8.29), (8.34) για k=K–1

$$\begin{aligned} V(x_{K-1},K-1) &= \max_{u_{K-1}} E\{\ln[\alpha x_{K-1} + (z_{K-1} - \alpha)u_{K-1}]\} \\ &= \max_{u_{K-1}} \{p \cdot \ln[\alpha x_{K-1} + (A-\alpha)u_{K-1}] + (1-p) \cdot \ln[\alpha x_{K-1} + (B-\alpha)u_{K-1}]\}. \end{aligned}$$

Για μεγιστοποίηση ως προς u_{K-1} θέτουμε την παράγωγο του παραπάνω όρου ίση με το μηδέν και έχουμε (παραλείποντας χάριν συντομίας τους δείκτες K–1)

$$p\frac{A-\alpha}{\alpha x+(A-\alpha)u}+(1-p)\frac{B-\alpha}{\alpha x+(B-\alpha)u}=0$$

εξού ο βέλτιστος κανόνας ελέγχου

$$u_{K-1}=\frac{(1-p)B+pA-\alpha}{(\alpha-A)(B-\alpha)}\alpha x_{K-1}. \qquad (8.35)$$

Επειδή η δεύτερη παράγωγος του ως άνω όρου είναι πάντα αρνητική, η (8.35) αντιστοιχεί σε απόλυτο μέγιστο. Αντικαθιστώντας στον αναδρομικό τύπο έχουμε

$$V(x_{K-1},K-1) = \ln(x_{K-1}) + \ln\alpha + \ln(B-A) + p\cdot\ln[p/(B-\alpha)] + (1-p)\cdot\ln[(1-p)/(\alpha-A)]. \quad (8.36)$$

Παρατηρούμε τα εξής:

- Η συνάρτηση $V(x_{K-1},K-1)$ παρουσιάζει στην (8.36) ένα μόνο μη σταθερό όρο (τον πρώτο) που οδηγεί για k=K–2 (και άρα και για k=K–3,...,0) σε ένα πανομοιότυπο πρόβλημα βελτιστοποίησης όπως αυτό που μόλις λύσαμε. Άρα η γενική λύση του προβλήματος δίδεται από την γενίκευση των (8.35), (8.36)

$$u_k=\frac{(1-p)B+pA-\alpha}{(\alpha-A)(B-\alpha)}\alpha x_k \qquad (8.37)$$

$$V(x_k,k) = \ln x_k + \{\ln\alpha + \ln(B-A) + p\cdot\ln[p/(B-\alpha)] + (1-p)\cdot\ln[(1-p)/(\alpha-A)]\}(K-k). \quad (8.38)$$

- Προφανώς η βέλτιστη λύση είναι μυωπική (με την έννοια που δώσαμε παραπάνω), πράγμα που αναμένονταν, αφού s = 0.
- Προφανώς η βέλτιστη λύση έχει τη μορφή (8.32) με s=0, b=1 και

$$\gamma=\frac{(1-p)B+pA-\alpha}{(\alpha-A)(B-\alpha)}.$$

- Η βέλτιστη πολιτική (8.37) δίδει $u_k \geq 0$ αν και μόνον αν $(1-p)B + pA - \alpha \geq 0$. Έτσι, δεδομένου ότι η αναμενόμενη απόδοση της αβέβαιης επένδυσης είναι (1–p)B+pA, η λύση εμπεριέχει ποσοστά αβέβαιης επένδυσης (δηλαδή $u_k>0$), αν και μόνον αν η αναμενόμενη απόδοση της αβέβαιης επένδυσης είναι μεγαλύτερη από το βέβαιο ποσοστό απόδοσης α. Η βέλτιστη πολιτική είναι επομένως εχθρική προς το ρίσκο, και αυτό κατά το Κεφάλαιο 8.1 οφείλεται στο κοίλο της συνάρτησης ωφέλειας. Σε περίπτωση ενεργοποίησης του περιορισμού $u_k \geq 0$ (δηλαδή αν $(1-p)B + pA - \alpha \leq 0$), έχουμε βέλτιστη λύση $u_k = 0$ (δηλαδή επένδυση όλων των χρημάτων στην βέβαιη επένδυση), οπότε $V(x_k,k) = \ln(\alpha^{K-k}x_k) = \ln x_k + (K-k)\ln\alpha$.
- Αν η βέλτιστη πολιτική (8.37) δίδει $u_k \geq x_k$, δηλαδή αν ισχύει για τα δεδομένα του προβλήματος

$$\frac{(1-p)B+pA-\alpha}{(\alpha-A)(B-\alpha)}\alpha x_k \geq x_k \Leftrightarrow \frac{B-\alpha}{B-A} \geq p\frac{\alpha}{A}$$

(π.χ. για p αρκούντως μικρό), τότε ενεργοποιείται ο αντίστοιχος περιορισμός και έχουμε αντί των (8.37), (8.38) την βέλτιστη λύση

$$u_k = x_k \tag{8.39}$$

$$V(x_k,k) = \ln x_k + [p \ln A + (1-p) \ln B]\,(K-k) \tag{8.40}$$

δηλαδή αβέβαιη επένδυση του συνολικού κεφαλαίου.

- Μπορεί εύκολα να πιστοποιηθεί ότι ο βέλτιστος κανόνας ελέγχου ισχύει και στην περίπτωση χρονικά μεταβλητών α_k, p_k, A_k, B_k.

■

Παράδειγμα 8.3: Είναι ενδιαφέρον να διερευνηθεί η συμπεριφορά του προβλήματος προγραμματισμού επενδύσεων του Παραδ. 8.2 όταν η συνάρτηση ωφέλειας είναι κυρτή (δηλαδή φιλική στο ρίσκο), π.χ. $\vartheta(x_K) = x_K^2$, ή γραμμική (δηλαδή ουδέτερη στο ρίσκο), π.χ. $\vartheta(x_K) = x_K$.

Στην πρώτη περίπτωση έχουμε από την αναδρομική εξίσωση (7.9) για $k = K-1$ και με $V(x_K, K) = x_K^2$

$$V(x_{K-1}, K-1) = \max_{u_{K-1}} E\{[\alpha x_{K-1} + (z_{K-1} - \alpha)\, u_{K-1}]^2\} =$$

$$= \max_{u_{K-1}} \{p[\alpha x_{K-1} + (A-\alpha)\, u_{K-1}]^2 + (1-p)\,[\alpha x_{K-1} + (B-\alpha)\, u_{K-1}]^2\}. \tag{8.41}$$

Ο υπό μεγιστοποίηση όρος αντιστοιχεί σε τετραγωνική παραβολή ως προς u_{K-1} με ένα μοναδικό ελάχιστο, πράγμα που μπορεί να πιστοποιηθεί λαμβάνοντας την πρώτη και δεύτερη παράγωγο ως προς u_{K-1}. Άρα η μεγιστοποίηση του όρου επιτυγχάνεται σε μια συνοριακή τιμή, δηλαδή είτε $u_{K-1} = x_{K-1}$, είτε $u_{K-1} = 0$. Έτσι, από την (8.41) μπορεί εύκολα να υπολογισθεί ο βέλτιστος έλεγχος

$$u_{K-1} = \begin{cases} x_{K-1}, \\ 0 \end{cases} \quad V(x_{K-1}, K-1) = \begin{cases} [pA^2 + (1-p)B^2]x_{K-1}^2 & \text{αν} \quad pA^2 + (1-p)B^2 - \alpha^2 \ge 0 \\ \alpha^2 x_{K-1}^2 & \text{άλλως.} \end{cases}$$

Επειδή η $V(x_{K-1}, K-1)$ έχει την ίδια δομή με την $V(x_K, K)$, το πρόβλημα βελτιστοποίησης και τα αποτελέσματα της βαθμίδας $K-2$, αλλά και όλων των προηγούμενων βαθμίδων, είναι παρόμοια, έχουμε δηλαδή

$$u_k = \begin{cases} x_k \\ 0 \end{cases}, \quad V(x_k, k) = \begin{cases} [pA^2 + (1-p)B^2]^{(K-k)}\, x_k^2 & \text{αν} \quad pA^2 + (1-p)B^2 - \alpha^2 \ge 0 \\ \alpha^{2(K-k)} x_k^2 & \text{άλλως.} \end{cases}$$

Στην γραμμική εξάλλου περίπτωση $V(x_K, K) = \vartheta(x_K) = x_K$, έχουμε από την αναδρομική εξίσωση (7.9) για $k = K-1$

$$V(x_{K-1}, K-1) = \max_{u_{K-1}} E\{\alpha x_{K-1} + (z_{K-1} - \alpha)\, u_{K-1}\} = \max_{u_{K-1}} \{\alpha x_{K-1} + [pA + (1-p)\, B - \alpha]\, u_{K-1}]\}.$$

Λόγω της γραμμικότητας του υπό μεγιστοποίηση όρου, το ζητούμενο μέγιστο θα βρίσκεται σε ένα από τα δύο επιτρεπτά όρια, δηλαδή $u_{K-1} = x_K$ ή $u_{K-1} = 0$, ανάλογα με το πρόσημο του όρου $pA + (1-p)\, B - \alpha$. Σε κάθε περίπτωση, η προκύπτουσα συνάρτηση $V(x_{K-1}, K-1)$ είναι γραμμική και άρα η λύση μπορεί να γενικευθεί για προηγούμενες βαθμίδες. Έχουμε έτσι

$$u_k = \begin{cases} x_k \\ 0 \end{cases}, \quad V(x_k, k) = \begin{cases} [pA + (1-p)B]^{(K-k)}\, x_k & \text{αν} \quad pA + (1-p)B - \alpha \ge 0 \\ \alpha^{(K-k)} x_k & \text{άλλως.} \end{cases}$$

Συμπερασματικά παρατηρούμε τα εξής:

- Η κοίλη συνάρτηση ωφέλειας του Παραδ. 8.2 οδηγεί σε μη συνοριακή λύση (8.37) (δηλαδή σε συνδυασμό βέβαιης και αβέβαιης επένδυσης) εφ' όσον

  $$\frac{A}{\alpha} \le p\frac{B-A}{B-\alpha} \le 1.$$

  Αντιθέτως η κυρτή ή γραμμική συνάρτηση ωφέλειας οδηγούν σε αποκλειστικά βέβαιη ή αποκλειστικά αβέβαιη επένδυση, ανάλογα με τα δεδομένα του προβλήματος.

- Η γραμμική συνάρτηση ωφέλειας οδηγεί σε αποκλειστικά αβέβαιη επένδυση, ακριβώς όταν $pA + (1-p)B-\alpha \ge 0$, δηλαδή ακριβώς όταν η αναμενόμενη απόδοση της αβέβαιης επένδυσης είναι μεγαλύτερη από τη βέβαιη απόδοση α, ενώ το αντίστοιχο κατώφλι στην περίπτωση της κυρτής συνάρτησης είναι $pA^2 + (1-p)B^2 - \alpha^2 \ge 0$. Επειδή το δεύτερο κατώφλι είναι μικρότερο από το πρώτο (δηλαδή, ισχύει $pA^2 + (1-p)B^2 - \alpha^2 < pA + (1-p)B-\alpha$ εφ' όσον $0 \le p \le 1$, $0 < A < \alpha < B$), η κυρτή συνάρτηση (φιλική στο ρίσκο) μπορεί να οδηγήσει σε αποκλειστικά αβέβαιη επένδυση ακόμη και αν η αναμενόμενη απόδοση είναι αρνητική, πράγμα που διαπιστώθηκε και στο Παραδ. 8.1.

■

8.4 Προβλήματα Βέλτιστης Στάσης

Η μεταβλητή ελέγχου σε **προβλήματα βέλτιστης στάσης** παίρνει συνήθως τιμές από ένα πεπερασμένο διακριτό πεδίο, όπου μια τιμή υποδηλώνει το τέλος (στάση) της εξέλιξης του συστήματος. Έτσι, σε κάθε χρονικό διάστημα, ο ελεγκτής παρατηρεί την παρούσα κατάσταση του συστήματος και αποφασίζει να συνεχίσει την διεργασία (ενδεχομένως με κάποιο κόστος) ή να την σταματήσει. Ένα παράδειγμα αυτού του είδους των προβλημάτων είδαμε ήδη στο πρόβλημα προθεσμίας των Παραδ. 7.4, 7.6.

Ένα παρόμοιο πρόβλημα της ίδιας κατηγορίας είναι το **πρόβλημα αγοράς εντός προθεσμίας**. Στο πρόβλημα αυτό θεωρούμε ότι υπάρχει ανάγκη αγοράς, το αργότερο σε κάποια συγκεκριμένη μελλοντική χρονική στιγμή, μιας ποσότητας κάποιου είδους. Αν η τιμή z του είδους μεταβάλλεται χρονικά, τίθεται σε κάθε διάστημα k το ερώτημα, με δεδομένη την παρούσα τιμή του, αν θα πρέπει η αναγκαία ποσότητα να αγορασθεί τώρα ($u(k) = 1$), ή σε κάποιο μελλοντικό διάστημα ($u(k) = 0$), οπότε η τιμή μπορεί να είναι υψηλότερη ή χαμηλότερη. Βέβαια, η αγορά πρέπει να γίνει το αργότατο κατά το χρονικό διάστημα K.

Κάνουμε την παραδοχή ότι οι διαδοχικές τιμές z(k) είναι ανεξάρτητες, με μη μεταβλητή και δεδομένη πιθανολογική κατανομή P(z). Όπως στα Παραδ. 7.4, 7.6 ορίζουμε την καταστατική εξίσωση

$$x(k+1) = \begin{cases} 0 & \text{αν} \quad u(k)=1 \text{ ή } x(k)=0 \\ z(k) & \text{άλλως} \end{cases}$$

και τις συναρτήσεις κόστους (προς ελαχιστοποίηση)

$$\vartheta\,[x(K)] = x(K)$$

$$\varphi[x(k), u(k)] = x(k)\cdot u(k).$$

Εφαρμόζοντας Στοχαστικό Δυναμικό Προγραμματισμό έχουμε

$$V[x(K),K] = x(K)$$

$$V[x(k),k] = \begin{cases} \min\{x(k),\ E\{V[z(k),k+1]\}\} & \text{αν} \quad x(k) \neq 0 \\ 0 & \text{αν} \quad x(k) = 0 \end{cases}$$

εξού η βέλτιστη πολιτική:

Αγορά αν $x(k) < E\{V[z(k),k+1]\} = \alpha(k)$
Μη αγορά αν $x(k) > E\{V[z(k),k+1]\} = \alpha(k)$.

