ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΙΓΑΙΟΥ

ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ

ΤΜΗΜΑ ΣΤΑΤΙΣΤΙΚΗΣ ΚΑΙ

ΑΝΑΛΟΓΙΣΤΙΚΩΝ-ΧΡΗΜΑΤΟΟΙΚΟΝΟΜΙΚΩΝ ΜΑΘΗΜΑΤΙΚΩΝ

ΒΕΛΤΙΣΤΕΣ ΜΟΝΟΤΟΝΕΣ ΠΟΛΙΤΙΚΕΣ ΣΕ ΠΡΟΒΛΗΜΑΤΑ ΕΛΕΓΧΟΥ ΣΤΟΧΑΣΤΙΚΩΝ ΔΙΑΔΙΚΑΣΙΩΝ

Θεοδόσης Δ. Δημητράκος

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ

ΣΑΜΟΣ 2005

Στους γονείς μου και στα αδέλφια μου,

Βασιλική και Κώστα.

ΠΕΡΙΛΗΨΗ

Στην παρούσα διατριβή μελετώνται διάφορα προβλήματα βέλτιστου ελέγχου στοχαστικών διαδικασιών τα οποία μπορούν να περιγραφούν με κατάλληλα Μαρκοβιανά ή ημι-Μαρκοβιανά μοντέλα αποφάσεων. Τα προβλήματα σχετίζονται με το βέλτιστο έλεγχο μιας διδιάστατης επιδημικής διαδικασίας, με τη βέλτιστη προληπτική συντήρηση ενός συστήματος παραγωγής, με τη βέλτιστη επισκευή ή αντικατάσταση ενός μηχανήματος και με το βέλτιστο έλεγχο ενός πληθυσμού παρασίτων. Ο κύριος στόχος είναι η εύρεση της πολιτικής η οποία, για κάθε αρχική κατάσταση της διαδικασίας, ελαχιστοποιεί τη μέση τιμή μιας προκαθορισμένης συνάρτησης του μελλοντικού κόστους.

Σε μερικά προβλήματα αποδεικνύουμε ότι η βέλτιστη πολιτική είναι μονότονη, δηλαδή θέτει σε λειτουργία το μηχανισμό ελέγχου της διαδικασίας αν και μόνο αν η κατάσταση της διαδικασίας (π.χ. αριθμός φορέων μιας ασθένειας, βαθμός επιδείνωσης ή ηλικία ενός μηχανήματος, πληθυσμιακό μέγεθος παρασίτων) είναι ίση ή υπερβαίνει μία κρίσιμη τιμή. Σε κάποιες περιπτώσεις είναι δυνατόν να βρεθεί η βέλτιστη κρίσιμη τιμή.

Σε άλλα προβλήματα κατασκευάζουμε κατάλληλους αλγορίθμους οι οποίοι αποσκοπούν στην εύρεση της βέλτιστης πολιτικής. Σε ορισμένες περιπτώσεις αποδεικνύουμε ότι ο αλγόριθμος συγκλίνει στη βέλτιστη πολιτική ενώ σε άλλες περιπτώσεις υπάρχουν ισχυρές αριθμητικές ενδείξεις ότι η τελική πολιτική που δημιουργεί ο αλγόριθμος είναι βέλτιστη.

ΠΕΡΙΕΧΟΜΕΝΑ

ΠΕΡΙΛΗΨΗ 3

ΠΕΡΙΕΧΟΜΕΝΑ 4

ΕΥΧΑΡΙΣΤΙΕΣ 7

ΚΕΦΑΛΑΙΟ 1. Εισαγωγή 8

ΚΕΦΑΛΑΙΟ 2. Στοιχεία της θεωρίας των Μαρκοβιανών διαδικασιών αποφάσεων

2.1 Εισαγωγή 14
2.2 Μαρκοβιανές διαδικασίες αποφάσεων σε διακριτό χρόνο 15
2.3 Προβλήματα πεπερασμένου χρονικού ορίζοντα 17
2.4 Προβλήματα άπειρου χρονικού ορίζοντα 18
2.4.1 Ελαχιστοποίηση συνολικού αναμενόμενου αποπληθωρισμένου κόστους 18
2.4.2 Ελαχιστοποίηση μακροπρόθεσμου αναμενόμενου μέσου κόστους ανά μονάδα χρόνου 20
2.5 Ημι-Μαρκοβιανές διαδικασίες αποφάσεων και Μαρκοβιανές διαδικασίες αποφάσεων σε συνεχή χρόνο 25
2.6 Η προσεγγιστική μέθοδος της Sennott 31

ΚΕΦΑΛΑΙΟ 3. Βέλτιστος έλεγχος δύο ανταγωνιζομένων ασθενειών ή ειδών

3.1 Εισαγωγή 33
3.2 Αλγόριθμοι δυναμικού προγραμματισμού για τα Προβλήματα 1, 3 και 4 38
3.3 Ντετερμινιστική επιδημική διαδικασία 45
3.4 Αβεβαιότητα στις τιμές των παραμέτρων 56

ΚΕΦΑΛΑΙΟ 4. Βέλτιστη προληπτική συντήρηση ενός συστήματος παραγωγής

4.1 Εισαγωγή 60
4.2 Περιγραφή του μοντέλου 64

4.3 Η μορφή της βέλτιστης πολιτικής 69
4.4 Στάσιμη επιδείνωση του μηχανισμού 78
4.5 Ένας αλγόριθμος όταν οι χρόνοι προληπτικής συντήρησης και επισκευής ακολουθούν τη Γεωμετρική κατανομή 80
4.6 Δύο γενικεύσεις του μοντέλου 90
4.7 Οι χρόνοι προληπτικής συντήρησης και επισκευής είναι συνεχείς τυχαίες μεταβλητές 92
4.8 Ένας αλγόριθμος όταν οι χρόνοι προληπτικής συντήρησης και επισκευής είναι συνεχείς τυχαίες μεταβλητές 97

ΚΕΦΑΛΑΙΟ 5. Ένας αλγόριθμος για τον υπολογισμό της βέλτιστης πολιτικής επισκευής ή αντικατάστασης ενός συστήματος

5.1 Εισαγωγή 107
5.2 Κατασκευή του μοντέλου 109
5.3 Ο αλγόριθμος 113
5.4 Αριθμητικά παραδείγματα 120

ΚΕΦΑΛΑΙΟ 6. Υπολογισμός της βέλτιστης πολιτικής για τον έλεγχο μιας σύνθετης διαδικασίας μετανάστευσης με την εισαγωγή ολοκληρωτικών καταστροφών

6.1 Εισαγωγή 123
6.2 Η μορφή της βέλτιστης πολιτικής 126
6.3 Η μορφή της συνάρτησης του μέσου κόστους υπό τον έλεγχο μιας μονότονης πολιτικής 132
6.4 Ο υπολογισμός της βέλτιστης πολιτικής 148
6.5 Διωνυμικές καταστροφές 157

ΚΕΦΑΛΑΙΟ 7. Υπολογισμός της βέλτιστης πολιτικής για τον έλεγχο μιας απλής διαδικασίας μετανάστευσης με την εισαγωγή ενός αρπακτικού

7.1 Εισαγωγή 159
7.2 Ο υπολογισμός της βέλτιστης πολιτικής 162
7.3 Η σύγκλιση του αλγορίθμου 165

ΠΑΡΑΡΤΗΜΑ 169

ΒΙΒΛΙΟΓΡΑΦΙΑ 179

ΕΥΧΑΡΙΣΤΙΕΣ

Η παρούσα διδακτορική διατριβή εκπονήθηκε στο Τμήμα Στατιστικής και Αναλογιστικών-Χρηματοοικονομικών Μαθηματικών του Πανεπιστημίου Αιγαίου. Στο σημείο αυτό νιώθω την ανάγκη να εκφράσω τις ευχαριστίες μου στους ανθρώπους που συνετέλεσαν στην πραγματοποίησή της.

Ευχαριστώ θερμά τον κ. Αθανάσιο Γιαννακόπουλο, Αναπληρωτή Καθηγητή του Τμήματος Στατιστικής και Αναλογιστικών-Χρηματοοικονομικών Μαθηματικών του Πανεπιστημίου Αιγαίου, ο οποίος με τίμησε αποδεχόμενος την επίβλεψη της διατριβής. Θα ήθελα επίσης να τον ευχαριστήσω για το διαρκές ενδιαφέρον του και την υποστήριξή του.

Θέλω να εκφράσω την ευγνωμοσύνη μου στο δάσκαλό μου, κ. Επαμεινώνδα Κυριακίδη, Αναπληρωτή Καθηγητή του Τμήματος Μηχανικών Οικονομίας και Διοίκησης του Πανεπιστημίου Αιγαίου, ο οποίος στάθηκε δίπλα μου σε όλα τα στάδια αυτής της εργασίας, από την πρόταση του θέματός της μέχρι την παρουσίασή της. Χωρίς τις χρήσιμες συμβουλές του, τις υποδείξεις του και τη συνεχή συμπαράστασή του, θα ήταν αδύνατον να ολοκληρωθεί αυτή η διατριβή.

Θα ήθελα επίσης να ευχαριστήσω τα μέλη της επιτροπής κρίσης της διατριβής, που δέχτηκαν να αφιερώσουν κάποιο από τον πολύτιμο χρόνο τους για την αξιολόγηση της εργασίας.

Θέλω να ευχαριστήσω το διδακτικό και διοικητικό προσωπικό του Τμήματος Στατιστικής και Αναλογιστικών-Χρηματοοικονομικών Μαθηματικών του Πανεπιστημίου Αιγαίου για τη βοήθεια και την υποστήριξη που μου προσέφεραν σε όλη τη διάρκεια εκπόνησης της διατριβής.

Θα ήταν παράλειψή μου να μην ευχαριστήσω τον κ. Χρήστο Ευθυμιόπουλο, μέλος της ερευνητικής ομάδας του Τομέα Αστροφυσικής, Αστρονομίας και Μηχανικής του Τμήματος Φυσικής του Πανεπιστημίου Αθηνών καθώς και τον κ. Χρήστο Τσαγγάρη, που ανήκει στο Ειδικό Εργαστηριακό Διδακτικό Προσωπικό του Τμήματος Μαθηματικών του Πανεπιστημίου Αιγαίου, οι οποίοι με βοήθησαν στα πρώτα προβλήματα που συνάντησα στο υπολογιστικό μέρος της εργασίας.

Τέλος, θα ήθελα να ευχαριστήσω με όλη μου την καρδιά, τους γονείς μου, την αδελφή μου Βασιλική και τον αδελφό μου Κώστα, για την αγάπη τους και τη συμπαράστασή τους. Το λιγότερο που μπορώ να κάνω είναι να τους αφιερώσω αυτή τη διατριβή.

ΚΕΦΑΛΑΙΟ 1

Εισαγωγή

Θεωρούμε το πρόβλημα του βέλτιστου ελέγχου μιας στοχαστικής διαδικασίας. Ο έλεγχος της διαδικασίας πραγματοποιείται από έναν υποτιθέμενο ελεγκτή ο οποίος σε συγκεκριμένες χρονικές στιγμές επιλέγει μία ενέργεια από ένα σύνολο εναλλακτικών ενεργειών. Σε κάθε χρονική στιγμή ελέγχου της διαδικασίας, η επιλογή μιας ενέργειας επιφέρει ένα κόστος. Μία πολιτική είναι ένας κανόνας σύμφωνα με τον οποίον επιλέγονται οι ενέργειες σε κάθε χρονική στιγμή ελέγχου της διαδικασίας. Η βέλτιστη πολιτική είναι εκείνη η πολιτική η οποία ελαχιστοποιεί μία προκαθορισμένη συνάρτηση κόστους για κάθε αρχική κατάσταση της διαδικασίας.

Το πρόβλημα της εύρεσης της βέλτιστης πολιτικής παρουσιάζει ιδιαίτερο ενδιαφέρον και εμφανίζεται στη βιβλιογραφία με διάφορες μορφές. Αποτελεί το αντικείμενο έρευνας πολλών ερευνητών σε διάφορα πεδία της επιστήμης, όπως για παράδειγμα, στην Επιχειρησιακή Έρευνα, στην Οικολογία, στη Βιολογία και στην Πληροφορική. Ειδικότερα, η βέλτιστη πολιτική αναζητείται σε προβλήματα συντήρησης και αντικατάστασης μηχανημάτων, σε προβλήματα ελέγχου αποθεμάτων, σε προβλήματα ελέγχου ουρών αναμονής, σε προβλήματα ελέγχου βιολογικών πληθυσμών και σε προβλήματα διαχείρισης δικτύων και τηλεπικοινωνιών.

Σε πολλές περιπτώσεις, η μαθηματική διατύπωση, η ανάλυση και η επίλυση του προβλήματος είναι εφικτή μέσω της κατασκευής ενός κατάλληλου μοντέλου, που είναι γνωστό ως Μαρκοβιανό μοντέλο αποφάσεων. Το Μαρκοβιανό μοντέλο αποφάσεων επινοήθηκε από τον Bellman (1957) και είναι ένα κατάλληλο μαθηματικό μοντέλο που χρησιμοποιείται συχνά για την περιγραφή μιας στοχαστικής διαδικασίας η οποία μπορεί να ελεγχθεί από μία ακολουθία ενεργειών.

Σε πολλά προβλήματα Μαρκοβιανών μοντέλων αποφάσεων έχει αποδειχθεί ότι η βέλτιστη πολιτική είναι μονότονη, δηλαδή ο ελεγκτής επεμβαίνει στην εξέλιξη της διαδικασίας αν και μόνο αν η κατάσταση της διαδικασίας (π.χ. ο βαθμός επιδείνωσης ή η ηλικία ενός μηχανήματος, το πλήθος των πελατών σε μία ουρά αναμονής, το μέγεθος ενός βιολογικού πληθυσμού) είναι μεγαλύτερη ή ίση με μία κρίσιμη τιμή (βλέπε π.χ. Blackburn (1972), Kawai (1983), So (1992), Douer and Yechiali (1994), Van der Duyn Schouten and Vanneste (1995), Federgruen and So

(1989, 1990, 1991), Benyamini and Yechiali (1999), Kyriakidis (1999a, 2004)). Το γεγονός ότι μία μονότονη πολιτική είναι βέλτιστη επιταχύνει σημαντικά τον υπολογισμό της. Αναφέρουμε τις εργασίες των Abakuks (1979), Federgruen and So (1989) και Love et al. (2000) στις οποίες σχεδιάστηκαν αποδοτικοί αλγόριθμοι για τον υπολογισμό μιας βέλτιστης μονότονης πολιτικής.

Στην παρούσα διατριβή μελετώνται διάφορα προβλήματα βέλτιστου ελέγχου στοχαστικών διαδικασιών τα οποία μπορούν να περιγραφούν με τη χρήση κατάλληλων Μαρκοβιανών μοντέλων αποφάσεων. Αναζητείται η βέλτιστη πολιτική και αποδεικνύεται σε κάποια από αυτά τα προβλήματα ότι η βέλτιστη πολιτική είναι μονότονη. Σε άλλα προβλήματα, στα οποία φαίνεται δύσκολο να αποδειχθεί ότι η βέλτιστη πολιτική είναι μονότονη, κατασκευάζονται κατάλληλοι αλγόριθμοι για την εύρεση της βέλτιστης πολιτικής. Τα αριθμητικά αποτελέσματα των αλγορίθμων παρέχουν ισχυρή ένδειξη ότι η βέλτιστη πολιτική είναι μονότονη.

Στο Κεφάλαιο 2 παρουσιάζονται με συνοπτικό τρόπο τα βασικά στοιχεία της θεωρίας των Μαρκοβιανών μοντέλων αποφάσεων τα περισσότερα των οποίων θα χρησιμοποιηθούν στα επόμενα κεφάλαια. Η παρουσίαση βασίζεται στα διάφορα κριτήρια βελτιστοποίησης που καθορίζονται από την επιλογή της συνάρτησης του κόστους. Ιδιαίτερη έμφαση δίνεται στο κριτήριο της ελαχιστοποίησης του μακροπρόθεσμου αναμενόμενου μέσου κόστους ανά μονάδα χρόνου που χρησιμοποιείται σε όλα τα κεφάλαια της διατριβής εκτός από το Κεφάλαιο 3 στο οποίο χρησιμοποιείται το κριτήριο της ελαχιστοποίησης του συνολικού αναμενόμενου κόστους.