Έχουμε λοιπόν

$$V[x(k)] = \begin{cases} x(k) & \text{αν} \quad x(k) < \alpha(k) \\ \alpha(k) & \text{αν} \quad x(k) > \alpha(k) \end{cases}$$

και κατά παρόμοιο τρόπο όπως στο Παράδ. 7.6 εξάγεται ότι η σειρά των αριθμών α(0), α(1),..., α(K–1) υπολογίζεται από την εξίσωση διακριτού χρόνου

$$\alpha(k) = \alpha(k+1)\ [1-P[\alpha(k+1)]] + \int_0^{\alpha(k+1)} z\ dP(z) \tag{8.42}$$

με την ακόλουθη συνοριακή τιμή που απορρέει από τον ορισμό του α(k) για k = K–1

$$\alpha(K-1) = \int_0^{\infty} z\ dP(z) = E\{z\}. \tag{8.43}$$

Το πρόβλημα βέλτιστης στάσης μπορεί να επιλυθεί κατά παρόμοιο τρόπο και στην περίπτωση διακριτού πεδίου τιμών Z και διακριτής πιθανολογικής κατανομής όπως στο Παράδ. 7.6.

Παράδειγμα 8.4: Ενδιαφερόμαστε για την αγορά ενός μεταχειρισμένου αυτοκινήτου κάποιου συγκεκριμένου τύπου για να ξεκινήσουμε για διακοπές σε 5 ημέρες. Κάθε ημέρα k=0, 1, 2, 3, 4 αποφασίζουμε για αγορά ή μη, ανάλογα με την καλύτερη προσφορά z(k) της προηγούμενης ημέρας, όπου οι προσφορές κάθε ημέρας μπορούν να θεωρηθούν ανεξάρτητες μεταξύ τους. Βάσει αντίστοιχων παλαιότερων αγγελιών σε εφημερίδες, καθορίζουμε την ακόλουθη συνάρτηση πυκνότητας προσφορών

$$p(z) = \begin{cases} (z-7)/9 & \text{για} \quad 7 \le z \le 10 \\ (13-z)/9 & \text{για} \quad 10 \le z \le 13 \\ 0 & \text{άλλως.} \end{cases}$$

Ζητείται η βέλτιστη πολιτική αποδοχής ή απόρριψης προσφορών. Από την (8.43) έχουμε καταρχήν α(4)=E{z}=10. Από την ως άνω συνάρτηση πυκνότητας έχουμε επίσης την κατανομή

$$P(z) = \begin{cases} z^2/18 - 7z/9 + 49/18 & \text{για} \quad 7 \le z \le 10 \\ -z^2/18 + 13z/9 - 151/18 & \text{για} \quad 10 \le z \le 13 \\ 0 & \text{για} \quad z \le 7 \\ 1 & \text{για} \quad z \ge 13. \end{cases}$$

Με

$$\int_0^{\alpha(4)} z\,p(z)\,dz = \int_0^{10} z\,p(z)\,dz = \left(\frac{z^3}{27} - \frac{7z^2}{18}\right)\Bigg|_7^{10} = 4.5$$

και P(10) = 0.5, έχουμε από την (8.42) α(3) = 9.5 και κατά παρόμοιο τρόπο α(2) = 9.2, α(1) = 9, α(0) = 8.85. Η βέλτιστη πολιτική επομένως είναι για k=0,...,4:

Αγορά αν x(k) < α(k)
Μη αγορά αν x(k) > α(k).

Είναι προφανές ότι η σειρά α(k), k=0,1,2,3,4, είναι συνεχώς αύξουσα, πράγμα που ανταποκρίνεται στη λογική του προβλήματος (γιατί;).

■

8.5 Στοχαστικός Γραμμικός-Τετραγωνικός Έλεγχος

Στο κεφάλαιο αυτό θα εξετάσουμε μια σημαντική ειδική περίπτωση του ΠΣΒΕ όπου η καταστατική εξίσωση έχει την ακόλουθη γραμμική μορφή

$$\mathbf{x}(k+1) = \mathbf{A}(k)\mathbf{x}(k) + \mathbf{B}(k)\mathbf{u}(k) + \mathbf{z}(k), \quad k=0,...,K-1 \tag{8.44}$$

και το κριτήριο κόστους είναι τετραγωνικό

$$J = E\{\|\mathbf{x}(K)\|_{\mathbf{S}}^2 + \sum_{k=0}^{K-1}[\|\mathbf{x}(k)\|_{\mathbf{Q}(k)}^2 + \|\mathbf{u}(k)\|_{\mathbf{R}(k)}^2]\} \tag{8.45}$$

και όπου $\mathbf{x}(k) \in R^n$, $\mathbf{u}(k) \in R^m$ ενώ οι πίνακες βάρους ικανοποιούν $\mathbf{S} \ge \mathbf{0}$, $\mathbf{Q}(k) \ge \mathbf{0}$, $\mathbf{R}(k) > \mathbf{0}$. Οι διαταραχές $\mathbf{z}(k)$ είναι ανεξάρτητες των $\mathbf{x}(k)$, $\mathbf{u}(k)$, $\mathbf{z}(k-1)$, $\mathbf{z}(k-2)$,... και έχουν γνωστή πιθανολογική κατανομή. Έχουμε επίσης $E\{\mathbf{z}(k)\}=\mathbf{0}$ και πεπερασμένη τιμή για $E\{\mathbf{z}(k)^2\}$. Το πρόβλημα **στοχαστικής γραμμικής-τετραγωνικής βελτιστοποίησης** όπως ορίζεται μέσω των (8.44), (8.45) είναι μια φυσική επέκταση του αντίστοιχου αιτιοκρατικού προβλήματος του Κεφαλαίου 5.4β).

Με εφαρμογή των (7.9), (7.10) (Στοχαστικός Δυναμικός Προγραμματισμός) στο παρόν πρόβλημα έχουμε

$$V[\mathbf{x}(K),K] = \|\mathbf{x}(K)\|_{\mathbf{S}}^2 \tag{8.46}$$

$$V[\mathbf{x}(k),k] = \min_{\mathbf{u}(k)}\{E[\|\mathbf{x}(k)\|_{\mathbf{Q}(k)}^2 + \|\mathbf{u}(k)\|_{\mathbf{R}(k)}^2 + V[\mathbf{A}(k)\mathbf{x}(k) + \mathbf{B}(k)\mathbf{u}(k) + \mathbf{z}(k), k+1]]\}. \tag{8.47}$$

Για την αναλυτική επίλυση των εξισώσεων αυτών αρχίζουμε με την βαθμίδα K–1

$$V[\mathbf{x}(K-1), K-1] = \min_{\mathbf{u}(K-1)} \{E[\|\mathbf{x}(K-1)\|^2_{\mathbf{Q}(K-1)} + \|\mathbf{u}(K-1)\|^2_{\mathbf{R}(K-1)}$$
$$+\|\mathbf{A}(K-1)\mathbf{x}(K-1) + \mathbf{B}(K-1)\mathbf{u}(K-1) + \mathbf{z}(K-1)\|^2_{\mathbf{S}}]\}.$$

Μετά από μερικές πράξεις και κάνοντας χρήση της $E\{\mathbf{z}(K-1)\} = \mathbf{0}$ έχουμε

$$V[\mathbf{x}(K-1), K-1] = \|\mathbf{x}(K-1)\|^2_{\mathbf{Q}(K-1)} + \min_{\mathbf{u}(K-1)} [\mathbf{u}(K-1)^T\mathbf{R}(K-1)\mathbf{u}(K-1)$$
$$+\mathbf{u}(K-1)^T\mathbf{B}(K-1)^T\mathbf{SB}(K-1)\mathbf{u}(K-1)+\mathbf{x}(K-1)^T\mathbf{A}(K-1)^T\mathbf{SA}(K-1)\mathbf{x}(K-1)$$
$$+2\mathbf{x}(K-1)^T\mathbf{A}(K-1)^T\mathbf{SB}(K-1)\mathbf{u}(K-1)] + E[\mathbf{z}(K-1)^T\mathbf{Sz}(K-1)]. \quad (8.48)$$

Η ελαχιστοποίηση που περιλαμβάνεται στην εξίσωση αυτή δίδει, μετά από παραγώγιση ως προς $\mathbf{u}(K-1)$, τον ακόλουθο γραμμικό κανόνα ελέγχου

$$\mathbf{u}(K-1) = -[\mathbf{R}(K-1) + \mathbf{B}(K-1)^T\mathbf{SB}(K-1)]^{-1}\mathbf{B}(K-1)^T\mathbf{SA}(K-1)\mathbf{x}(K-1) \quad (8.49)$$

όπου ο αντιστρεφόμενος πίνακας είναι πλήρους βαθμού καθότι $\mathbf{R}(K-1) > \mathbf{0}$. Αντικαθιστώντας την (8.49) στην (8.48) έχουμε

$$V[\mathbf{x}(K-1),K-1] = \mathbf{x}(K-1)^T\mathbf{P}(K-1)\mathbf{x}(K-1) + \mathbf{E}\{\mathbf{z}(K-1)^T\mathbf{Sz}(K-1)\} \quad (8.50)$$

με

$$\mathbf{P}(K-1) = \mathbf{A}(K-1)^T\{\mathbf{S}-\mathbf{SB}(K-1)[\mathbf{B}(K-1)^T\mathbf{SB}(K-1) + \mathbf{R}(K-1)]^{-1}\mathbf{B}(K-1)^T\mathbf{S}\}\mathbf{A}(K-1) + \mathbf{Q}(K-1).$$

Ο πίνακας $\mathbf{P}(K-1)$ είναι προφανώς συμμετρικός. Ο ίδιος πίνακας είναι επίσης θετικά ημιορισμένος. Το τελευταίο γίνεται φανερό αν παρατηρήσουμε ότι η ακόλουθη σχέση ισχύει για κάθε $\mathbf{x}\in R^n$

$$\mathbf{x}^T\mathbf{P}(K-1)\mathbf{x} = \min_{\mathbf{u}}[\|\mathbf{x}\|^2_{\mathbf{Q}(K-1)} + \|\mathbf{u}\|^2_{\mathbf{R}(K-1)} + \|\mathbf{A}(K-1)\mathbf{x} + \mathbf{B}(K-1)\mathbf{u}\|^2_{\mathbf{S}}].$$

Επειδή όμως $\mathbf{Q}(K-1)$, $\mathbf{R}(K-1)$, $\mathbf{S}$ είναι θετικά ημιορισμένοι, ο παραπάνω όρος είναι μη αρνητικός και παραμένει βέβαια τέτοιος και μετά την ελαχιστοποίηση ως προς $\mathbf{u}$, εξού και συνεπάγεται $\mathbf{P}(K-1) \geq \mathbf{0}$.

Συμπερασματικά διαπιστώσαμε ότι ο κανόνας ελέγχου (8.49) είναι γραμμικός και ότι η $V[\mathbf{x}(K-1), K-1]$, που επεισέρχεται στην ελαχιστοποίηση της επόμενης βαθμίδας K–2, είναι μια θετικά ημιορισμένη, τετραγωνική συνάρτηση, ακριβώς όπως και η $V[\mathbf{x}(K),K]$ που συμπεριλήφθηκε στην ελαχιστοποίηση της βαθμίδας K–1. (Ο σταθερός όρος στην (8.50) δεν επηρεάζει την παρούσα επιχειρηματολογία.) Κατά συνέπεια θα έχουμε κατά την ελαχιστοποίηση της επόμενης βαθμίδας K–2 κατά παρόμοιο τρόπο ένα γραμμικό κανόνα ελέγχου και μια θετικά ημιορισμένη τετραγωνική συνάρτηση $V[\mathbf{x}(K-2), K-2]$ κ.ο.κ. Με τον τρόπο αυτό μπορεί να υπολογισθεί τελικά ο ακόλουθος βέλτιστος κανόνας ελέγχου

$$\mathbf{u}(k) = -\mathbf{L}(k)\mathbf{x}(k), \quad k = 0,...,K-1 \quad (8.51)$$

όπου **L**(k) είναι ο χρονικά μεταβλητός **πίνακας ελέγχου** που προσδιορίζεται μέσω της ακόλουθης εξίσωσης

$$\mathbf{L}(k) = [\mathbf{B}(k)^T \mathbf{P}(k+1)\, \mathbf{B}(k) + \mathbf{R}(k)]^{-1}\, \mathbf{B}(k)^T \mathbf{P}(k+1)\, \mathbf{A}(k) \qquad (8.52)$$

και όπου ο **P**(k) ικανοποιεί την ακόλουθη εξίσωση διαφορών

$$\mathbf{P}(k) = \mathbf{A}(k)^T \mathbf{P}(k+1)\, \mathbf{A}(k) + \mathbf{Q}(k) - \mathbf{L}(k)^T \mathbf{B}(k)^T \mathbf{P}(k+1)\, \mathbf{A}(k)$$
$$\mathbf{P}(K) = \mathbf{S}. \qquad (8.53)$$

Οι εξισώσεις (8.51) - (8.53) είναι πανομοιότυπες με τις αντίστοιχες (5.62) - (5.65) του αιτιοκρατικού προβλήματος. Παρατηρούμε επίσης ότι το πρόβλημα αιτιοκρατικού ΓΤ Ελέγχου του Κεφαλαίου 5.4β απορρέει από το παρόν στοχαστικό πρόβλημα αν αντικαταστήσουμε τις στοχαστικές διαταραχές **z**(k) με τις αναμενόμενες τιμές τους, δηλαδή μηδέν. Οι διαπιστώσεις αυτές οδηγούν στο ερώτημα, αν όλα τα ΠΣΒΕ μπορούν να αναχθούν σε αντίστοιχα αιτιοκρατικά προβλήματα αντικαθιστώντας τις στοχαστικές διαταραχές με τις αναμενόμενες τιμές τους. Η απάντηση στο ερώτημα αυτό είναι αρνητική. Ο αναφερθείς μετασχηματισμός ενός στοχαστικού σε ένα αιτιοκρατικό πρόβλημα βελτιστοποίησης, γνωστός σαν **Αρχή Ισοδυναμίας Βεβαιότητας**, οδηγεί στην ακριβή λύση του στοχαστικού προβλήματος μόνο σε ειδικές περιπτώσεις, όπως στο πρόβλημα Στοχαστικού Γραμμικού-Τετραγωνικού Ελέγχου. Στα περισσότερα στοχαστικά προβλήματα η Αρχή Ισοδυναμίας Βεβαιότητας δεν οδηγεί σε λύση του αντίστοιχου στοχαστικού προβλήματος (βλ. επί παραδείγματι το αιτιοκρατικό και το στοχαστικό πρόβλημα εισοδηματία στο Παράδ. 3.8 και Κεφάλαιο 8.2 αντίστοιχα), βλ. Κεφάλαιο 9.