Στο Κεφάλαιο 3 μελετάται το πρόβλημα του βέλτιστου ελέγχου μιας διδιάστατης επιδημικής διαδικασίας. Υποτίθεται ότι ένας πληθυσμός ατόμων είναι δυνατόν να προσβληθεί από δύο μεταδοτικές ασθένειες, την ασθένεια 1 και την ασθένεια 2. Θεωρούμε ότι η ασθένεια 1 είναι μία σοβαρή ασθένεια και ότι η ασθένεια 2 είναι μία ήπια ασθένεια. Η παρουσία ενός ατόμου που έχει προσβληθεί από τη σοβαρή ασθένεια 1 επιφέρει κάποιο κόστος ενώ η παρουσία ενός ατόμου που έχει προσβληθεί από την ήπια ασθένεια 2 δεν επιφέρει κάποιο κόστος. Θεωρούνται πολιτικές, οι οποίες εμβολιάζουν με την ήπια ασθένεια 2 τα επιδεκτικά άτομα που έχουν απομείνει στον πληθυσμό και δεν έχουν προσβληθεί από καμία από τις δύο ασθένειες, ή απομονώνουν τα άτομα που έχουν προσβληθεί από τη σοβαρή ασθένεια 1. Υποτίθεται ότι ο εμβολιασμός των επιδεκτικών και η απομόνωση των προσβληθέντων ατόμων επιφέρουν αντίστοιχα κόστη. Γενικεύονται τα μοντέλα που επινοήθηκαν από τον Kyriakidis (1995, 1999c). Ορίζεται μία στοχαστική επιδημική διαδικασία στην οποία οι ρυθμοί προσβολής εξαρτώνται από μία δύναμη του αριθμού των προσβληθέντων ατόμων. Κατασκευάζονται κατάλληλοι αλγόριθμοι του δυναμικού προγραμματισμού για τον αριθμητικό υπολογισμό της βέλτιστης πολιτικής. Για την αντίστοιχη ντετερμινιστική επιδημική διαδικασία η μορφή της βέλτιστης πολιτικής αποδεικνύεται αναλυτικά σε δύο περιπτώσεις και συγκρίνεται αριθμητικά με τη βέλτιστη πολιτική της στοχαστικής επιδημικής διαδικασίας. Επίσης παρουσιάζεται μία τροποποίηση της στοχαστικής επιδημικής διαδικασίας στην οποία θεωρείται ότι το πηλίκο δύο παραμέτρων είναι μία τυχαία μεταβλητή που ακολουθεί μία γνωστή κατανομή. Η βέλτιστη πολιτική υπολογίζεται αριθμητικά για την τροποποιημένη διαδικασία και συγκρίνεται με την αντίστοιχη βέλτιστη πολιτική της στοχαστικής επιδημικής διαδικασίας.

Στο Κεφάλαιο 4 προτείνεται μία γενίκευση του μοντέλου που επινοήθηκε από τους Van der Duyn Schouten and Vanneste (1995). Υποτίθεται ότι ένα σύστημα παραγωγής αποτελείται από ένα μηχανισμό τροφοδοσίας, μία μονάδα παραγωγής και έναν ενδιάμεσο αποθηκευτικό χώρο. Ο μηχανισμός τροφοδοτεί τη μονάδα παραγωγής με ένα ακατέργαστο υλικό. Στο μοντέλο των Van der Duyn Schouten and Vanneste θεωρείται ότι η επιδείνωση του μηχανισμού είναι στάσιμη υπό την έννοια ότι οι πιθανότητες μετάβασης εξαρτώνται μόνο από το βαθμό επιδείνωσης του μηχανισμού. Στο προτεινόμενο μοντέλο θεωρείται ότι η επιδείνωση του μηχανισμού δεν είναι στάσιμη διότι εξαρτάται από το βαθμό επιδείνωσης και από την ηλικία του μηχανισμού. Υποτίθεται ότι η λειτουργία του μηχανισμού, μία προληπτική συντήρησή του ή μία επισκευή του επιφέρουν αντίστοιχα κόστη για κάθε μονάδα χρόνου κατά την οποία ο μηχανισμός λειτουργεί, συντηρείται προληπτικά ή επισκευάζεται. Υποτίθεται επίσης ότι υπάρχει ένα κόστος αποθήκευσης του ακατέργαστου υλικού στον αποθηκευτικό χώρο και ένα κόστος όταν ο αποθηκευτικός χώρος είναι κενός. Ορίζονται κατάλληλες συνθήκες οι οποίες αφορούν τα κόστη της λειτουργίας, τα κόστη της προληπτικής συντήρησης, τα κόστη της επισκευής, τις πιθανότητες μετάβασης, τους αναμενόμενους χρόνους της προληπτικής συντήρησης και τους αναμενόμενους χρόνους της επισκευής του μηχανισμού. Αποδεικνύεται αναλυτικά ότι η βέλτιστη πολιτική είναι μονότονη διότι, για σταθερό περιεχόμενο του αποθηκευτικού χώρου και σταθερή ηλικία του μηχανισμού, θέτει σε λειτουργία μία προληπτική συντήρηση αν και μόνο αν ο βαθμός επιδείνωσης του μηχανισμού είναι μεγαλύτερος ή ίσος με μία κρίσιμη τιμή. Στην περίπτωση της στάσιμης επιδείνωσης του μηχανισμού επισημαίνεται ότι η βέλτιστη πολιτική είναι επίσης μονότονη για σταθερό περιεχόμενο του αποθηκευτικού χώρου. Σχεδιάζεται ένας αποδοτικός αλγόριθμος βελτίωσης των πολιτικών για την περίπτωση αυτή ο οποίος παράγει μία ακολουθία βελτιωμένων πολιτικών που έχουν τη μονότονη μορφή. Υπάρχει ισχυρή ένδειξη, βάσει αριθμητικών αποτελεσμάτων, ότι η τελική πολιτική που παράγει ο αλγόριθμος είναι βέλτιστη. Μελετώνται ακόμα δύο γενικεύσεις του στάσιμου μοντέλου στις οποίες υπολογίζεται αριθμητικά η βέλτιστη πολιτική. Παρουσιάζεται επίσης μία τροποποίηση του στάσιμου μοντέλου στην οποία υποθέτουμε ότι οι απαιτούμενοι χρόνοι για μία προληπτική συντήρηση και μία επισκευή του μηχανισμού είναι συνεχείς τυχαίες μεταβλητές. Σχεδιάζεται ένας αποδοτικός αλγόριθμος βελτίωσης των πολιτικών για το τροποποιημένο μοντέλο ο οποίος εφαρμόζεται στο σύνολο των μονότονων πολιτικών και παράγει μία ακολουθία βελτιωμένων πολιτικών που έχουν τη μονότονη μορφή. Υπάρχει πάλι ισχυρή ένδειξη ότι η τελική πολιτική που παράγει ο αλγόριθμος είναι βέλτιστη.

Στο Κεφάλαιο 5 μελετάται ένα μοντέλο που επινοήθηκε από τους Kijima et al. (1988) και γενικεύτηκε από τους Makis and Jardine (1993). Υποτίθεται ότι ένα σύστημα επιδεινώνεται με την πάροδο του χρόνου και ότι η λειτουργία του διακόπτεται εξαιτίας ενδεχόμενων βλαβών. Υποτίθεται επίσης ότι, όταν το σύστημα έχει υποστεί μία βλάβη, μπορεί να επισκευαστεί ή να αντικατασταθεί από ένα καινούργιο σύστημα. Η επισκευή και η αντικατάσταση του συστήματος επιφέρουν αντίστοιχα κόστη. Οι Makis and Jardine όρισαν κατάλληλες συνθήκες οι οποίες αφορούν τα κόστη της επισκευής, τα κόστη της αντικατάστασης και το ρυθμό επιδείνωσης του συστήματος. Απέδειξαν ότι η βέλτιστη πολιτική είναι μονότονη διότι αντικαθιστά το σύστημα όταν έχει υποστεί την $n-$οστή βλάβη αν και μόνο αν η ηλικία του είναι μεγαλύτερη ή ίση με μία κρίσιμη τιμή η οποία εξαρτάται από τον αριθμό n. Σχεδιάζουμε έναν αποδοτικό αλγόριθμο βελτίωσης των πολιτικών ο οποίος εφαρμόζεται στο σύνολο των μονότονων πολιτικών και παράγει μία ακολουθία βελτιωμένων πολιτικών που έχουν τη μονότονη μορφή. Ο αλγόριθμος είναι κατά πολύ ταχύτερος ενός παρόμοιου αλγορίθμου που αναπτύχθηκε από τους Love et al. (2000). Πολλά αριθμητικά παραδείγματα παρέχουν ισχυρή ένδειξη ότι ο αλγόριθμος συγκλίνει στη βέλτιστη πολιτική.