Προβλήματα

8.1 Ένας επενδυτής με κεφάλαιο x_k έχει την ευκαιρία κατά τα χρονικά διαστήματα k = 0, 1,..., K–1 είτε να διατηρήσει το κεφάλαιό του ως έχει (σίγουρη επένδυση), είτε να επενδύσει ένα μέρος του κεφαλαίου με πιθανότητες:

- p_1 να έχει ποσοστό απόδοσης A_1
- p_2 να έχει ποσοστό απόδοσης A_2
- $1-p_1-p_2$ να διατηρήσει ως έχει το επενδυθέν τμήμα του κεφαλαίου.

Έχουμε συνάρτηση ωφέλειας $\vartheta(x_K) = \ln x_K$ και ζητούμε τη βέλτιστη πολιτική επένδυσης $u_k = R(x_k)$, k = 0,..., K–1, που μεγιστοποιεί το κριτήριο $J = E[\vartheta(x_K)]$ για $p_1=3/11$, $p_2 = 4/11$, $A_1 = 0$, $A_2 = 11$. Ποιό είναι το βέλτιστο αναμενόμενο κριτήριο ωφέλειας $V(x_k, k)$;

(**Υπόδειξη**: Αρχίστε την επίλυση με τη βαθμίδα K–1 και γενικεύστε κατάλληλα τα αποτελέσματα για κάθε k με την αντίστοιχη αιτιολόγηση).

8.2 Σ' ένα καζίνο παίζεται το εξής ηλεκτρονικό παιγνίδι για έναν μόνο παίκτη. Σε κάθε προσπάθειά του ο παίκτης πατά ένα κουμπί και κερδίζει ένα ποσό x(k+1) = z(k), όπου z(k) $\in$ {0,1,2,3,4} είναι μια τυχαία μεταβλητή με γνωστή εκ των προτέρων πιθανολογική κατανομή, και όπου k ο αριθμός της προσπάθειας του παίκτου. Ο μέγιστος αριθμός προσπαθειών που επιτρέπεται να κάνει ο παίκτης είναι K=3. Στο τέλος της κάθε προσπάθειας ο παίκτης πρέπει να αποφασίσει αν θα σταματήσει να παίζει, αποδεχόμενος έτσι το ποσό που κερδίζει εκείνη

τη στιγμή, ή αν θα παραιτηθεί απ' αυτό το ποσό και συνεχίσει να παίζει προσδοκώντας κάποιο μεγαλύτερο κέρδος στις επόμενες προσπάθειες. Μετά την τελευταία του προσπάθεια ο παίκτης υποχρεούται να αποδεχθεί το αντίστοιχο αποτέλεσμα z(K–1). Η κατανομή του z(k) είναι διαφορετική για κάθε προσπάθεια και έχει ως εξής:

- Πρώτη προσπάθεια (k=0):
  $p[z(0) = 0] = \frac{10}{100}$, $p[z(0)=1] = \frac{20}{100}$, $p[z(0)=2] = \frac{10}{100}$, $p[z(0)=3] = \frac{30}{100}$, $p[z(0) = 4] = \frac{30}{100}$.
- Δεύτερη προσπάθεια (k=1):
  $p[z(1) = 0] = \frac{10}{100}$, $p[z(1)=1] = \frac{25}{100}$, $p[z(1)=2] = \frac{20}{100}$, $p[z(1)=3] = \frac{25}{100}$, $p[z(1) = 4] = \frac{20}{100}$.
- Τρίτη προσπάθεια (k=2):
  $p[z(2) = 0] = \frac{20}{100}$, $p[z(2)=1] = \frac{25}{100}$, $p[z(2)=2] = \frac{15}{100}$, $p[z(2)=3] = \frac{20}{100}$, $p[z(2) = 4] = \frac{20}{100}$.

Ο παίκτης προσπαθεί να μεγιστοποιήσει το αναμενόμενο κέρδος του από το παιγνίδι.

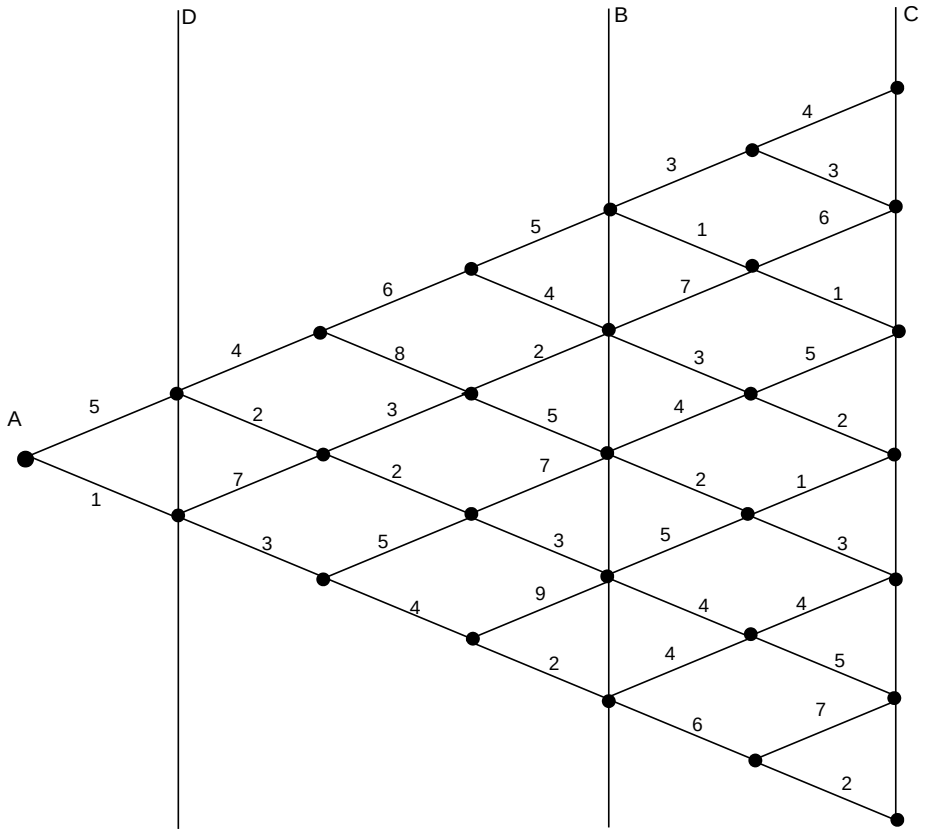
(a) Μοντελοποιήστε το πρόβλημα αυτό σαν ένα πρόβλημα στοχαστικού βέλτιστου ελέγχου αναγράφοντας:
   (i) Το κριτήριο κέρδους
   (ii) Την καταστατική εξίσωση
   (iii) Το πεδίο τιμών της μεταβλητής κατάστασης
   (iv) Τους περιορισμούς ελέγχου.

(b) Προσδιορίστε τη βέλτιστη πολιτική αποφάσεων για όλη τη διάρκεια κάθε παιγνιδιού.

8.3 Θεωρούμε το δίκτυο οδών που απεικονίζεται στην επόμενη σελίδα στο οποίο όλες οι μεταβάσεις νοούνται από αριστερά προς τα δεξιά και τα αναγραφόμενα κόστη μεταβάσεων αντιπροσωπεύουν τα μήκη των αντίστοιχων διαδρομών. Όταν ξεκινάμε από τον κόμβο A και μέχρι την γραμμή D δεν γνωρίζουμε με βεβαιότητα αν η κίνηση μας θα τερματίσει όταν φτάσουμε στη γραμμή B ή όταν φτάσουμε στη γραμμή C. Η πιθανότητα να σταματήσουμε στη γραμμή B είναι p ενώ η πιθανότητα να σταματήσουμε στη γραμμή C είναι 1−p. Θεωρούμε ότι όταν φτάσουμε στη γραμμή D, μαθαίνουμε σε ποια γραμμή (την B ή την C) θα σταματήσουμε την κίνηση μας.

(a) Θεωρώντας p=0.4, επιλύστε το συγκεκριμένο στοχαστικό πρόβλημα και προσδιορίστε τις βέλτιστες ακολουθητέες κατευθύνσεις σε κάθε κόμβο του δικτύου καθώς και τα ελάχιστα κόστη μεταφοράς μέχρι τον τερματισμό της διαδρομής. Ποια είναι η βέλτιστη απόφαση στον κόμβο A;

(b) Προσδιορίστε την ελάχιστη τιμή της πιθανότητας p για την οποία η βέλτιστη απόφαση που βρήκατε στο ερώτημα (α) για τον κόμβο A δεν μεταβάλλεται. Δώστε τη βέλτιστη απόφαση στον κόμβο A για κάθε τιμή της πιθανότητας $p \in [0,1]$.



8.4 Ζητούμε να προσδιορίσουμε τον βέλτιστο έλεγχο για την ελαχιστοποίηση του τετραγωνικού κριτηρίου

$$J = \frac{1}{2}\left[x_1(K)^2 + x_2(K)^2\right] + \frac{1}{2}\sum_{k=0}^{K-1}\left[0.75u(k)^2 + x_2(k)^2\right]$$

λαμβάνοντας υπόψη το γραμμικό δυναμικό σύστημα

$x_1(k+1)=2x_1(k)+0.5u(k),\ \ x_1(0)=0.5$
$x_2(k+1)=0.5x_2(k)+x_1(k),\ \ x_2(0)=-3$

(a) Προσδιορίστε τους πίνακες συστήματος **A** και εισόδου **B** καθώς και τους πίνακες βαρών **Q**, **R** και **S**.

(b) Θεωρώντας K=2 υπολογίστε τον βέλτιστο έλεγχο u(k) για k=0,1.

Ερωτήσεις θεωρίας/κατανόησης:

(c) Θεωρήστε τώρα ότι στην πρώτη ως άνω καταστατική εξίσωση προστίθεται μια στοχαστική διαταραχή z(k), ανεξάρτητη των x(k), u(k), z(k−1), z(k−2),…, με μέσο όρο μηδέν και $E\{z(k)^2\}=10$. Προσδιορίστε τον βέλτιστο έλεγχο u(k), k=0,1, για ελαχιστοποίηση της αναμενόμενης τιμής του ως άνω κριτηρίου.

(d) Ερμηνεύστε το ως άνω αντικειμενικό κριτήριο.

8.5 Ένας παίκτης καζίνου έχει στην διάθεση του μια μάρκα. Πρόκειται να στοιχηματίσει 3 φορές, και κάθε φορά μπορεί να παίζει το πολύ όσες μάρκες έχει στην διάθεση του. Σε κάθε στοίχημα (γύρο) είτε κερδίζει είτε χάνει τόσες μάρκες όσες στοιχηματίζει. Σε κάθε γύρο η πιθανότητα επιτυχίας είναι 0.6 και η πιθανότητα αποτυχίας 0.4. Αντικειμενικός σκοπός του παίκτη είναι να έχει στη διάθεση του τουλάχιστον 3 μάρκες στο τέλος των τριών γύρων, ώστε να μπορεί να πληρώσει το ξενοδοχείο του. Προς το σκοπό αυτό, συμβολίζοντας με x(K) (K=3) τον αριθμό των μαρκών που έχει στα χέρια του μετά τους τρεις γύρους, θεωρούμε μεγιστοποίηση του αντικειμενικού κριτηρίου $J = E\{\vartheta[x(K)]\}$ όπου

$$\vartheta[x(K)]=\begin{cases}1 & \text{αν } x(k)\geq 3\\ 0 & \text{αλλιώς.}\end{cases}$$

(a) Μοντελοποιήστε το πρόβλημα αυτό σαν ένα πρόβλημα στοχαστικού διακριτού δυναμικού προγραμματισμού αναγράφοντας (i) το κριτήριο κόστους, (ii) την καταστατική εξίσωση και όλους τους περιορισμούς.

(b) Αναγράψτε τον αναδρομικό τύπο Bellman για το συγκεκριμένο πρόβλημα στοχαστικού διακριτού δυναμικού προγραμματισμού.

(c) Προσδιορίστε την βέλτιστη στρατηγική του παίκτη (βέλτιστος κανόνας ελέγχου).

Ερωτήσεις θεωρίας/κατανόησης:

(d) Αλλάζει η λύση του προβλήματος αν στο ως άνω αντικειμενικό κριτήριο η συνάρτηση ϑ παίρνει τιμές 10 και 0 αντί 1 και 0, αντίστοιχα; Αιτιολογήστε την απάντηση σας.

(e) Σε ένα πρόβλημα στοχαστικού διακριτού δυναμικού προγραμματισμού εμπεριέχονται p στοχαστικές μεταβλητές $\mathbf{z}(k)\in R^p$ με γνωστή πιθανολογική κατανομή. Θεωρούμε μια αιτιοκρατική εκδοχή του προβλήματος αντικαθιστώντας στη θέση των $\mathbf{z}(k)$ τους αντίστοιχους μέσους όρους. Πόσες φορές υψηλότερος είναι ο υπολογιστικός φόρτος επίλυσης του στοχαστικού προβλήματος σε σχέση με την αιτιοκρατική εκδοχή του;

8.6 Ένας παίκτης καζίνου θέλει να υπολογίσει πόσα χρήματα από το κεφάλαιό του θα στοιχηματίσει σε κάθε γύρο για ένα παιχνίδι που θα παίξει συνολικά K φορές (γύρους). Σε κάθε στοίχημα (γύρο) μπορεί να παίξει το πολύ όσο είναι το κεφάλαιο που έχει στην διάθεσή του, και είτε κερδίζει είτε χάνει το ποσό που στοιχηματίζει. Σε κάθε γύρο η πιθανότητα επιτυχίας είναι p και η πιθανότητας αποτυχίας 1−p, όπου 0.5<p<1.

Συμβολίζοντας με x(K) το κεφάλαιο του παίκτη μετά από K γύρους, θεωρούμε μεγιστοποίηση της συνάρτησης ωφέλειας J = E{ ϑ [x(K)]} όπου ϑ [x(K)]=ln[x(K)].