Στο Κεφάλαιο 6 επεκτείνεται ένα μοντέλο που επινοήθηκε από τον Economou (2003). Ο Economou υπέθεσε ότι ένας πληθυσμός παρασίτων αναπτύσσεται στοχαστικά σύμφωνα με μία σύνθετη διαδικασία Poisson και μπορεί να ελεγχθεί από ένα μηχανισμό ο οποίος, όταν τίθεται σε λειτουργία, καταστρέφει ολοκληρωτικά τον πληθυσμό. Υπέθεσε επίσης ότι ο ρυθμός του κόστους για τη λειτουργία του μηχανισμού της καταστροφής είναι σταθερός και ότι ο ρυθμός του κόστους που προξενούν τα παράσιτα είναι μία αύξουσα συνάρτηση ως προς το πληθυσμιακό τους μέγεθος. Χρησιμοποίησε την τεχνική της ομοιομορφοποίησης (βλέπε π.χ. Sennott (1999)) και απέδειξε ότι, όταν τα κόστη είναι άνω φραγμένα, η βέλτιστη πολιτική είναι μονότονη διότι θέτει σε λειτουργία το μηχανισμό της καταστροφής αν και μόνο αν το μέγεθος του πληθυσμού των παρασίτων είναι μεγαλύτερο ή ίσο με μία κρίσιμη τιμή. Στην παρούσα διατριβή χρησιμοποιούμε τη μέθοδο των διαδοχικών προσεγγίσεων (βλέπε π.χ. Ross (1992)) και δίνουμε μία διαφορετική απόδειξη του αποτελέσματος του Economou το οποίο γενικεύεται για την περίπτωση κατά την οποία τα κόστη δεν είναι άνω φραγμένα. Αποδεικνύουμε επίσης ότι το μακροπρόθεσμο αναμενόμενο μέσο κόστος ανά μονάδα χρόνου, υπό τον έλεγχο μιας μονότονης πολιτικής, είναι μία μονοκόρυφη συνάρτηση ως προς την κρίσιμη τιμή. Το αποτέλεσμα αυτό επιτρέπει την εφαρμογή δύο αποδοτικών αλγορίθμων για τον αριθμητικό υπολογισμό της βέλτιστης μονότονης πολιτικής, δηλαδή της μεθόδου της διχοτόμησης και ενός κατάλληλου αλγορίθμου που παράγει βελτιωμένες μονότονες πολιτικές. Προτείνεται επίσης μία μέθοδος για τον ακριβή υπολογισμό των στάσιμων πιθανοτήτων, υπό τον έλεγχο μιας μονότονης πολιτικής. Επιπλέον μελετάμε ένα γενικότερο μοντέλο στο οποίο, όταν ο μηχανισμός τίθεται σε λειτουργία, προξενεί μία διωνυμική καταστροφή αντί μιας ολοκληρωτικής καταστροφής του πληθυσμού των παρασίτων. Κατασκευάζεται ένας κατάλληλος αλγόριθμος για τον αριθμητικό υπολογισμό της βέλτιστης πολιτικής ο οποίος βασίζεται στην προσεγγιστική μέθοδο της Sennott (1997). Πολλά παραδείγματα παρέχουν ισχυρή ένδειξη ότι η βέλτιστη πολιτική, για το μοντέλο με τις διωνυμικές καταστροφές, είναι επίσης μονότονη.

Στο Κεφάλαιο 7 προτείνεται μία τροποποίηση του μοντέλου που επινοήθηκε από τον Kyriakidis (2003). Ο Kyriakidis υπέθεσε ότι ένας πληθυσμός παρασίτων αναπτύσσεται στοχαστικά σε ένα βιότοπο σύμφωνα με μία απλή διαδικασία Poisson και μπορεί να ελεγχθεί με την εισαγωγή ενός αρπακτικού. Υπέθεσε επίσης ότι ο ρυθμός του κόστους για την εισαγωγή του αρπακτικού στο βιότοπο είναι σταθερός και ότι ο ρυθμός του κόστους που προξενούν τα παράσιτα είναι μία αύξουσα συνάρτηση ως προς το πληθυσμιακό τους μέγεθος. Θεώρησε ότι το αρπακτικό μπορεί να αποδημήσει από το βιότοπο μόνο όταν έχει εξοντώσει όλα τα παράσιτα. Στο προτεινόμενο μοντέλο υποτίθεται ότι το αρπακτικό μπορεί να αποδημήσει από το βιότοπο πριν εξοντώσει όλα τα παράσιτα. Κατασκευάζεται ένας κατάλληλος αλγόριθμος για τον αριθμητικό υπολογισμό της βέλτιστης πολιτικής ο οποίος βασίζεται στην προσεγγιστική μέθοδο της Sennott (1997). Επαληθεύονται οι συνθήκες που εγγυώνται τη σύγκλιση του αλγορίθμου στη βέλτιστη πολιτική. Πολλά παραδείγματα παρέχουν ισχυρή ένδειξη ότι η βέλτιστη πολιτική είναι μονότονη διότι εισάγει το αρπακτικό στο βιότοπο αν και μόνο αν το μέγεθος του πληθυσμού των παρασίτων είναι μεγαλύτερο ή ίσο με μία κρίσιμη τιμή.

Σημειώνεται ότι οι αλγόριθμοι της παρούσας διατριβής υλοποιήθηκαν με χρήση του μαθηματικού πακέτου Matlab και εκτελέστηκαν σε υπολογιστή τύπου Acer Aspire 1605DLC.

ΚΕΦΑΛΑΙΟ 2

Στοιχεία της θεωρίας των Μαρκοβιανών διαδικασιών αποφάσεων

2.1 Εισαγωγή

Στις αρχές του $20^{ου}$ αιώνα ο Ρώσος Μαθηματικός A. A. Markov στην προσπάθειά του να ερμηνεύσει την «αβεβαιότητα» στην εναλλαγή φωνηέντων και συμφώνων γραμμάτων στο ποίημα «Onegin» του Pushkin εισήγαγε τη θεωρία των *Μαρκοβιανών διαδικασιών*.

Ο Bellman (1957) εισήγαγε τη θεωρία του *δυναμικού προγραμματισμού*. Ανέπτυξε μία αναδρομική διαδικασία η οποία υπολογίζει βέλτιστες τιμές συναρτήσεων κέρδους ή κόστους μέσω μιας κατάλληλης συναρτησιακής εξίσωσης. Ο δυναμικός προγραμματισμός χρησιμοποιείται σε προβλήματα πεπερασμένου ή άπειρου χρονικού ορίζοντα στα οποία μία στοχαστική διαδικασία ελέγχεται από μία ακολουθία ενεργειών. Ο κύριος στόχος είναι η εύρεση ενός κανόνα επιλογής των ενεργειών που ελέγχει τη διαδικασία με το βέλτιστο τρόπο. Ο Howard (1960) συνδύασε ιδέες του δυναμικού προγραμματισμού με στοιχεία της θεωρίας των στοχαστικών διαδικασιών και κατασκεύασε έναν αλγόριθμο βελτίωσης των πολιτικών για να υπολογίσει τη βέλτιστη πολιτική σε προβλήματα ελέγχου διαδικασιών σε άπειρο χρονικό ορίζοντα.

Οι *Μαρκοβιανές διαδικασίες αποφάσεων* εισήχθησαν από τον Bellman (1957) και ήταν αποτέλεσμα του συνδυασμού της θεωρίας των Μαρκοβιανών διαδικασιών και του δυναμικού προγραμματισμού. Κατά τη διάρκεια των τελευταίων τεσσάρων δεκαετιών αποτελούν το αντικείμενο έρευνας πολλών ερευνητών. Έχουν βρει εφαρμογή σε διάφορα πεδία της επιστήμης, όπως για παράδειγμα στην Επιχειρησιακή Έρευνα, στη Βιολογία, στην Οικολογία και στην Πληροφορική. Ειδικότερα, έχουν αποδειχθεί πολύ χρήσιμες σε προβλήματα βέλτιστου ελέγχου αποθεμάτων, βέλτιστου ελέγχου ουρών αναμονής και βιολογικών πληθυσμών, βέλτιστης συντήρησης και αντικατάστασης μηχανημάτων, βέλτιστης διαχείρισης δικτύων και τηλεπικοινωνιών.

Στα βιβλία των Ross (1983, 1992), Puterman (1994), Sennott (1999) και Bather (2000) παρουσιάζονται με λεπτομέρεια βασικά αποτελέσματα της θεωρίας των Μαρκοβιανών διαδικασιών αποφάσεων. Οι αλγόριθμοι του δυναμικού προγραμματισμού και οι βελτιώσεις τους παρουσιάζονται λεπτομερώς με διάφορες εφαρμογές τους στα βιβλία των Puterman (1994), Tijms (1994) και Heyman and Sobel (2004).

Στο παρόν κεφάλαιο θα παρουσιάσουμε με συνοπτικό τρόπο στοιχεία της θεωρίας των Μαρκοβιανών διαδικασιών αποφάσεων, τα περισσότερα των οποίων θα χρησιμοποιηθούν στα επόμενα κεφάλαια. Στο Εδάφιο 2.2 περιγράφουμε τις Μαρκοβιανές διαδικασίες αποφάσεων σε διακριτό χρόνο και εισάγουμε βασικές έννοιες. Στα Εδάφια 2.3 και 2.4 αναφέρουμε γνωστά αποτελέσματα για τα προβλήματα του πεπερασμένου και του άπειρου χρονικού ορίζοντα, αντίστοιχα. Στο Εδάφιο 2.5 παρουσιάζουμε τις ημι-Μαρκοβιανές διαδικασίες αποφάσεων και μία υποκατηγορία τους, τις Μαρκοβιανές διαδικασίες αποφάσεων σε συνεχή χρόνο. Στο Εδάφιο 2.6 περιγράφουμε με συνοπτικό τρόπο την προσεγγιστική μέθοδο της Sennott (1997).

2.2 Μαρκοβιανές διαδικασίες αποφάσεων σε διακριτό χρόνο

Έστω μία στοχαστική διαδικασία $X_n, n = 0, 1, 2, ...$ όπου η τυχαία μεταβλητή X_n αναπαριστά την κατάσταση ενός συστήματος τη χρονική στιγμή n. Το σύνολο των καταστάσεων του συστήματος είναι πεπερασμένο ή άπειρο αριθμήσιμο. Χάριν απλότητας, στο παρόν κεφάλαιο μπορούμε να υποθέσουμε ότι είναι το σύνολο των μη-αρνητικών ακεραίων αριθμών $0, 1, 2, ...$ Το σύστημα επιθεωρείται τις χρονικές στιγμές $n = 0, 1, 2, ...$ οι οποίες θεωρούμε ότι ισαπέχουν μεταξύ τους. Η κατάσταση του συστήματος παρατηρείται σε κάθε χρονική στιγμή επιθεώρησης και μία ενέργεια επιλέγεται από ένα σύνολο εναλλακτικών ενεργειών. Έστω ότι σε κάποια χρονική στιγμή επιθεώρησης n το σύστημα βρίσκεται στην κατάσταση i και η ενέργεια a επιλέγεται από ένα σύνολο εναλλακτικών ενεργειών $A(i)$. Υποθέτουμε ότι το σύνολο $A(i)$ είναι πεπερασμένο.

Το σύστημα που περιγράψαμε παραπάνω είναι μία *Μαρκοβιανή διαδικασία αποφάσεων σε διακριτό χρόνο*, αν:

(α) Υπάρχει ένα κόστος $C(i,a)$ το οποίο εξαρτάται μόνον από την κατάσταση i και την ενέργεια a ως οικονομική συνέπεια της επιλογής της ενέργειας a τη χρονική στιγμή κατά την οποία το σύστημα βρίσκεται στην κατάσταση i.
(β) Την επόμενη χρονική στιγμή η κατάσταση του συστήματος είναι η κατάσταση j με πιθανότητα $p_{ij}(a)$ η οποία εξαρτάται μόνον από την ενέργεια a και τις καταστάσεις i και j.

Ο όρος «Μαρκοβιανή» δικαιολογείται από το γεγονός ότι το κόστος $C(i,a)$ και η πιθανότητα μετάβασης $p_{ij}(a)$ εξαρτώνται από το «παρελθόν» της διαδικασίας μόνο μέσω της τρέχουσας κατάστασης i της διαδικασίας και της ενέργειας a που επιλέγεται στην κατάσταση i.

Μία *πολιτική* π είναι ένας κανόνας με τον οποίον επιλέγονται οι ενέργειες κατά τις χρονικές στιγμές $n = 0, 1, 2, \ldots$. Έστω $\Im$ το σύνολο όλων των πολιτικών. Υπάρχουν διάφορα είδη πολιτικών (βλέπε σελ. 22 του βιβλίου του Puterman (1994)). Η ταξινόμησή τους εξαρτάται από το αν είναι ή όχι «τυχαιοποιημένες» καθώς και από το αν εξαρτώνται από την «ιστορία» της διαδικασίας. Με τον όρο «τυχαιοποιημένη» θεωρούμε εκείνη την πολιτική σύμφωνα με την οποία, όταν η διαδικασία βρίσκεται στην κατάσταση i, μία ενέργεια a επιλέγεται με πιθανότητα P_a, $a \in A(i)$, σε κάποια χρονική στιγμή επιλογής των ενεργειών. Στην παρούσα διατριβή θα μας απασχολήσουν οι Μαρκοβιανές πολιτικές καθώς επίσης και μία σημαντική υποκατηγορία τους, οι στάσιμες πολιτικές.

Μία *Μαρκοβιανή πολιτική* είναι μία πολιτική σύμφωνα με την οποία η επιλογή μιας ενέργειας σε κάθε χρονική στιγμή $n = 0, 1, 2, \ldots$ εξαρτάται μόνον από τη χρονική στιγμή n και από την κατάσταση της διαδικασίας σ' αυτή τη χρονική στιγμή.

Μία *στάσιμη πολιτική* είναι μία πολιτική σύμφωνα με την οποία η επιλογή μιας ενέργειας σε κάθε χρονική στιγμή $n = 0, 1, 2, \ldots$ εξαρτάται μόνον από την κατάσταση της διαδικασίας σ' αυτή τη χρονική στιγμή. Επομένως μία στάσιμη πολιτική f καθορίζεται πλήρως από μία ακολουθία $\{f_i\}$, $i = 0, 1, 2, \ldots$ όπου $f_i \in A(i)$ είναι η ενέργεια που επιλέγεται οποτεδήποτε η διαδικασία βρίσκεται στην κατάσταση i σε μία χρονική στιγμή επιλογής ενέργειας.

Στη γενική του μορφή το πρόβλημα που θα μας απασχολήσει είναι η εύρεση της πολιτικής η οποία, για κάθε αρχική κατάσταση της διαδικασίας, ελαχιστοποιεί μία προκαθορισμένη συνάρτηση κόστους.

Η συνάρτηση του κόστους ορίζει το *κριτήριο βελτιστοποίησης* του προβλήματος. Τα κριτήρια βελτιστοποίησης τα οποία χρησιμοποιούνται πιο συχνά είναι η ελαχιστοποίηση του συνολικού αναμενόμενου αποπληθωρισμένου κόστους και η ελαχιστοποίηση του μακροπρόθεσμου αναμενόμενου μέσου κόστους ανά μονάδα χρόνου. Επίσης θα μας απασχολήσει το κριτήριο της ελαχιστοποίησης του συνολικού αναμενόμενου κόστους.

Αν υπάρχει ένας σταθερός ακέραιος αριθμός $N \geq 1$ τέτοιος ώστε οι ενέργειες για τον έλεγχο μιας Μαρκοβιανής διαδικασίας αποφάσεων σε διακριτό χρόνο επιλέγονται τις χρονικές στιγμές $0, 1, 2, \ldots, N-1$ και η διαδικασία σταματά τη χρονική στιγμή N, τότε το πρόβλημα είναι *πεπερασμένου χρονικού ορίζοντα* N *βημάτων*. Διαφορετικά, αν το σύνολο των χρονικών στιγμών επιλογής ενέργειας είναι άπειρο, το πρόβλημα είναι *άπειρου χρονικού ορίζοντα*.

2.3 Προβλήματα πεπερασμένου χρονικού ορίζοντα

Έστω a_j η ενέργεια που επιλέγεται τη χρονική στιγμή j. Το συνολικό αναμενόμενο αποπληθωρισμένο κόστος $V_\alpha(i, N, \pi)$ όταν απομένουν N βήματα μέχρι το τερματισμό της διαδικασίας, υπό τον έλεγχο της πολιτικής π, δοθέντος ότι η αρχική κατάσταση της διαδικασίας είναι η κατάσταση i, ορίζεται ως εξής:

$$V_\alpha(i, N, \pi) = E_\pi\left[\sum_{j=0}^{N-1} \alpha^j C(X_j, a_j) + \alpha^N F(X_N) \mid X_0 = i\right],$$

όπου, E_π αναπαριστά την υπό-συνθήκη αναμενόμενη τιμή, δοθέντος ότι η πολιτική π έχει υιοθετηθεί για τον έλεγχο της διαδικασίας. Η συνάρτηση $F(\cdot)$ είναι μία γνωστή μη-αρνητική συνάρτηση κόστους η οποία ορίζεται στο χώρο καταστάσεων της διαδικασίας και αντιπροσωπεύει ένα τελικό κόστος $F(X_N)$ όταν η διαδικασία σταματά τη χρονική στιγμή N. Η σταθερά $\alpha \in (0, 1]$ είναι ο *αποπληθωριστικός παράγοντας*.