(a) Μοντελοποιήστε το πρόβλημα αυτό σαν ένα πρόβλημα στοχαστικού διακριτού δυναμικού προγραμματισμού αναγράφοντας (i) το κριτήριο κόστους, (ii) την καταστατική εξίσωση και όλους τους περιορισμούς.

(b) Αναγράψτε τον αναδρομικό τύπο Bellman για το συγκεκριμένο πρόβλημα στοχαστικού διακριτού δυναμικού προγραμματισμού.

(c) Προσδιορίστε αναλυτικά την βέλτιστη στρατηγική του παίκτη (βέλτιστος κανόνας ελέγχου) u(k)=R[x(k)] για k=0, 1, …, K−1. Ποιο είναι το βέλτιστο αναμενόμενο κριτήριο ωφέλειας V[x(k), k]; (Υπόδειξη: Αρχίστε την επίλυση με την βαθμίδα K−1 και γενικεύστε κατάλληλα τα αποτελέσματα για κάθε k με την αντίστοιχη αιτιολόγηση. Τυπολόγιο: (lnx)΄=1/x, ln(xy)=lnx+lny.)

Ερωτήσεις θεωρίας/κατανόησης:

(d) Η ως άνω συνάρτηση ωφέλειας είναι φιλική ή εχθρική προς το ρίσκο; Θεωρούμε τώρα τη συνάρτηση ωφέλειας ϑ [x(K)]=x(K)^2. Τι είδους παίκτη (φιλικό ή εχθρικό προς το ρίσκο) χαρακτηρίζει η συνάρτηση αυτή; Αιτιολογήστε τις απαντήσεις σας.

8.7 Ενδιαφερόμαστε για την αγορά μίας κάρτας μνήμης για τη φωτογραφική μας μηχανή μέσω της ηλεκτρονικής σελίδας αγορών κάποιου καταστήματος για να την χρησιμοποιήσουμε σε ένα ταξίδι που έχουμε προγραμματίσει σε K=3 ημέρες. Οι προσφορές που έχει το κατάστημα για ηλεκτρονική αγορά των προϊόντων του, ισχύουν μόνο για μία ημέρα, και σε περίπτωση αγοράς κάποιου προϊόντος η παράδοσή του γίνεται αυθημερόν. Θεωρούμε επίσης, ότι οι προσφορές που παρέχει το κατάστημα κάθε ημέρα είναι ανεξάρτητες μεταξύ τους, με γνωστή εκ των προτέρων πιθανολογική κατανομή. Κάθε ημέρα k=0,1,2 αποφασίζουμε για την αγορά ή μη της κάρτας μνήμης ανάλογα με την τρέχουσα προσφορά z(k). Αν μέχρι την τελευταία ημέρα k=2 πριν την αναχώρησή μας δεν έχουμε κάνει την αγορά, τότε είμαστε αναγκασμένοι να αποδεχτούμε την τιμή z(K−1) και να αγοράσουμε την κάρτα μνήμης. Η πιθανολογική κατανομή P του z(k) έχει ως εξής:

$$P(z) = \begin{cases} \dfrac{z^2}{50} - \dfrac{3z}{5} + \dfrac{9}{2} & \text{για } 15 \le z \le 20 \\ -\dfrac{z^2}{50} + z - \dfrac{23}{2} & \text{για } 20 \le z \le 25 \\ 0 & \text{για } z \le 15 \\ 1 & \text{για } z \ge 25 \end{cases}$$

Ζητείται η βέλτιστη πολιτική αποδοχής ή απόρριψης των προσφορών.

(a) Μοντελοποιήστε το πρόβλημα σαν ένα πρόβλημα Στοχαστικού Βέλτιστου Ελέγχου με το κριτήριο κόστους, την καταστατική εξίσωση και όλους τους περιορισμούς.

(b) Προσδιορίστε τη βέλτιστη πολιτική απόφασης αποδοχής ή απόρριψης των προσφορών για κάθε k=0,1,2 (πρόβλημα βέλτιστης στάσης).

9. ΕΠΙΛΟΓΟΣ Η ΕΠΙΛΟΓΕΣ ΠΡΑΚΤΙΚΗΣ ΕΦΑΡΜΟΓΗΣ

Η επίλυση πρακτικών προβλημάτων αιτιοκρατικής ή στοχαστικής βελτιστοποίησης απαιτεί την διαθεσιμότητα μερικών δεδομένων, όπως ενός μαθηματικού πρότυπου της αντίστοιχης διεργασίας, μετρήσεις κατάστασης πραγματικού χρόνου, πρόβλεψη των μελλοντικών διαταραχών (στην αιτιοκρατική περίπτωση) ή της πιθανολογικής κατανομής τους (στην στοχαστική περίπτωση). Φυσικά, τα δεδομένα αυτά είναι συνήθως (περισσότερο ή λιγότερο ακριβείς) προσεγγίσεις της πραγματικότητας. Επί πλέον, η επίλυση ενός προβλήματος εφαρμογής θέτει συχνά το ερώτημα του απαιτούμενου υπολογιστικού φόρτου.

Αν για κάποια συγκεκριμένη εφαρμογή ο υπολογιστικός φόρτος δεν αποτελεί εμπόδιο, τότε συνιστάται η εφαρμογή της στοχαστικής βελτιστοποίησης και μάλιστα με την μορφή ΒΚΕ, που όπως είδαμε στο Κεφάλαιο 7.1 δίδει πάντα καλύτερα-ίσα αποτελέσματα σε σχέση με ΒΤΕ και ΒΤΕΚΒ. Θα πρέπει όμως να τονισθεί ότι το γεγονός αυτό ισχύει μόνον αν η πιθανολογική κατανομή που χρησιμοποιήθηκε για τον υπολογισμό του ΒΚΕ είναι αρκούντως ακριβής. Σε αντίθετη περίπτωση μπορεί η ΒΤΕΚΒ, εφαρμοζόμενη κατ΄ επανάληψη, με **ενημερωμένα**, ακριβέστερα δεδομένα πιθανολογικής κατανομής, να οδηγήσει σε καλύτερα αποτελέσματα από τον ΒΚΕ. Ο τελευταίος μπορεί όμως να επαυξηθεί (αν ο υπολογιστικός φόρτος δεν γίνεται υπέρμετρος), όπως περιγράφηκε στο Κεφάλαιο 7.5γ, ώστε να συμπεριλάβει ενδεχόμενες πληροφορίες πραγματικού χρόνου που αφορούν την μελλοντική πιθανολογική κατανομή των διαταραχών.

Εφαρμόζοντας την Αρχή Ισοδυναμίας Βεβαιότητας (βλ. Κεφάλαιο 8.5), αντικαθιστώντας δηλαδή όλες τις στοχαστικές μεταβλητές του ΠΣΒΕ με τις αντίστοιχες αναμενόμενες τιμές, λαμβάνουμε ένα αιτιοκρατικό ΠΒΕ. Όπως είδαμε στο Κεφάλαιο 8.5, η Αρχή Ισοδυναμίας Βεβαιότητας οδηγεί στην ακριβή επίλυση του ΠΣΒΕ μόνο σε ειδικές περιπτώσεις, όπως π.χ. στην περίπτωση στοχαστικού ΓΤ Ελέγχου. Σε όλες τις άλλες περιπτώσεις, η επίλυση του προκύπτοντος αιτιοκρατικού ΠΒΕ οδηγεί το πολύ σε κατά προσέγγιση λύση του αρχικού ΠΣΒΕ. Το προφανές πλεονέκτημα του ΠΒΕ είναι βέβαια η μείωση του υπολογιστικού φόρτου. Επί πλέον όμως, ακόμη και αν οι πιθανολογικές κατανομές στο ΠΣΒΕ είναι ακριβείς, ο υπολογισμός αιτιοκρατικών βέλτιστων τροχιών σε πραγματικό χρόνο, με ενημερωμένες, αρκούντως ακριβείς προγνώσεις διαταραχών, μπορεί να οδηγήσει τελικά σε καλύτερα αποτελέσματα από την επίλυση του ΠΣΒΕ.

Το νόημα των παραπάνω διαπιστώσεων είναι το εξής. Η θεωρία βελτιστοποίησης γεννά πιθανόν στον άπειρο σχεδιαστή την εντύπωση ότι κάθε πρακτικό πρόβλημα μπορεί να διατυπωθεί κατά έναν και μοναδικό τρόπο και να επιλυθεί άμεσα, οδηγώντας αυτομάτως σε βέλτιστα αποτελέσματα. Στην πραγματικότητα όμως, αυστηρά βέλτιστες λύσεις πρακτικών προβλημάτων είναι κατά κανόνα ανέφικτες, κυρίως λόγω έλλειψης επαρκών ή ακριβών δεδομένων. Έτσι, αυστηρά βέλτιστες λύσεις υπάρχουν ουσιαστικά μόνον στα μαθηματικά. Για την πρακτική εφαρμογή, υπάρχει ένα φάσμα επιστημονικά τεκμηριωμένων μεθόδων βελτιστοποίησης που στην πράξη οδηγούν σε περισσότερο ή λιγότερο αποτελεσματικές λύσεις κατά περίπτωση. Βέβαια η κατάλληλη εφαρμογή των μεθόδων βελτιστοποίησης παρέχει κατά κανόνα σημαντικά πρακτικά πλεονεκτήματα συγκριτικά με μεθόδους εμπειρικές ή στηριζόμενες στην κοινή λογική (έμπειρα συστήματα, ασαφής λογική κλπ.). Τα πλεονεκτήματα αυτά αναφέρονται τόσο στην ποιότητα των αποτελεσμάτων, όσο και στο εύρος του πεδίου εφαρμογής, στην ευελιξία, αλλά συχνά και στην απλότητα των αντίστοιχων στρατηγικών. Σε προβλήματα Δυναμικού Προγραμματισμού, υπό ιδανικές συνθήκες ακριβών δεδομένων (πράγμα που σπάνια συναντάται σε πρακτικές εφαρμογές), κάποιες μέθοδοι (π.χ. ΒΚΕ του ΠΣΒΕ) υπερτερούν άλλων (π.χ. ΒΤΕ ή ΒΤΕΚΒ). Υπό πρακτικές συνθήκες κατά προσέγγιση δεδομένων όμως, και λαμβάνοντας υπόψη το ερώτημα του υπολογιστικού φόρτου,

μπορεί μια από τις θεωρητικά υποβέλτιστες προσεγγίσεις να δίδει τελικά καλύτερα αποτελέσματα σε κάποια συγκεκριμένη πρακτική εφαρμογή. Ποιά προσέγγιση είναι τελικά η καλύτερη για κάποιο συγκεκριμένο πρόβλημα, είναι δύσκολο να απαντηθεί γενικά. Η θεωρία Βελτιστοποίησης γενικά και Δυναμικού Προγραμματισμού ειδικότερα παραθέτουν μια σειρά **μεθοδολογικών εργαλείων** για την συστηματική και αποτελεσματική επίλυση προβλημάτων προγραμματισμού, αποφάσεων ή ελέγχου. Η επιλογή του καταλληλότερου εργαλείου για κάποιο συγκεκριμένο πρόβλημα, επαφίεται στις γνώσεις και την εμπειρία του σχεδιαστή και προϋποθέτει βέβαια επαρκή γνώση τόσο των θεωρητικών μεθόδων όσο και των ιδιαιτεροτήτων και απαιτήσεων της συγκεκριμένης εφαρμογής. Ας μη ξεχνάμε εξάλλου ότι παρομοίως δημιουργική συνδρομή του σχεδιαστή απαιτείται συχνά και στην **διατύπωση** (μοντελοποίηση) κάποιου συγκεκριμένου προβλήματος, δηλαδή στην κατάλληλη επιλογή του υπό βελτιστοποίηση κριτηρίου, του μαθηματικού προτύπου, των περιορισμών κλπ. Αν οι επιλογές της διατύπωσης του προβλήματος ή της μεθόδου λύσης δεν είναι προφανείς, ενδέχεται ο σχεδιαστής να χρειασθεί να διερευνήσει μια σειρά εναλλακτικών προσεγγίσεων, επιλέγοντας τελικά εκείνη που ανταποκρίνεται κατά τον καλύτερο τρόπο στις ανάγκες της δεδομένης εφαρμογής. Ο βαθμός ευστοχίας της επιλογής του κατά την πρακτική εφαρμογή χαρακτηρίζει την ποιότητα του σχεδιαστή.

ΠΡΟΤΕΙΝΟΜΕΝΗ ΒΙΒΛΙΟΓΡΑΦΙΑ

Bellman, R.E. (1957) **Dynamic Programming**. Princeton University Press, Princeton.

Bertsekas, D.P. (1995) **Dynamic Programming and Optimal Control - Volume I**. Athena Scientific, Belmont, Massachusetts.

Larson, R.E., Casti, J.L. (1978) **Principles of Dynamic Programming - Part I: Basic Analytic and Computational Methods**. Dekker, New York.

Larson, R.E., Casti, J.L. (1982) **Principles of Dynamic Programming - Part II: Advanced Theory and Applications**. Dekker, New York.

Papageorgiou, M., Leibold, M., Buss, M. (2015) **Optimierung - statische, dynamische, stochastische Verfahren für die Anwendung** (4η Εκδοση). Springer, Berlin.