Για κάθε κατάσταση i της διαδικασίας, η *βέλτιστη συνάρτηση* $V_\alpha(i, N)$ *του αναμενόμενου αποπληθωρισμένου κόστους όταν απομένουν* N *βήματα μέχρι το τερματισμό της διαδικασίας*, ορίζεται ως εξής:

$$V_\alpha(i, N) = \inf_{\pi \in \Im} V_\alpha(i, N, \pi), \quad i \geq 0.$$

Το παρακάτω θεώρημα παρέχει μία εξίσωση που υπολογίζει αναδρομικά τη βέλτιστη συνάρτηση για κάθε κατάσταση της διαδικασίας.

Θεώρημα 2.1. Η ποσότητα $V_\alpha(i,N),\ i \geq 0,$ ικανοποιεί για κάθε $N = 1, 2, \ldots$ την εξίσωση:

$$V_\alpha(i,N) = \min_{a \in A(i)} \left\{ C(i,a) + \alpha \sum_{j=0}^{\infty} p_{ij}(a) V_\alpha(j, N-1) \right\}, \tag{2.1}$$

όπου, $V_\alpha(i,0) = F(i).$

Επιπλέον υπάρχει μία βέλτιστη Μαρκοβιανή πολιτική η οποία, οποτεδήποτε η διαδικασία βρίσκεται στην κατάσταση $i,$ επιλέγει εκείνη την ενέργεια a που ελαχιστοποιεί το δεξιό μέλος της εξίσωσης (2.1).

Η εξίσωση (2.1) είναι γνωστή ως εξίσωση βελτιστοποίησης για το πρόβλημα της ελαχιστοποίησης του συνολικού αναμενόμενου αποπληθωρισμένου κόστους σε πεπερασμένο χρονικό ορίζοντα. Το Θεώρημα 2.1 μπορεί να αποδειχθεί με επαγωγή ως προς N και η απόδειξη βρίσκεται στις σελ. 37-39 του βιβλίου της Sennott (1999).

Στην περίπτωση κατά την οποία ο αποπληθωριστικός παράγοντας α είναι ίσος με τη μονάδα έχουμε το πρόβλημα της ελαχιστοποίησης του συνολικού αναμενόμενου κόστους σε πεπερασμένο χρονικό ορίζοντα. Για το πρόβλημα αυτό οι προαναφερθέντες ορισμοί και το Θεώρημα 2.1 ισχύουν, αν θέσουμε $\alpha = 1.$

2.4 Προβλήματα άπειρου χρονικού ορίζοντα

2.4.1 Ελαχιστοποίηση συνολικού αναμενόμενου αποπληθωρισμένου κόστους

Στο παρόν εδάφιο υποθέτουμε ότι ο αποπληθωριστικός παράγοντας α είναι γνησίως μικρότερος της μονάδας. Επίσης θεωρούμε ότι υπάρχει ένας θετικός πραγματικός αριθμός M τέτοιος ώστε για κάθε ενέργεια a και κάθε κατάσταση i της διαδικασίας, ισχύει ότι: $|C(i,a)| < M.$

Το συνολικό αναμενόμενο αποπληθωρισμένο κόστος $V_\alpha(i,\pi)$ σε άπειρο χρονικό ορίζοντα, υπό τον έλεγχο της πολιτικής $\pi,$ δοθέντος ότι η αρχική κατάσταση της διαδικασίας είναι η κατάσταση $i,$ ορίζεται ως εξής:

$$V_\alpha(i,\pi) = E_\pi\left[\sum_{j=0}^{\infty} \alpha^j C(X_j, a_j) \mid X_0 = i\right].$$

Οι παραπάνω υποθέσεις εξασφαλίζουν ότι: $V_\alpha(i,\pi) < \infty$.

Για κάθε κατάσταση i της διαδικασίας, η *βέλτιστη συνάρτηση* $V_\alpha(i)$ *του αναμενόμενου αποπληθωρισμένου κόστους σε άπειρο χρονικό ορίζοντα,* ορίζεται ως εξής:

$$V_\alpha(i) = \inf_{\pi\in\Im} V_\alpha(i,\pi), \quad i \geq 0.$$

Το Θεώρημα 2.2 παρέχει μία εξίσωση η οποία ικανοποιείται από τη βέλτιστη συνάρτηση $V_\alpha(i),\ i \geq 0.$

Θεώρημα 2.2. Η ποσότητα $V_\alpha(i),\ i \geq 0,$ ικανοποιεί την εξίσωση:

$$V_\alpha(i) = \min_{a\in A(i)}\left\{C(i,a) + \alpha\sum_{j=0}^{\infty} p_{ij}(a)V_\alpha(j)\right\}. \tag{2.2}$$

Επιπλέον υπάρχει μία βέλτιστη στάσιμη πολιτική $f \equiv \{f_i\},\ i \geq 0,$ η οποία, οποτεδήποτε η διαδικασία βρίσκεται στην κατάσταση $i,$ επιλέγει εκείνη την ενέργεια $a = f_i$ που ελαχιστοποιεί το δεξιό μέλος της εξίσωσης (2.2) και ικανοποιεί τη σχέση $V_\alpha(i) = V_\alpha(i, f),\ i \geq 0.$

Η εξίσωση (2.2) είναι γνωστή ως *εξίσωση βελτιστοποίησης* για το πρόβλημα της ελαχιστοποίησης του συνολικού αναμενόμενου αποπληθωρισμένου κόστους σε άπειρο χρονικό ορίζοντα. Η απόδειξη του Θεωρήματος 2.2 βρίσκεται στις σελ. 121 και 124 του βιβλίου του Ross (1992).

Το Θεώρημα Σταθερού Σημείου για συστολές προσφέρει μία διαφορετική προσέγγιση για τον υπολογισμό της βέλτιστης συνάρτησης $V_\alpha(i),\ i \geq 0.$ Με χρήση γνωστών αποτελεσμάτων αυτής της θεωρίας (βλέπε σελ. 125-128 του βιβλίου του Ross (1992)) αποδεικνύεται ότι, για κάθε κατάσταση i της διαδικασίας, η βέλτιστη συνάρτηση $V_\alpha(i)$ είναι η μοναδική λύση της εξίσωσης βελτιστοποίησης (2.2).

Με βάση το παρακάτω θεώρημα υπολογίζουμε, για κάθε κατάσταση i της διαδικασίας, τη βέλτιστη συνάρτηση $V_\alpha(i)$ σε άπειρο χρονικό ορίζοντα μέσω της βέλτιστης συνάρτησης $V_\alpha(i,N)$ όταν απομένουν N βήματα μέχρι το τερματισμό της διαδικασίας. Η μέθοδος υπολογισμού είναι γνωστή ως *μέθοδος των διαδοχικών προσεγγίσεων*.

Θεώρημα 2.3. Έστω ότι η μη-αρνητική συνάρτηση κόστους $F(\cdot)$ η οποία ορίζεται στο χώρο καταστάσεων της διαδικασίας είναι φραγμένη. Τότε ισχύει ότι: $\lim_{N\to\infty} V_\alpha(i,N) = V_\alpha(i),\ \ i \geq 0.$

Η απόδειξη του Θεωρήματος 2.3 βρίσκεται στη σελ. 128 του βιβλίου του Ross (1992). Το θεώρημα είναι αρκετά χρήσιμο καθώς σε πολλές περιπτώσεις μας δίνει τη δυνατότητα να αντιμετωπίσουμε το πρόβλημα του άπειρου χρονικού ορίζοντα μέσω του αντίστοιχου προβλήματος σε πεπερασμένο χρονικό ορίζοντα. Αποτελέσματα τα οποία σχετίζονται με τη βέλτιστη συνάρτηση $V_\alpha(i)$ σε άπειρο χρονικό ορίζοντα αποδεικνύονται με τη χρήση του Θεωρήματος 2.3, αφού πρώτα έχουν αποδειχθεί για την αντίστοιχη βέλτιστη συνάρτηση $V_\alpha(i,N),\ i \geq 0,$ όταν απομένουν N βήματα μέχρι το τερματισμό της διαδικασίας.

Ο αλγόριθμος βελτίωσης των πολιτικών, ο αλγόριθμος των διαδοχικών προσεγγίσεων ο οποίος βασίζεται στο Θεώρημα 2.3 και η μέθοδος του γραμμικού προγραμματισμού αποτελούν τις βασικές υπολογιστικές τεχνικές για την εύρεση της βέλτιστης στάσιμης πολιτικής στο πρόβλημα της ελαχιστοποίησης του συνολικού αναμενόμενου αποπληθωρισμένου κόστους σε άπειρο χρονικό ορίζοντα. Στα βιβλία των Puterman (1994), Tijms (1994) και Heyman and Sobel (2004) παρουσιάζονται αναλυτικά αυτές οι υπολογιστικές τεχνικές με πολλές αριθμητικές εφαρμογές τους.

2.4.2 Ελαχιστοποίηση μακροπρόθεσμου αναμενόμενου μέσου κόστους ανά μονάδα χρόνου

Το μακροπρόθεσμο αναμενόμενο μέσο κόστος $g(i,\pi)$ ανά μονάδα χρόνου, υπό τον έλεγχο της πολιτικής $\pi,$ δοθέντος ότι η αρχική κατάσταση της διαδικασίας είναι η κατάσταση $i,$ ορίζεται ως εξής:

$$g(i,\pi)=\limsup_{n\to\infty}\frac{E_{\pi}\left[\sum_{j=0}^{n-1}C(X_j,a_j)\mid X_0=i\right]}{n}.$$

Για κάθε κατάσταση i της διαδικασίας, μία πολιτική π^* είναι βέλτιστη, αν:

$$g(i,\pi^*)=\min_{\pi\in\Im} g(i,\pi),\quad i\geq 0.$$

Σε αντίθεση με το πρόβλημα της ελαχιστοποίησης του συνολικού αναμενόμενου (αποπληθωρισμένου) κόστους, στο πρόβλημα της ελαχιστοποίησης του μακροπρόθεσμου αναμενόμενου μέσου κόστους ανά μονάδα χρόνου, η ύπαρξη μιας βέλτιστης πολιτικής δεν είναι βέβαιη. Υπάρχουν προβλήματα ελέγχου μιας στοχαστικής διαδικασίας στα οποία η βέλτιστη πολιτική είτε δεν υπάρχει είτε και αν ακόμη υπάρχει δεν είναι μία στάσιμη πολιτική. Παραδείγματα τέτοιων περιπτώσεων αναφέρονται στις σελ. 141-144 του βιβλίου του Ross (1992) και στις σελ. 128-132 του βιβλίου της Sennott (1999).

Πολλοί ερευνητές έχουν ασχοληθεί με το θέμα της κατασκευής κατάλληλων υποθέσεων οι οποίες εξασφαλίζουν την ύπαρξη μιας βέλτιστης στάσιμης πολιτικής. Στο Κεφάλαιο 7 του βιβλίου της Sennott (1999) γίνεται μία αναλυτική παρουσίαση των πρόσφατων αποτελεσμάτων σχετικά με το θέμα αυτό με αρκετές αναφορές σε προηγούμενες εργασίες.

Στο παρόν εδάφιο υποθέτουμε πάλι ότι υπάρχει ένας θετικός πραγματικός αριθμός M τέτοιος ώστε για κάθε ενέργεια a και κάθε κατάσταση i της διαδικασίας, ισχύει ότι: $|C(i,a)|<M$.

Το Θεώρημα 2.4 παρέχει μία ικανή συνθήκη για την ύπαρξη μιας βέλτιστης στάσιμης πολιτικής.

Θεώρημα 2.4. Έστω ότι υπάρχει μία άνω φραγμένη ακολουθία αριθμών $\{h_i\},\ i\geq 0,$ και μία σταθερά g έτσι ώστε:

$$h_i=\min_{a\in A(i)}\left\{C(i,a)-g+\sum_{j=0}^{\infty}p_{ij}(a)h_j\right\},\quad i\geq 0. \tag{2.3}$$

Τότε υπάρχει μία βέλτιστη στάσιμη πολιτική $f \equiv \{f_i\},\ \ i \geq 0,$ η οποία, οποτεδήποτε η διαδικασία βρίσκεται στην κατάσταση $i,$ επιλέγει εκείνη την ενέργεια $a = f_i$ που ελαχιστοποιεί το δεξιό μέλος της εξίσωσης (2.3). Επιπλέον η σταθερά g είναι ίση με $g(i, f),\ \ i \geq 0.$

Η εξίσωση (2.3) είναι γνωστή ως *εξίσωση βελτιστοποίησης* για το πρόβλημα της ελαχιστοποίησης του μακροπρόθεσμου αναμενόμενου μέσου κόστους ανά μονάδα χρόνου και οι τιμές $h_i,\ i \geq 0,$ είναι γνωστές ως οι *σχετικές τιμές* της βέλτιστης στάσιμης πολιτικής. Η απόδειξη του Θεωρήματος 2.4 βρίσκεται στη σελ. 144 του βιβλίου του Ross (1992).

Για κάθε κατάσταση i της διαδικασίας, το παρακάτω θεώρημα παρέχει μία συνθήκη η οποία εγγυάται την ύπαρξη της ακολουθίας των σχετικών τιμών $\{h_i\}$ μέσω της βέλτιστης συνάρτησης $V_\alpha(i),\ \ i \geq 0.$

Θεώρημα 2.5. Έστω ότι για μία κατάσταση της διαδικασίας (π.χ. για την κατάσταση 0) υπάρχει μία σταθερά B τέτοια ώστε:

$$|V_\alpha(i) - V_\alpha(0)| < B, \tag{2.4}$$

για κάθε $i \geq 0$ και κάθε $\alpha \in (0,1).$ Τότε:

(α) Υπάρχει μία φραγμένη ακολουθία αριθμών $\{h_i\},\ i \geq 0,$ και μία σταθερά g που ικανοποιούν την εξίσωση βελτιστοποίησης (2.3).

(β) Υπάρχει μία ακολουθία αριθμών $\{\alpha_n\}$ τέτοια ώστε $\lim_{n\to\infty} \alpha_n = 1,$ για την οποία ισχύει ότι: $h_i = \lim_{n\to\infty}\{V_{\alpha_n}(i) - V_{\alpha_n}(0)\}$ και

(γ) $\lim_{\alpha\to 1}(1-\alpha)V_\alpha(0) = g.$

Η απόδειξη του Θεωρήματος 2.5 βρίσκεται στις σελ. 146-149 του βιβλίου του Ross (1992) στο οποίο περιέχονται επίσης αποτελέσματα που παρέχουν ικανές συνθήκες τέτοιες ώστε να ισχύει η ανισότητα (2.4).

Στο πρόβλημα της ελαχιστοποίησης του μακροπρόθεσμου αναμενόμενου μέσου κόστους ανά μονάδα χρόνου, η ακόλουθη Υπόθεση UC χρειάζεται να εισαχθεί έτσι ώστε να είναι εφικτός ο υπολογισμός μιας βέλτιστης στάσιμης πολιτικής.

Υπόθεση UC: Για κάθε στάσιμη πολιτική f υπάρχει μία κατάσταση r (η οποία μπορεί να εξαρτάται από την πολιτική f) τέτοια ώστε ο αναμενόμενος χρόνος και το αναμενόμενο κόστος που απαιτούνται για τη μετάβαση στην κατάσταση r από οποιαδήποτε αρχική κατάσταση της διαδικασίας, υπό τον έλεγχο της πολιτικής f, είναι πεπερασμένα.

Η Υπόθεση UC εξασφαλίζει επίσης ότι το μακροπρόθεσμο αναμενόμενο μέσο κόστος ανά μονάδα χρόνου οποιασδήποτε στάσιμης πολιτικής που υιοθετείται για τον έλεγχο της διαδικασίας είναι ανεξάρτητο της αρχικής κατάστασης της διαδικασίας.

Όπως στο πρόβλημα της ελαχιστοποίησης του συνολικού αναμενόμενου αποπληθωρισμένου κόστους έτσι και στο πρόβλημα της ελαχιστοποίησης του μακροπρόθεσμου αναμενόμενου μέσου κόστους ανά μονάδα χρόνου, ο αλγόριθμος βελτίωσης των πολιτικών, ο αλγόριθμος των διαδοχικών προσεγγίσεων και η μέθοδος του γραμμικού προγραμματισμού συνιστούν τις βασικές υπολογιστικές τεχνικές. Στα βιβλία των Puterman (1994), Tijms (1994) και Heyman and Sobel (2004) περιέχεται το θεωρητικό υπόβαθρο αυτών των υπολογιστικών τεχνικών καθώς και αρκετές εφαρμογές τους. Ο αλγόριθμος βελτίωσης των πολιτικών βασίζεται στα ακόλουθα Θεωρήματα 2.6 και 2.7 τα οποία με τις αποδείξεις τους βρίσκονται στις σελ. 191-217 του βιβλίου του Tijms (1994).