ΠΑΡΑΡΤΗΜΑ Α: ΔΥΝΑΜΙΚΑ ΣΥΣΤΗΜΑΤΑ

Α.1 Γενικοί Ορισμοί

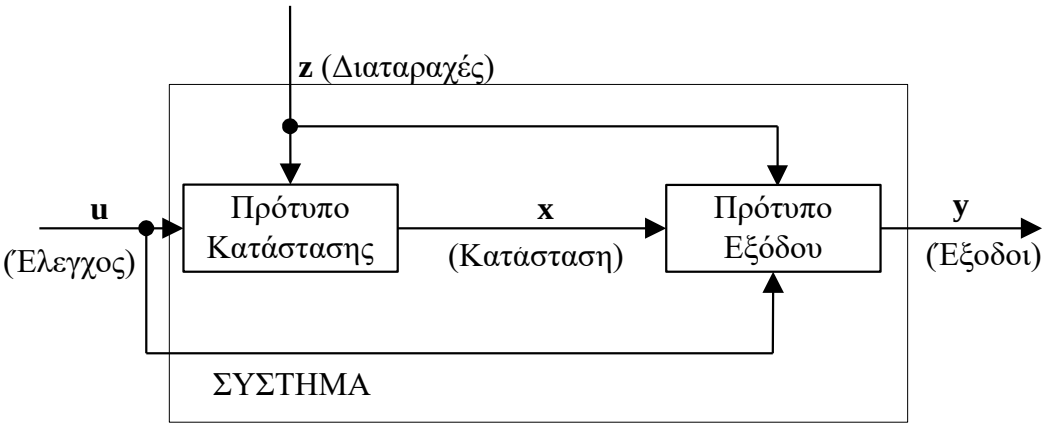
Ονομάζουμε **σύστημα** ένα σύνολο πραγματικών φαινομένων ή μαθηματικών σχηματισμών με συγκεκριμένη εσωτερική λειτουργία και σαφείς αλληλεπιδράσεις με το εξωσυστημικό περιβάλλον. Σύμφωνα με το Σχ. Α.1, το περιβάλλον επιδρά στη λειτουργία του συστήματος μέσω των **εισόδων**, ενώ το ενδιαφέρον μας σχετικά με το σύστημα επικεντρώνεται στις **εξόδους** του συστήματος. Οι είσοδοι του συστήματος μπορούν να ταξινομηθούν σε **ελεγχόμενες εισόδους** ή **εισόδους ελέγχου** ή κατ΄ εξοχήν **εισόδους** $\mathbf{u}\in R^m$ και σε μη ελεγχόμενες εισόδους ή **διαταραχές** $\mathbf{z}\in R^p$. Πέραν αυτού διακρίνουμε τις εσωτερικές **μεταβλητές κατάστασης** του συστήματος $\mathbf{x}\in R^n$ από τις (συνήθως μετρήσιμες) **εξόδους** $\mathbf{y}\in R^q$, βλ. Σχ. Α.2, βλ. Πρόβλημα Α.1.

Τα συστήματα διακρίνονται σε γραμμικά και μη γραμμικά, σε στατικά και δυναμικά, σε αιτιοκρατικά και στοχαστικά. Μαθηματικοί υπολογισμοί των αλληλεπιδράσεων είναι δυνατοί μέσω ενός **μαθηματικού πρότυπου** (μοντέλου) του συστήματος. Ονομάζουμε συστήματα **μιας** εισόδου ελέγχου (m = 1) και **μιας** εξόδου (q = 1) **SISO-συστήματα** (**S**ingle-**I**nput-**S**ingle-**O**utput) και συστήματα **πολλαπλών** εισόδων ελέγχου και **πολλαπλών** εξόδων **MIMO-συστήματα** (**M**ultiple-**I**nput-**M**ultiple-**O**utput). Δυναμικά συστήματα μπορούν να περιγραφούν σε συνεχή χρόνο ή σε διακριτό χρόνο. Περιοριζόμαστε εδώ σε δυναμικά συστήματα συγκεντρωμένων παραμέτρων, που μπορούν να περιγραφούν μέσω συνήθων διαφορικών εξισώσεων ή εξισώσεων διαφορών.

Α.2 Δυναμικά Συστήματα Συνεχούς Χρόνου

Θεωρούμε στην κατηγορία αυτή δυναμικά συστήματα που περιγράφονται μέσω των ακόλουθων εξισώσεων

Σχήμα Α.1: Γενική μορφή συστήματος.



Σχήμα Α.2: Γενική δομή συστήματος.

Πρότυπο κατάστασης: $\dot{\mathbf{x}}(t) = \mathbf{f}[\mathbf{x}(t), \mathbf{u}(t), \mathbf{z}(t), t], \ \mathbf{x}(t_0) = \mathbf{x_0}$ (A.1)

Πρότυπο εξόδου: $\mathbf{y}(t) = \mathbf{c}[\mathbf{x}(t), \mathbf{u}(t), \mathbf{z}(t), t]$ (A.2)

όπου (A.1) είναι η **διαφορική εξίσωση κατάστασης**, (A.2) είναι η **εξίσωση εξόδου** του συστήματος, t είναι η ανεξάρτητη μεταβλητή χρόνου και **f**, **c** είναι εν γένει μη γραμμικές συναρτήσεις.

Αν θεωρήσουμε δεδομένες τις χρονικές συναρτήσεις των εισόδων $\mathbf{u}(t)$, $\mathbf{z}(t)$, $t\in[t_0, t_e]$, και δεδομένη την **αρχική κατάσταση** $\mathbf{x}(t_0) = \mathbf{x}_0$, μπορούμε (υπό ορισμένες προϋποθέσεις που στις περισσότερες πρακτικά ενδιαφέρουσες εφαρμογές πράγματι εκπληρούνται) να υπολογίσουμε μέσω ολοκλήρωσης της (A.1) και αντικατάστασης στην (A.2) τις μοναδικές χρονικές συναρτήσεις $\mathbf{x}(t)$, $\mathbf{y}(t)$, $t\in[t_0, t_e]$. Ονομάζουμε μια ομάδα χρονικών συναρτήσεων $[\mathbf{u}_s(t), \mathbf{z}_s(t), \mathbf{x}_s(t), \mathbf{y}_s(t), t_0 \le t \le t_e]$ **συστημική** αν εκπληροί τις (A.1), (A.2).

Ονομάζουμε ένα δυναμικό σύστημα (A.1) **ελεύθερο**, αν δεν περιλαμβάνει εισόδους ελέγχου ούτε διαταραχές. **Σημεία ισορροπίας** $\bar{\mathbf{x}}$ ενός ελεύθερου δυναμικού συστήματος έχουν την ιδιότητα $\mathbf{f}(\bar{\mathbf{x}},t) = \mathbf{0}$. Το δυναμικό σύστημα ονομάζεται **ασυμπτωτικά ευσταθές** στην περιοχή $\mathcal{X}(t_0)$ περί το σημείο ισορροπίας $\bar{\mathbf{x}}$, αν για κάθε $\mathbf{x}(t_0)\in\mathcal{X}(t_0)$ η επίλυση της (A.1) εκπληροί $\lim_{t\to\infty}\mathbf{x}(t) = \bar{\mathbf{x}}$.

Παράδειγμα Α.1: Θεωρούμε το μαθηματικό πρότυπο ενός οχήματος εδάφους κινουμένου επί μιάς ευθείας με θέση x_1 και ταχύτητα x_2. Η διαφορική εξίσωση κατάστασης αποδίδεται από

$$\dot{x}_1 = x_2 \qquad (A.3)$$

$$\dot{x}_2 = -\alpha_1 x_2 |x_2| + u + z \qquad (A.4)$$

όπου u αντιστοιχεί στην δύναμη ώθησης του κινητήρα, $-\alpha_1 x_2 |x_2|$ στην αντίσταση του αέρα (με σταθερή παράμετρο $\alpha_1 \ge 0$) και z στην δύναμη που ασκεί στην κατεύθυνση κίνησης ενδεχόμενος άνεμος. Στο πρότυπο κατάστασης (A.3), (A.4) θεωρούμε $\mathbf{x} = [x_1\ x_2]^T$ το διάνυσμα κατάστασης, u τον έλεγχο και z την διαταραχή του συστήματος. Σε περίπτωση μετρήσιμης ταχύτητας, έχουμε το πρότυπο εξόδου

$$y = x_2. \qquad (A.5)$$

■

Α.3 Δυναμικά Συστήματα Διακριτού Χρόνου

Θεωρούμε στην κατηγορία αυτή δυναμικά συστήματα που περιγράφονται μέσω των ακόλουθων εξισώσεων

Πρότυπο κατάστασης: $\mathbf{x}(k+1) = \mathbf{f}[\mathbf{x}(k), \mathbf{u}(k), \mathbf{z}(k), k], \ \mathbf{x}(k_0) = \mathbf{x}_0$ (A.6)

Πρότυπο εξόδου: $\mathbf{y}(k) = \mathbf{c}[\mathbf{x}(k), \mathbf{u}(k), \mathbf{z}(k), k]$ (A.7)

όπου $k = k_0, k_0+1,...$ είναι ο δείκτης διακριτού χρόνου. Για όλες τις μεταβλητές ισχύει $.(k) = .(kT)$, όπου T είναι η **περίοδος** ή **διάστημα διακριτού χρόνου**. Η (Α.6) είναι η **καταστατική εξίσωση διαφορών**, η (Α.7) είναι η **εξίσωση εξόδου** του συστήματος και **f**, **c** είναι εν γένει μη γραμμικές συναρτήσεις. Το σύστημα (Α.6), (Α.7) αναφέρεται συχνά και ως **σύστημα αποφάσεων πολλαπλών βαθμίδων**, όπου οι αποφάσεις αντιστοιχούν στον έλεγχο $\mathbf{u}(k)$ του συστήματος.

Αν θεωρήσουμε δεδομένες τις χρονικά διακριτές τιμές των εισόδων $\mathbf{u}(k)$, $\mathbf{z}(k)$, $k\in[k_0,K-1]$, και δεδομένη την αρχική κατάσταση $\mathbf{x}(k_0) = \mathbf{x}_0$, μπορούμε να υπολογίσουμε μέσω των (Α.6), (Α.7) τις μοναδικές χρονικές συναρτήσεις $\mathbf{x}(k)$, $\mathbf{y}(k-1)$, $k\in[k_0,K]$.

Οι ορισμοί της συστημικής ομάδας διακριτών χρονικών συναρτήσεων, των ελεύθερων συστημάτων και των ασυμπτωτικά ευσταθών συστημάτων είναι ανάλογοι με τους αντίστοιχους ορισμούς στο Κεφάλαιο Α.2.

Δυναμικά πρότυπα διακριτού χρόνου προέρχονται είτε άμεσα, από την περιγραφή αντίστοιχων πραγματικών συστημάτων, είτε έμμεσα, από την διακριτή προσέγγιση προτύπων συνεχούς χρόνου. Αν εφαρμόσουμε, για παράδειγμα, τον κανόνα Euler στην (Α.1) έχουμε

$$\mathbf{x}(k+1) = \mathbf{x}(k) + T \cdot \mathbf{f}[\mathbf{x}(k), \mathbf{u}(k), \mathbf{z}(k), k\cdot T] \qquad (A.7)$$

δηλαδή μια διακριτή προσέγγιση της διαφορικής εξίσωσης (Α.1).

Παράδειγμα Α.2: Θεωρούμε έναν εισοδηματία ο οποίος στην αρχή κάθε νέας χρονιάς αποφασίζει να επενδύσει ένα μέρος του κεφαλαίου του και να ξοδεύσει το υπόλοιπο. Έστω x(k) το ύψος του κεφαλαίου στην αρχή της χρονιάς k και u(k) το προς επένδυση τμήμα του. Το επενδυμένο κεφάλαιο αυξάνει με σταθερό επιτόκιο κάθε χρονιά έτσι ώστε

$$x(k+1) = \alpha \cdot u(k) \qquad (A.8)$$

όπου $\alpha > 1$. Από το μη επενδυθέν τμήμα του κεφαλαίου κάθε χρονιάς καλύπτει ο εισοδηματίας ανάγκες ιατρικής περίθαλψης, ύψους z(k), ενώ το υπόλοιπο $x(k) - u(k) - z(k)$ χρησιμοποιείται για άλλες προσωπικές ανάγκες. Ο εισοδηματίας θεωρεί ότι η ποιότητα ζωής του κάθε χρόνο, y(k), είναι άμεση συνάρτηση αυτού του τελευταίου ποσού, δηλαδή

$$y(k) = c[x(k) - u(k) - z(k)]. \qquad (A.9)$$

■

Α4. Ελεγξιμότητα και Παρατηρησιμότητα Γραμμικών Συστημάτων

Θεωρούμε ένα γραμμικό σύστημα συνεχούς χρόνου της μορφής

$$\dot{\mathbf{x}}(t) = \mathbf{A}(t)\,\mathbf{x}(t) + \mathbf{B}(t)\,\mathbf{u}(t),\ \mathbf{x}(t_0) = \mathbf{x}_0 \qquad (A.10)$$

$$\mathbf{y}(t) = \mathbf{C}(t)\,\mathbf{x}(t) \qquad (A.11)$$

ή ένα γραμμικό σύστημα διακριτού χρόνου της μορφής

$$\mathbf{x}(k+1) = \mathbf{A}(k)\,\mathbf{x}(k) + \mathbf{B}(k)\,\mathbf{u}(k),\ \mathbf{x}(k_0) = \mathbf{x}_0 \qquad (A.12)$$

$$\mathbf{y}(k) = \mathbf{C}(k)\,\mathbf{x}(k). \qquad (A.13)$$

Το σύστημα ονομάζεται **ελέγξιμο** αν οποιαδήποτε αρχική κατάσταση $\mathbf{x}_0$ μπορεί να μεταφερθεί, μέσω κατάλληλων τιμών του ελέγχου **u**, σε οποιοδήποτε σημείο του χώρου R^n σε πεπερασμένο χρόνο. Αν οι πίνακες **A**, **B** δεν εξαρτώνται από τον χρόνο, μπορεί να αποδειχθεί ότι ικανή και αναγκαία συνθήκη για να είναι ένα γραμμικό σύστημα (συνεχούς ή διακριτού χρόνου) ελέγξιμο, είναι ο ορθογώνιος πίνακας

$$[\mathbf{B}\ \ \mathbf{AB}\ \ \mathbf{A}^2\mathbf{B}\ \ \ldots\ \ \mathbf{A}^{n-1}\mathbf{B}] \qquad (A.14)$$

να είναι πλήρους βαθμού, δηλαδή βαθμού n.

Ένα γραμμικό σύστημα ονομάζεται **παρατηρήσιμο** αν η κατάστασή του **x** μπορεί να προσδιορισθεί μέσω παρατηρήσεων της εξόδου του **y** για ένα πεπερασμένο χρονικό διάστημα. Αν οι πίνακες **A**, **C** δεν εξαρτώνται από τον χρόνο, μπορεί να αποδειχθεί ότι ικανή και αναγκαία συνθήκη για να είναι ένα γραμμικό σύστημα (συνεχούς ή διακριτού χρόνου) παρατηρήσιμο, είναι ο ορθογώνιος πίνακας

$$[\mathbf{C}^T\ \ \mathbf{A}^T\mathbf{C}^T\ \ (\mathbf{A}^T)^2\mathbf{C}^T\ \ \ldots\ \ (\mathbf{A}^T)^{n-1}\,\mathbf{C}^T] \qquad (A.15)$$

να είναι πλήρους βαθμού, δηλαδή βαθμού n.

Προβλήματα

Α.1 Θεωρούμε το σύστημα θερμικής συμπεριφοράς ενός κλιματιζόμενου κλειστού χώρου. Δώστε παραδείγματα ελεγχόμενων εισόδων, διαταραχών, μεταβλητών κατάστασης και εξόδων του συστήματος.

ΠΑΡΑΡΤΗΜΑ Β: ΒΕΛΤΙΣΤΟΣ ΕΛΕΓΧΟΣ

Β.1 Διατύπωση του Προβλήματος Βέλτιστου Ελέγχου

Το **Πρόβλημα Βέλτιστου Ελέγχου** (**ΠΒΕ**) τίθεται σε σχέση με δυναμικά συστήματα συνεχούς ή διακριτού χρόνου. Στην περίπτωση συνεχούς χρόνου έχουμε το ακόλουθο ΠΒΕ[1]:

Δεδομένων της αρχικής κατάστασης $\mathbf{x}(0) = \mathbf{x}_0$ και των διαταραχών[2] $\mathbf{z}(t)$, $0 \le t \le t_e^*$, ζητούνται οι χρονικές συναρτήσεις των μεταβλητών ελέγχου $\mathbf{u}^*(t)$, $0 \le t \le t_e^*$, και κατάστασης $\mathbf{x}^*(t)$, $0 \le t \le t_e^*$, καθώς και ο τελικός χρόνος t_e^* που ελαχιστοποιούν το **κριτήριο κόστους**

$$J = \vartheta[\mathbf{x}(t_e), t_e] + \int_0^{t_e} \varphi[\mathbf{x}(t), \mathbf{u}(t), t]dt \qquad (B.1)$$

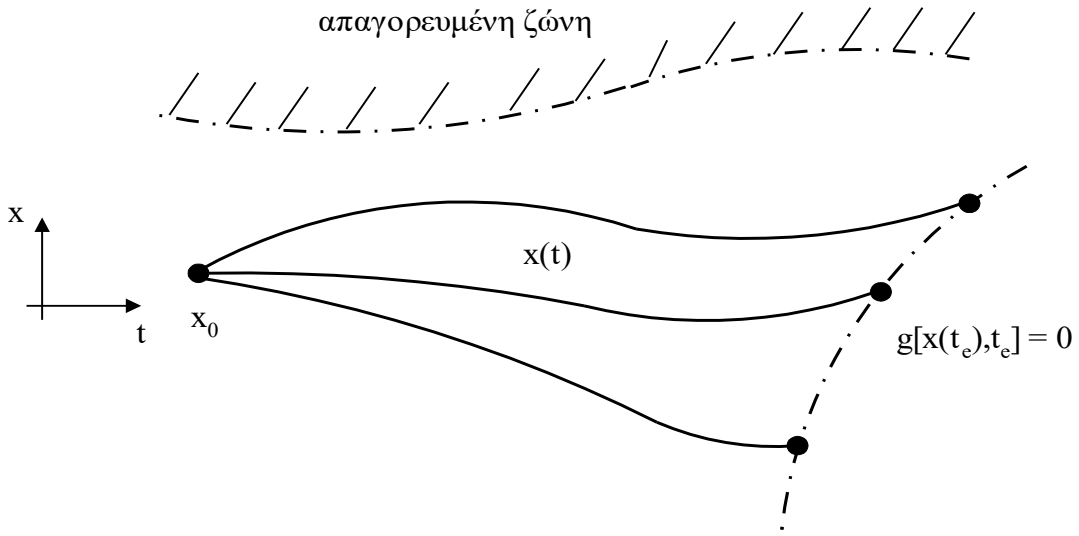
λαμβάνοντας υπόψη τους ακόλουθους περιορισμούς $\forall\ t \in [0,t_e]$

$$\dot{\mathbf{x}} = \mathbf{f}[\mathbf{x}(t), \mathbf{u}(t), t] \qquad (B.2)$$

$$\mathbf{h}[\mathbf{x}(t), \mathbf{u}(t), t] \le \mathbf{0} \qquad (B.3)$$

καθώς και τον τελικό στόχο

$$\mathbf{g}[\mathbf{x}(t_e), t_e] = \mathbf{0}. \qquad (B.4)$$



Σχήμα Β.1: Επεξήγηση του ΠΒΕ.

[1]Χωρίς περιορισμό της γενικότητας του προβλήματος θα θεωρήσουμε τον αρχικό χρόνο t_0=0.

[2]Καθότι οι διαταραχές θεωρούνται δεδομένες, μπορούν, χωρίς περιορισμό της γενικότητας, να απαλειφθούν από την διατύπωση του ΠΒΕ.

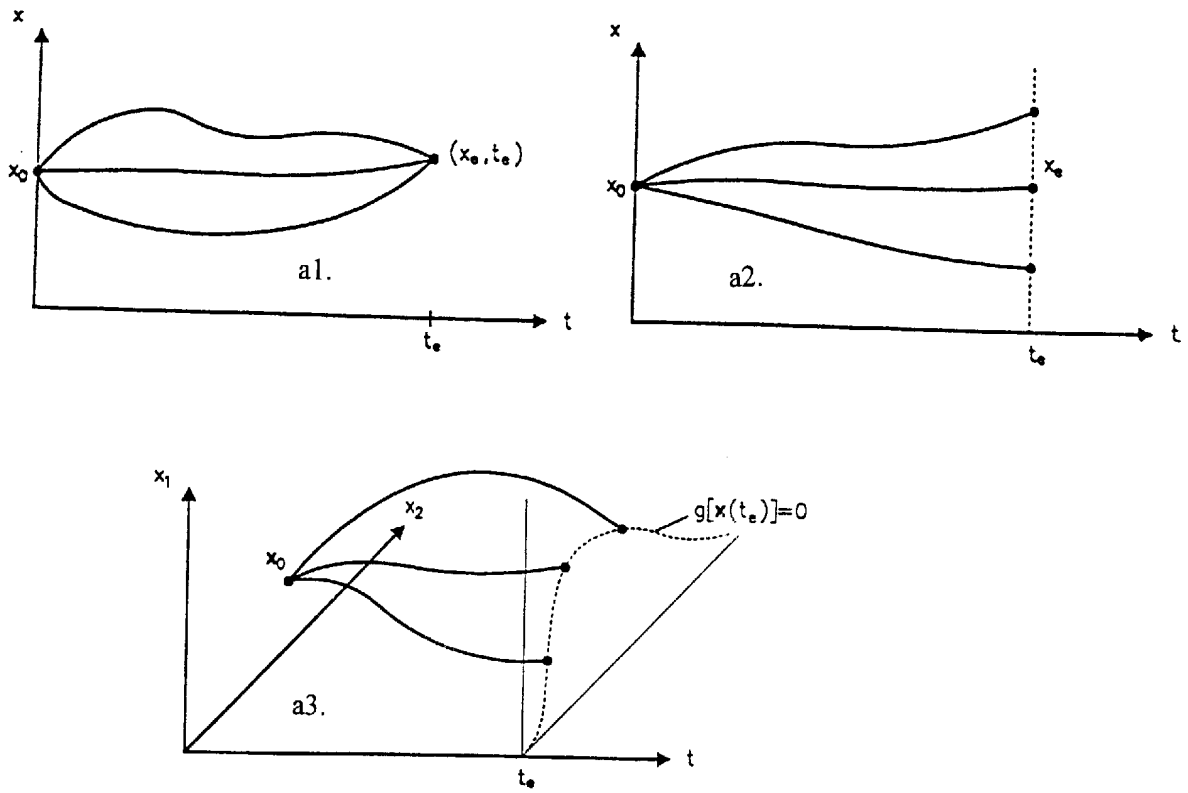
Το Σχ. Β.1 επεξηγεί σχηματικά το ΠΒΕ για ένα υποθετικό, μονοδιάστατο σύστημα. Η δεδομένη αρχική κατάσταση $\mathbf{x}_0$ του συστήματος (Β.2) πρέπει να μεταφερθεί στην τελική τροχιά (Β.4) αποφεύγοντας τις απαγορευμένες περιοχές που προσδιορίζονται από το σύνολο ανισοτήτων (Β.3). Υπάρχουν συνήθως πολλές επιτρεπτές χρονικές συναρτήσεις ελέγχου $\mathbf{u}(t)$, $0 \le t \le t_e 0$, που καθιστούν δυνατή την αναφερθείσα μεταφορά, και το Σχ. Β.1 απεικονίζει σχηματικά μερικές δυνατότητες. Μεταξύ αυτών, η λύση του ΠΒΕ επιτυγχάνεται από εκείνες τις χρονικές συναρτήσεις $\mathbf{u}^*(t)$, $\mathbf{x}^*(t)$, $0 \le t \le t_e^*$, που ελαχιστοποιούν το κριτήριο κόστους (Β.1).

Το ΠΒΕ διατυπώθηκε παραπάνω σε γενική μορφή. Ανάλογα με τα δεδομένα της εκάστοτε εφαρμογής, έχουμε μια σειρά ειδικών περιπτώσεων ως εξής:

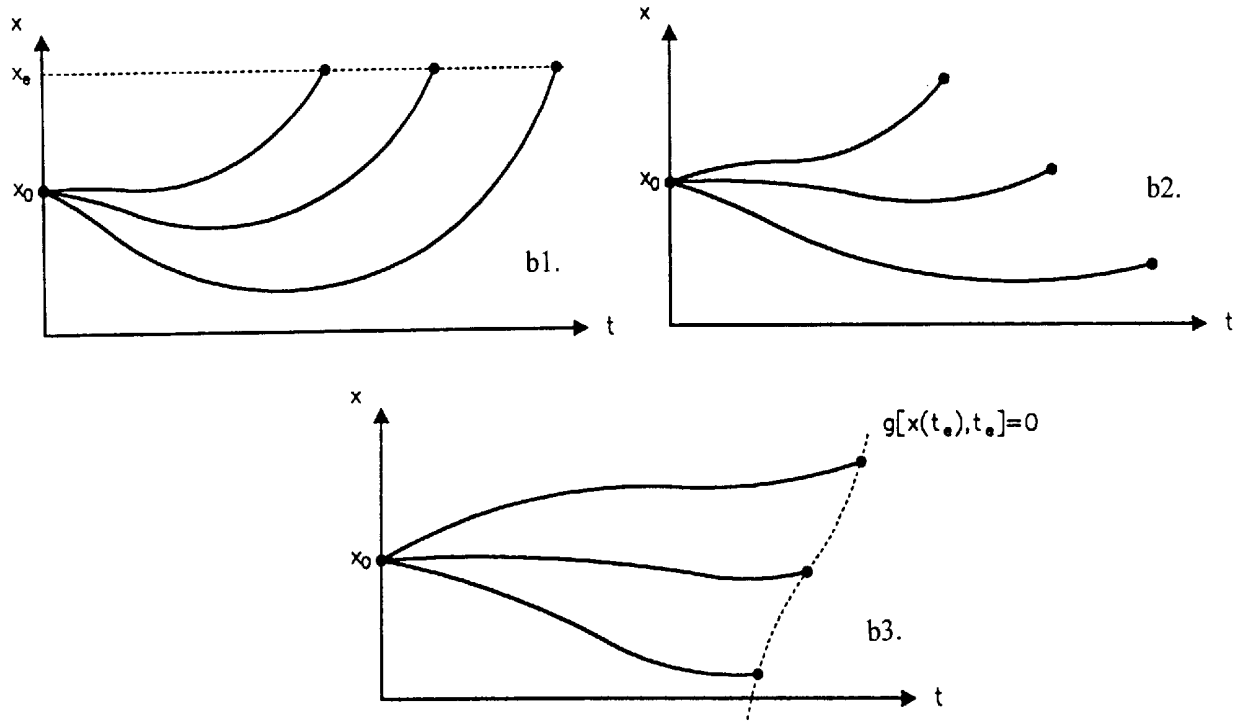
a. Δεδομένος τελικός χρόνος t_e:
   a1. Δεδομένη τελική κατάσταση $\mathbf{x}(t_e) = \mathbf{x}_e$
   a2. Ελεύθερη τελική κατάσταση
   a3. Τελική τροχιά $\mathbf{g}[\mathbf{x}(t_e)] = \mathbf{0}$.

b. Ελεύθερος τελικός χρόνος t_e:
   b1. Δεδομένη τελική κατάσταση $\mathbf{x}(t_e) = \mathbf{x}_e$
   b1. Ελεύθερη τελική κατάσταση
   b3. Τελική τροχιά (Β.4): η γενικότατη περίπτωση.

Οι ειδικές αυτές περιπτώσεις απεικονίζονται στα Σχ. Β.2 και Β.3.



Σχήμα Β.2: Ειδικές περιπτώσεις τελικής τροχιάς με δεδομένο τελικό χρόνο t_e.



Σχήμα Β.3: Ειδικές περιπτώσεις τελικής τροχιάς με ελεύθερο τελικό χρόνο t_e.

Παράδειγμα Β.1: Θεωρούμε το όχημα του Παραδ. Α.1 με σταθερή ώθηση ανέμου z = 1. Οι καταστατικές εξισώσεις (Α.3), (Α.4) προφανώς αντιστοιχούν στην (Β.2). Ένα ΠΒΕ για το όχημα αυτό μπορεί να διατυπωθεί κατά διάφορους τρόπους, ανάλογα με τα ενδιαφέροντα της ειδικής εφαρμογής. Καταρχήν είναι φυσικό ότι η δύναμη ώθησης u του κινητήρα δεν είναι απεριόριστη. Μπορούμε, επί παραδείγματι, να θεωρήσουμε

$$u_{min} \le u(t) \le u_{max}. \qquad (B.5)$$

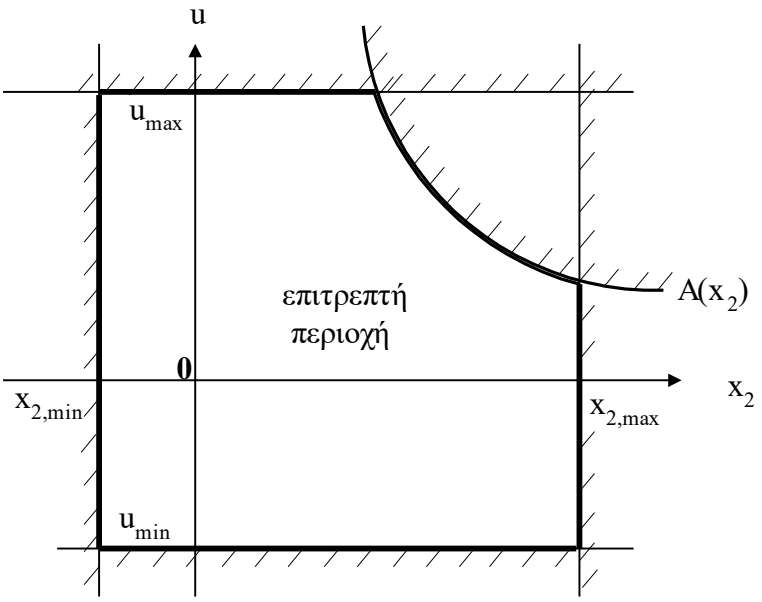
Πέραν αυτού θεωρούμε ότι η u(t) περιορίζεται επί πλέον ανάλογα με την αντίστοιχη ταχύτητα $x_2(t)$ ως εξής

$$u(t) \le A[x_2(t)] \qquad (B.6)$$

όπου Α είναι μια γνωστή συνάρτηση. Τέλος, το όχημα δεν επιτρέπεται να υπερβεί μια ανώτατη ταχύτητα προς κάθε κατεύθυνση

$$x_{2,min} \le x_2(t) \le x_{2,max}. \qquad (B.7)$$

Οι (Β.5), (Β.6), (Β.7) μπορούν να συγχωνευθούν στη μορφή (Β.3) και εκφράζουν τους περιορισμούς ανισοτήτων του προβλήματος. Το Σχ. Β.4 απεικονίζει την επιτρεπτή περιοχή ελέγχου.



Σχήμα Β.4: Επιτρεπτή περιοχή ελέγχου.

Ως παράδειγμα τελικού στόχου θεωρούμε ότι το όχημα πρέπει να εκκινήσει από την αρχική κατάσταση $\mathbf{x}_0 = [0\ 0]^T$ και να προλάβει ένα άλλο όχημα που κινείται με τροχιά

$$x_1(t_e) = 500 + 10 \cdot t_e. \qquad \text{(B.8)}$$

Η τελική ταχύτητα $x_2(t_e)$ και ο τελικός χρόνος t_e είναι ελεύθερα. Το πρόβλημα ανάγεται επομένως στην γενική περίπτωση b3. Σαν εναλλακτική αποστολή του οχήματος μπορούμε να θεωρήσουμε ότι ο τελικός χρόνος t_e είναι δεδομένος. Τότε η (Β.8) προκαθορίζει τον τόπο συνάντησης των δύο οχημάτων και έχουμε επομένως την περίπτωση a1 για την τελική θέση $x_1(t_e)$ και την περίπτωση a2 για την τελική ταχύτητα του οχήματος, άρα έχουμε συνολικά την περίπτωση a3.

Ας υποθέσουμε τώρα ότι η τελική ταχύτητα $x_2(t_e)$ είναι μεν ελεύθερη, δεν πρέπει όμως να διαφέρει πολύ από την ταχύτητα του προπορευόμενου οχήματος που είναι ίση με 10 (βλ. (Β.8)). Για να επιτευχθεί αυτό, μπορούμε να τιμωρήσουμε ενδεχόμενες διαφορές με την επιλογή (βλ. (Β.1))

$$\vartheta\,[\mathbf{x}(t_e)] = \alpha_1 \cdot [x_2(t_e) - 10]^2 \qquad \text{(B.9)}$$

όπου η σταθερή παράμετρος (βάρος) $\alpha_1 \geq 0$ αντανακλά την σχετική σπουδαιότητα της αντίστοιχης προδιαγραφής.

Τα καύσιμα Ε που καταναλίσκει το όχημα κατά την κίνησή του δίδονται από την συνάρτηση

$$E = x_2 \cdot u^2. \qquad \text{(B.10)}$$

Έτσι, αν θέλουμε να συντομεύσουμε το χρόνο συνάντησης των δύο οχημάτων, αλλά και να περιορίσουμε την αντίστοιχη κατανάλωση καυσίμων, μπορούμε να κάνουμε την επιλογή (βλ. (Β.1))

$$\varphi[\mathbf{x}(t), u(t), t] = \alpha_2 \cdot [x_1(t) - 500 - 10t]^2 + \alpha_3 \cdot x_2(t) \cdot u(t)^2 \qquad \text{(B.11)}$$

όπου $\alpha_2 \geq 0$, $\alpha_3 \geq 0$ είναι επίσης παράμετροι βάρους των αντίστοιχων προδιαγραφών.

Καταλήγουμε λοιπόν σε ένα ΠΒΕ όπου ζητούμε τις βέλτιστες τροχιές $x_1^*(t)$, $x_2^*(t)$, $u^*(t)$, $0 \leq t \leq t_e^*$, και τον βέλτιστο τελικό χρόνο t_e^* του αρχικά εν στάσει οχήματος, έτσι ώστε να ελαχιστοποιηθεί το κριτήριο κόστους (Β.1), όπως συγκεκριμενοποιήθηκε μέσω των (Β.9), (Β.11), λαμβάνοντας υπόψη την διαφορική εξίσωση κίνησης (Α.3), (Α.4) με z = 1 και την επιτρεπτή περιοχή ελέγχου (Β.5), (Β.6), (Β.7), και καταλήγοντας στην τελική τροχιά (Β.8) του προπορευόμενου οχήματος.

■

Ερωτήματα:

- Μπορούμε, χωρίς περιορισμό της γενικότητας του προβλήματος, να θέσουμε $\alpha_1 = 1$;
- Επιλέξτε το κριτήριο κόστους που ελαχιστοποιεί τον τελικό χρόνο t_e συνάντησης των δύο οχημάτων χωρίς άλλες προδιαγραφές.

■

Το ΠΒΕ διατυπώνεται κατ΄ανάλογο τρόπο σε σχέση με δυναμικά συστήματα διακριτού χρόνου ως εξής:

Δεδομένων της αρχικής κατάστασης $\mathbf{x}(0) = \mathbf{x}_0$ και των διαταραχών $\mathbf{z}(k)$, $0 \leq k \leq K^*-1$, ζητούνται οι χρονικές συναρτήσεις των μεταβλητών ελέγχου $\mathbf{u}^*(k)$, $0 \leq k \leq K^*-1$, και κατάστασης $\mathbf{x}^*(k)$, $1 \leq k \leq K^*$, καθώς και ο τελικός χρόνος K^* που ελαχιστοποιούν το κριτήριο κόστους

$$J = \vartheta[\mathbf{x}(K),K] + \sum_{k=0}^{K-1} \varphi[\mathbf{x}(k), \mathbf{u}(k), k] \qquad (B.12)$$

λαμβάνοντας υπόψη τους ακόλουθους περιορισμούς $\forall k \in [0,K-1]$

$$\mathbf{x}(k+1) = \mathbf{f}[\mathbf{x}(k), \mathbf{u}(k), k] \qquad (B.13)$$

$$\mathbf{h}[\mathbf{x}(k), \mathbf{u}(k), k] \leq \mathbf{0} \qquad (B.14)$$

καθώς και τον τελικό στόχο

$$\mathbf{g}[\mathbf{x}(K),K] = \mathbf{0}. \qquad (B.15)$$

Όλες οι επεξηγήσεις και ειδικές περιπτώσεις που δόθηκαν για το ΠΒΕ σε συνεχή χρόνο ισχύουν και στην περίπτωση διακριτού χρόνου.

Παράδειγμα Β.2: Θεωρούμε τον εισοδηματία του Παραδ. Α.2, συμπεριλαμβάνοντας όμως τις ανάγκες ιατρικής περίθαλψης στις γενικότερες προσωπικές ανάγκες, αντικαθιστούμε δηλαδή την (Α.9) με

$$y(k) = c[x(k) - u(k)]. \qquad (B.16)$$

Η καταστατική εξίσωση (Α.8) προφανώς αντιστοιχεί στην (Β.13). Το προς επένδυση ποσό u(k) κάθε χρονιάς είναι βέβαια μη αρνητικό. Επί πλέον, ο εισοδηματίας δεν επιθυμεί δανεισμό, άρα η επιτρεπτή περιοχή λήψης αποφάσεων καθορίζεται από

$$0 \le u(k) \le x(k) \qquad \text{(B.17)}$$

που προφανώς αντιστοιχεί στους περιορισμούς ανισοτήτων (Β.14).

Ο εισοδηματίας καταρτίζει το πλάνο του για μια δεκαετία (K=10) έχοντας ένα αρχικό κεφάλαιο x_0. Στο τέλος της δεκαετίας επιθυμεί να έχει κεφάλαιο $x(K)=x_e$. Το πρόβλημα βρίσκεται επομένως στην ειδική περίπτωση a1.

Δεδομένων αυτών, ο εισοδηματίας επιθυμεί να μεγιστοποιήσει την ποιότητα ζωής του κατά την διάρκεια της δεκαετίας και επιλέγει

$$\varphi[x(k), u(k)] = c[x(k) - u(k)]. \qquad \text{(B.18)}$$

Καταλήγουμε λοιπόν σε ένα ΠΒΕ όπου, δεδομένου του αρχικού κεφαλαίου x_0, ζητούμε την βέλτιστη πορεία επενδύσεων και κεφαλαίου u*(k), x*(k+1), k=0,...,9, έτσι ώστε να μεγιστοποιηθεί το κριτήριο κόστους (Β.12), όπως συγκεκριμενοποιήθηκε μέσω της ποιότητας ζωής (Β.18), λαμβάνοντας υπόψη την εξίσωση διαφορών (Α.8) που περιγράφει την μεταβολή του κεφαλαίου και την επιτρεπτή περιοχή χειρισμών (Β.17) και καταλήγοντας σε τελικό κεφάλαιο $x(K) = x_e$.

■

Ερωτήματα:

- Ο εισοδηματίας επιλέγει αντί της (Β.18) την εξής συνάρτηση κόστους

$$\varphi[x(k), u(k)] = k \cdot c[x(k) - u(k)]. \qquad \text{(B.19)}$$

  Τι επιδιώκει με την επιλογή αυτή;
- Ο εισοδηματίας θεωρεί ότι το ποσό που χρησιμοποιεί για προσωπικές ανάγκες δεν μπορεί σε καμιά περίπτωση να είναι μικρότερο του x_{min}. Πως μπορεί να συμπεριληφθεί η προδιαγραφή αυτή στο ΠΒΕ;
- Ο εισοδηματίας δεν καθορίζει τελικό κεφάλαιο $x(K) = x_e$, μεταβάλλει όμως το κριτήριο κόστους ως εξής

$$J = \alpha_1 x(10) + \alpha_2 \sum_{k=0}^{9} c[x(k) - u(k)] \qquad \text{(B.20)}$$

  με παραμέτρους $\alpha_1 \ge 0$, $\alpha_2 \ge 0$. Τι επιδιώκει με την επιλογή αυτή; Τι σημαίνει η επιλογή $\alpha_2 = 0$; Τι σημαίνει η επιλογή $\alpha_1 = 0$;

■

Το ΠΒΕ του Παραδ. Β.2 επιδιώκει μεγιστοποίηση αντί ελαχιστοποίηση του κριτηρίου κόστους. Το γεγονός αυτό δεν ενοχλεί, καθότι κάθε πρόβλημα μεγιστοποίησης μπορεί να αναχθεί άμεσα σε πρόβλημα ελαχιστοποίησης χρησιμοποιώντας την προφανή ισοτιμία

$$\max_{\mathbf{x}\in\mathscr{X}} J(\mathbf{x}) = -\min_{\mathbf{x}\in\mathscr{X}}[-J(\mathbf{x})] \qquad (B.21)$$

όπου $\mathscr{X}$ είναι η επιτρεπτή περιοχή της μεταβλητής **x**.

Να σημειώσουμε τέλος ότι η επιτρεπτή περιοχή ΠΒΕ μπορεί να καθορίζεται και κατ΄άλλους τρόπους πέραν των περιορισμών ανισότητας (Β.3) και (Β.14). Συγκεκριμένα μπορούν να λαμβάνονται υπόψη περιορισμοί ισότητας

$$\mathbf{G}[\mathbf{x}(t), \mathbf{u}(t), t] = \mathbf{0} \qquad (B.22)$$

διακριτές επιτρεπτές περιοχές

$$u_i(t) \in \mathscr{U}_i = \{u_{i,1}, u_{i,2}, ...\} \qquad (B.23)$$

καθώς και διάφορες άλλες δυνατότητες.

Β.2 Προσεγγίσεις στο Πρόβλημα Βέλτιστου Ελέγχου

Προβλήματα βέλτιστου ελέγχου μπορούν να επιλυθούν μέσω δύο εναλλακτικών μεθόδων. Η πρώτη μέθοδος είναι η **Αρχή του Μέγιστου** που αναπτύχθηκε από την ομάδα του σοβιετικού μαθηματικού **Λ.Σ. Ποντριάγκιν** το 1956 και στηρίζεται στη θεωρία του **Λογισμού Μεταβολών**[3]. Η δεύτερη μέθοδος, η οποία και εκτίθεται αναλυτικότερα στο κυρίως μέρος αυτού του συγγράμματος, είναι ο **Δυναμικός Προγραμματισμός** που αναπτύχθηκε από τον αμερικανό μαθηματικό **R. E. Bellman** στην αρχή της δεκαετίας του 1950. Κάθε μια από τις δύο μεθόδους έχει πλεονεκτήματα και μειονεκτήματα που αναφέρονται εκτενέστερα στο κύριο μέρος του συγγράμματος.

Ανεξάρτητα από την μέθοδο επίλυσης (Αρχή του Μέγιστου ή Δυναμικός Προγραμματισμός) τίθεται το ερώτημα, αν ένα συγκεκριμένο ΠΒΕ απαιτεί χρήση ή όχι υπολογιστή για την επίλυσή του. Η απάντηση στο ερώτημα αυτό είναι ότι μόνο απλά προβλήματα μικρής διάστασης (n=1 ή 2) μπορούν ενδεχομένως να επιλυθούν αναλυτικά, χωρίς χρήση υπολογιστή. Προβλήματα βέλτιστου ελέγχου με οποιεσδήποτε διαστάσεις αλλά απλή δομή (π.χ. γραμμικό σύστημα, χωρίς περιορισμούς ανισοτήτων, με τετραγωνικό κριτήριο κόστους) είναι επιλύσιμα με περιορισμένη χρήση υπολογιστή. Όλα τα άλλα ΠΒΕ απαιτούν για την επίλυσή τους χρήση υπολογιστή (αριθμητική επίλυση). Όταν πρόκειται για προβλήματα πολύ μεγάλων διαστάσεων, η επίλυση μπορεί να είναι προβληματική ή και αδύνατη, ακόμη και με χρήση υπολογιστή.

Στρέφουμε τώρα την προσοχή μας σε μια άλλη ενδιαφέρουσα πτυχή του ΠΒΕ. Όπως τέθηκε το πρόβλημα στο Κεφάλαιο Β.1, ζητούνται (εκτός ενδεχομένως του τελικού χρόνου t_e*) οι χρονικές εξελίξεις (τροχιές) των μεταβλητών ελέγχου $\mathbf{u}^*(t)$, $0 \le t \le t_e^*$, και των μεταβλητών κατάστασης $\mathbf{x}^*(t)$, $0 \le t \le t_e^*$, ως λύση του προβλήματος. Τι σημαίνει αυτό για την πρακτική εφαρμογή των αποτελεσμάτων; Σημαίνει πως ο βέλτιστος έλεγχος του οχήματος του Παραδ. Β.1, $u^*(t)$, $0 \le t \le t_e^*$, και ο βέλτιστος χρόνος t_e^*, που βέβαια εξαρτώνται από την αρχική κατάσταση $\mathbf{x}_0$, πρέπει να είναι γνωστοί πριν από τον χρόνο t = 0 ώστε να τεθούν σε εφαρμογή

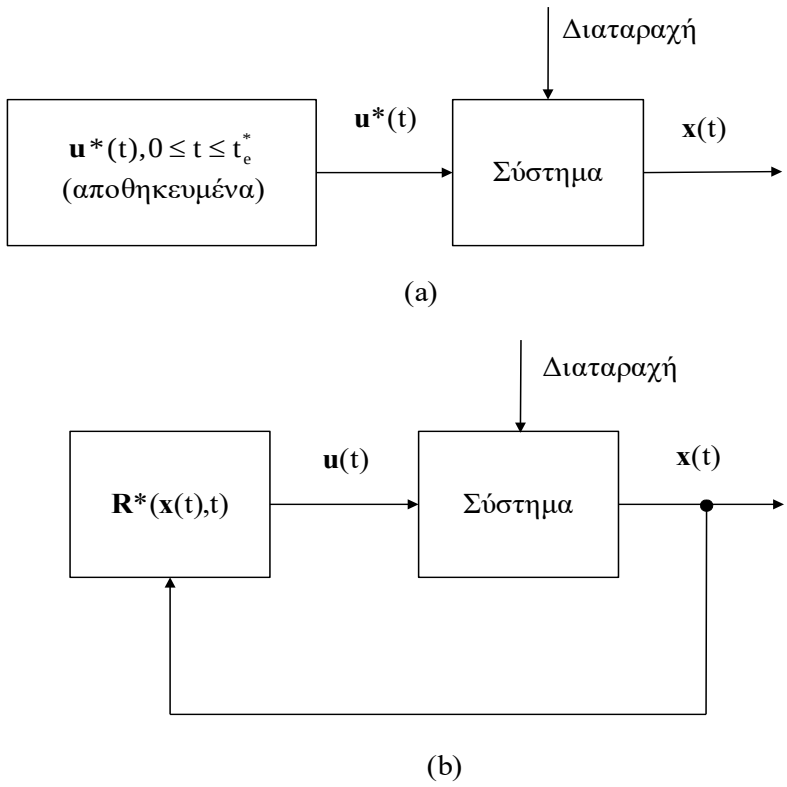
[3] Σημαντικό ρόλο στην ανάπτυξη της θεωρίας αυτής και στις εφαρμογές της σε προβλήματα δυναμικής βελτιστοποίησης έπαιξε ο μεγάλος ελληνικής καταγωγής μηχανικός και μαθηματικός **Κωνσταντίνος Καραθεοδωρή.**

κατά τη διάρκεια της πορείας του οχήματος. Η εφαρμογή του ελέγχου μπορεί να γίνει μέσω ενός υπολογιστή όπου έχει αποθηκευθεί η χρονική του εξέλιξη (βλ. Σχ. Β.5α). Αυτή είναι η περίπτωση **βέλτιστου ελέγχου ανοικτού βρόχου**. Το μειονέκτημα της πρακτικής αυτής οφείλεται στην πιθανότητα εμφάνισης απρόβλεπτων διαταραχών κατά την διάρκεια της εφαρμογής του ελέγχου που οδηγούν ενδεχομένως σε μη βέλτιστα αποτελέσματα. Το όχημα του Παραδ. Β.1 πιθανόν να υποστεί προς στιγμήν την ώθηση ισχυρότερου ανέμου απ΄ ό,τι είχε προβλεφθεί. Ο εισοδηματίας του Παραδ. Β.2 πιθανόν να κερδίσει απρόβλεπτα ένα ποσό στο λαχείο κατά τη διάρκεια $0 \le k \le K$. Όπως δείχνει το Σχ. Β.6a, οι αλλαγές αυτές μπορούν να οδηγήσουν σε αστοχία της τελικής τροχιάς και γενικά σε μη βέλτιστα αποτελέσματα.

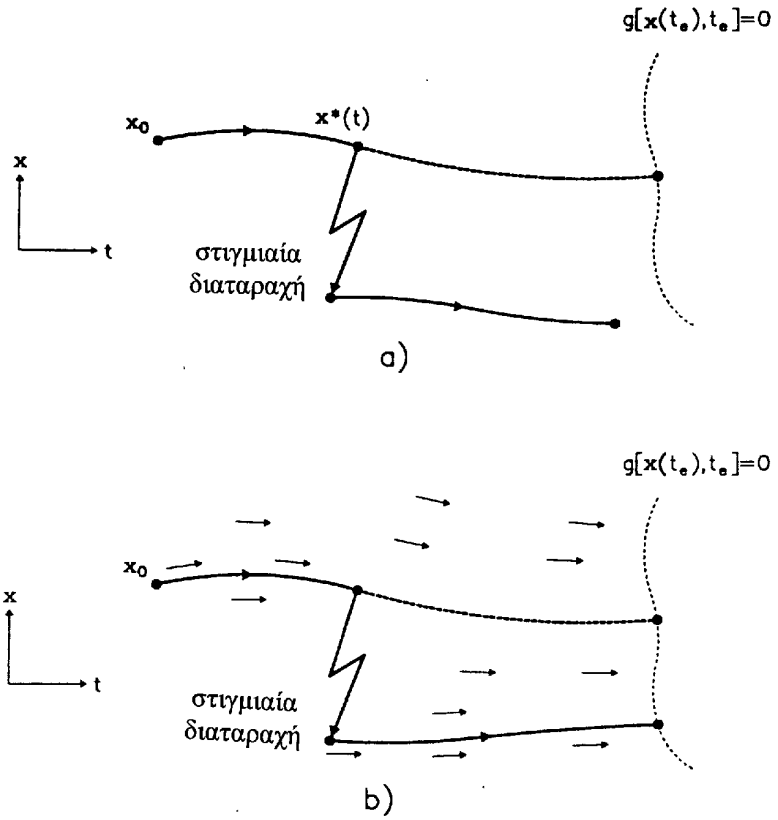
Η προβληματική αυτή μεταβάλλεται αν θεωρήσουμε αντί του ΠΒΕ το λεγόμενο **Πρόβλημα Σύνθεσης** Βέλτιστου Ελέγχου, στο οποίο ζητούμε ως λύση όχι την χρονική συνάρτηση ελέγχου $\mathbf{u}^*(t)$, $0 \le t \le t_e^*$, αλλά μια συνάρτηση $\mathbf{R}^*$ (**βέλτιστος κανόνας ελέγχου**), έτσι ώστε ο βέλτιστος έλεγχος να υπολογίζεται ανεξάρτητα από την αρχική κατάσταση $\mathbf{x}_0$ μέσω

$$\mathbf{u}^*(t) = \mathbf{R}^*[\mathbf{x}^*(t), t]. \qquad (B.24)$$

Η πρακτική εφαρμογή της λύσης αυτής γίνεται κατά το Σχ. Β.5b, δηλαδή ο υπολογισμός του βέλτιστου ελέγχου $\mathbf{u}^*(t)$ σε κάθε χρονική στιγμή γίνεται σε **πραγματικό χρόνο**, δηλαδή κατά τη διάρκεια της ελεγχόμενης διεργασίας, και βασίζεται σε **πραγματικές μετρήσεις** της κατάστασης $\mathbf{x}(t)$, οπότε και έχουμε **βέλτιστο έλεγχο κλειστού βρόχου**. Εφ΄ όσον δεν εμφανίζονται απρόβλεπτες διαταραχές, και αν το πρότυπο του συστήματος είναι ακριβές, ο βέλτιστος έλεγχος ανοικτού ή κλειστού βρόχου αποδίδει όμοια αποτελέσματα. Το Σχ. Β.6b δείχνει όμως ότι ο βέλτιστος έλεγχος κλειστού βρόχου αντιδρά καλύτερα υπό την επίδραση



Σχήμα Β.5: Βέλτιστος έλεγχος ανοιχτού βρόχου (a) και κλειστού βρόχου (b).



Σχήμα Β.6: Στιγμιαία διαταραχή σε βέλτιστο έλεγχο ανοικτού (a) και κλειστού (b) βρόχου.

στιγμιαίων διαταραχών. Συγκεκριμένα, η απόδοση του συστήματος είναι βέλτιστη μετά την επίδραση της διαταραχής και ο τελικός στόχος επιτυγχάνεται. Που οφείλεται αυτό; Οφείλεται στο γεγονός ότι η λύση του Προβλήματος Σύνθεσης είναι πολύ γενικότερη από την λύση του ΠΒΕ, καθότι είναι ανεξάρτητη από οποιαδήποτε συγκεκριμένη αρχική κατάσταση $\mathbf{x}_0$, αφού ορίζει ένα **πεδίο** στον χώρο ($\mathbf{x}$, t) που αναπαριστάται στο Σχ. Β.6b με μικρά βέλη. Έτσι, μετά την επίδραση της στιγμιαίας διαταραχής, ο έλεγχος ακολουθεί τα βέλη του πεδίου και το σύστημα καταλήγει κατά βέλτιστο τρόπο στον επιθυμητό τελικό στόχο.

Βέβαια, ο βέλτιστος έλεγχος κλειστού βρόχου αστοχεί και αυτός και δίδει μη βέλτιστα αποτελέσματα αν η διαταραχή δεν είναι στιγμιαία αλλά διαρκής. Τρόποι αντιμετώπισης τέτοιων καταστάσεων αναφέρονται στο κύριο μέρος του συγγράμματος. Στο σημείο αυτό αρκεί να συμπεράνουμε ότι ο βέλτιστος έλεγχος κλειστού βρόχου είναι καταλληλότερος για την αντιμετώπιση απρόβλεπτων διαταραχών ή σοβαρών ανακριβειών του προτύπου του συστήματος και επομένως προτιμότερος από τον βέλτιστο έλεγχο ανοικτού βρόχου. Επειδή όμως απαιτεί τη γενικότερη λύση του Προβλήματος Σύνθεσης, η ανάπτυξη συστημάτων βέλτιστου ελέγχου κλειστού βρόχου είναι δυσκολότερη, και για πολύπλοκα συστήματα συχνά ανέφικτη.

Β.3 Απόλυτο και Σχετικό Ελάχιστο

Κατά τη διατύπωση του ΠΒΕ αναφερθήκαμε στην ελαχιστοποίηση του κριτηρίου κόστους. Υπάρχουν διάφοροι ορισμοί του ελαχίστου μιας συνάρτησης J μιάς μεταβλητής $\mathbf{x}\in\mathscr{X}$, όπου $\mathscr{X}$ είναι η επιτρεπτή περιοχή βελτιστοποίησης.

Το **απόλυτο (ή γενικό) ελάχιστο** ορίζεται

$$J(\mathbf{x}^*) \le J(\mathbf{x}) \quad \forall \mathbf{x} \in \mathscr{X} \tag{B.25}$$

και ανταποκρίνεται στην συνήθη έννοια του ελάχιστου. Αν ισχύει

$$J(\mathbf{x}^*) < J(\mathbf{x}) \quad \forall \mathbf{x} \in \mathscr{X} - \{\mathbf{x}^*\} \tag{B.26}$$

έχουμε ένα **μοναδικό** απόλυτο ελάχιστο.

Το **σχετικό** (ή **τοπικό**) **ελάχιστο** ορίζεται μέσω $J(\mathbf{x}^*) \le J(\mathbf{x})$ για κάθε επιτρεπτό $\mathbf{x}$ από μια αρκούντως μικρή περιοχή περί το $\mathbf{x}^*$. Αν ισχύει $J(\mathbf{x}^*) < J(\mathbf{x})$ για κάθε επιτρεπτό $\mathbf{x} \ne \mathbf{x}^*$ από μια αρκούντως μικρή περιοχή περί το $\mathbf{x}^*$, τότε έχουμε ένα **αυστηρό σχετικό ελάχιστο**.

Παράδειγμα Β.3: Η συνάρτηση $J(x) = e^{-x}\sin x$, $x \ge 0$, έχει ένα μοναδικό απόλυτο ελάχιστο και άπειρα αυστηρά σχετικά ελάχιστα.

■

Ικανές συνθήκες για την ύπαρξη ενός απόλυτου ελάχιστου είναι οι εξής:
- Η επιτρεπτή περιοχή $\mathscr{X}$ είναι ένα σύνολο κλειστό, περιορισμένο και μη μηδενικό.
- Η συνάρτηση $J(\mathbf{x})$ είναι συνεχής.

Ας σημειώσουμε πως η ύπαρξη σχετικών ελαχίστων δεν σημαίνει υποχρεωτικά ότι ένα εξ αυτών είναι απόλυτο ελάχιστο. Για να ισχύει αυτό θα πρέπει να είναι βέβαιο ότι **υπάρχει** απόλυτο ελάχιστο.

Παράδειγμα Β.4: Η συνάρτηση $J(x) = (0.5 + x^2)e^{-x^2}$ έχει ένα σχετικό ελάχιστο στο σημείο $x^*=0$, δεν έχει όμως απόλυτο ελάχιστο.

■