Θεώρημα 2.6. Έστω ότι υπάρχει μία κατάσταση r η οποία ικανοποιεί την Υπόθεση UC. Έστω g_f το μακροπρόθεσμο αναμενόμενο μέσο κόστος ανά μονάδα χρόνου, υπό τον έλεγχο της στάσιμης πολιτικής $f \equiv \{f_i\},\ i \geq 0,$ και

$$h_i(f) = K_i(f) - g_f T_i(f), \quad i \geq 0, \tag{2.5}$$

όπου, $T_i(f)$ και $K_i(f)$ είναι ο αναμενόμενος χρόνος και το αναμενόμενο κόστος, αντίστοιχα, που απαιτούνται μέχρι η διαδικασία να επιστρέψει στην κατάσταση r, αν αρχικά βρισκόταν στην κατάσταση i και η πολιτική f έχει υιοθετηθεί για τον έλεγχο της διαδικασίας.

Τότε:

Οι ποσότητες $h_i(f)$, $i \geq 0$, και g_f είναι η μοναδική λύση του ακόλουθου συστήματος των γραμμικών εξισώσεων με αγνώστους g και h_i, $i \geq 0$:

$$h_i = C(i, f_i) - g + \sum_{j=0}^{\infty} p_{ij}(f_i)h_j, \quad i \geq 0, \tag{2.6}$$

$$h_r = 0.$$

Θεώρημα 2.7. Έστω ότι υπάρχει μία κατάσταση r η οποία ικανοποιεί την Υπόθεση UC. Έστω g_f και $g_{\tilde{f}}$ τα μακροπρόθεσμα αναμενόμενα μέσα κόστη ανά μονάδα χρόνου, υπό τον έλεγχο των στάσιμων πολιτικών $f \equiv \{f_i\}$ και $\tilde{f} \equiv \{\tilde{f}_i\}$, $i \geq 0$, αντίστοιχα. Υποθέτουμε ότι:

$$C(i, \tilde{f}_i) - g_f + \sum_{j=0}^{\infty} p_{ij}(\tilde{f}_i)h_j(f) \leq h_i(f), \quad i \geq 0, \tag{2.7}$$

όπου οι ποσότητες $h_i(f)$, $i \geq 0$, έχουν οριστεί μέσω της εξίσωσης (2.5). Τότε:

$$g_{\tilde{f}} \leq g_f. \tag{2.8}$$

Το θεώρημα ισχύει και στην περίπτωση κατά την οποία οι ανισότητες (2.7) και (2.8) έχουν αντίθετη φορά. Επίσης η ανισότητα (2.8) ισχύει αυστηρά όταν η ανισότητα (2.7) ισχύει αυστηρά για μία τουλάχιστον κατάσταση i της διαδικασίας η οποία είναι έμμονη θετική υπό τον έλεγχο της πολιτικής $\tilde{f}$.

Οι ποσότητες $h_i(f),\ \ i \geq 0,$ όπως έχουν οριστεί στην εξίσωση (2.5) είναι γνωστές ως οι *σχετικές τιμές* της πολιτικής f.

2.5 Ημι-Μαρκοβιανές διαδικασίες αποφάσεων και Μαρκοβιανές διαδικασίες αποφάσεων σε συνεχή χρόνο

Σε αντίθεση με το Εδάφιο 2.2, στο παρόν εδάφιο θεωρούμε ότι για τον έλεγχο του συστήματος οι ενέργειες επιλέγονται σε τυχαίες χρονικές στιγμές. Έστω X_n η κατάσταση του συστήματος τη χρονική στιγμή $n,\ \ n \geq 0,$ και $t_n,\ \ n \geq 1,$ το χρονικό διάστημα που μεσολαβεί μεταξύ των χρονικών στιγμών $n-1$ και $n.$ Υποθέτουμε ότι $t_0 = 0$ και ότι τη χρονική στιγμή n το σύστημα βρίσκεται στην κατάσταση i στην οποία η ενέργεια a επιλέγεται από ένα σύνολο εναλλακτικών ενεργειών $A(i).$

Το σύστημα είναι μία *ημι-Μαρκοβιανή διαδικασία αποφάσεων*, αν:

(α) Την επόμενη χρονική στιγμή η κατάσταση του συστήματος είναι η κατάσταση j με πιθανότητα $p_{ij}(a)$ η οποία εξαρτάται μόνον από την ενέργεια a και τις καταστάσεις i και $j.$

(β) Δοθέντος ότι την επόμενη χρονική στιγμή η κατάσταση του συστήματος είναι η κατάσταση $j,$ το χρονικό διάστημα που μεσολαβεί μέχρι το σύστημα να μεταβεί από την κατάσταση i στην κατάσταση j είναι μία τυχαία μεταβλητή με συνάρτηση πιθανότητας $F_{ij}(\cdot \mid a).$

(γ) Δοθέντος ότι η μετάβαση του συστήματος από την κατάσταση i στην κατάσταση j διαρκεί t χρονικές μονάδες, υπάρχει ένα άμεσο κόστος ίσο με $k_i(a)$ και ένα κόστος ίσο με $c_i(a)$ ανά μονάδα χρόνου με αποτέλεσμα το συνολικό κόστος να είναι ίσο με $k_i(a) + tc_i(a)$ ως οικονομική συνέπεια της επιλογής της ενέργειας a τη χρονική στιγμή κατά την οποία το σύστημα βρίσκεται στην κατάσταση $i.$

Υποθέτουμε ότι για κάθε ενέργεια a και κάθε κατάσταση i της διαδικασίας, το άμεσο κόστος $k_i(a)$ και το κόστος $c_i(a)$ ανά μονάδα χρόνου είναι φραγμένα.

Στην παρούσα διατριβή οι ημι-Μαρκοβιανές διαδικασίες αποφάσεων θα μας απασχολήσουν στο πρόβλημα της ελαχιστοποίησης του συνολικού αναμενόμενου αποπληθωρισμένου κόστους σε άπειρο χρονικό ορίζοντα καθώς επίσης στο πρόβλημα της ελαχιστοποίησης του μακροπρόθεσμου αναμενόμενου μέσου κόστους ανά μονάδα χρόνου.

Δοθέντος ότι η διαδικασία βρίσκεται στην κατάσταση i και η ενέργεια $a \in A(i)$ επιλέγεται, ο αναμενόμενος χρόνος $T(i,a)$ και το αναμενόμενο αποπληθωρισμένο κόστος $C_\alpha(i,a)$ μέχρι την επόμενη χρονική στιγμή επιλογής ενέργειας, ορίζονται αντίστοιχα, ως εξής:

$$T(i,a) = \sum_{j=0}^{\infty} p_{ij}(a) \int_0^{\infty} t dF_{ij}(t \mid a),$$

$$C_\alpha(i,a) = k_i(a) + \sum_{j=0}^{\infty} p_{ij}(a) \int_0^{\infty} \left[\int_0^t c_i(a) e^{-\alpha s} ds \right] dF_{ij}(t \mid a),$$

όπου, $\alpha > 0$ είναι ο αποπληθωριστικός παράγοντας. Ορίζουμε $C(i,a) \equiv C_0(i,a)$ και θεωρούμε ότι υπάρχει ένας θετικός πραγματικός αριθμός M τέτοιος ώστε για κάθε ενέργεια a και κάθε κατάσταση i της διαδικασίας, ισχύει ότι: $|C(i,a)| < M$.

Το συνολικό αναμενόμενο αποπληθωρισμένο κόστος $V_\alpha(i,N,\pi)$ όταν απομένουν N βήματα μέχρι το τερματισμό της διαδικασίας, το συνολικό αναμενόμενο αποπληθωρισμένο κόστος $V_\alpha(i,\pi)$ σε άπειρο χρονικό ορίζοντα και το μακροπρόθεσμο αναμενόμενο μέσο κόστος $g(i,\pi)$ ανά μονάδα χρόνου, υπό τον έλεγχο της πολιτικής π, δοθέντος ότι η αρχική κατάσταση της διαδικασίας είναι η κατάσταση i, ορίζονται αντίστοιχα, ως εξής:

$$V_\alpha(i,N,\pi) = E_\pi \left[\sum_{j=0}^{N-1} e^{-\alpha \sum_{i=0}^{j} t_i} C_\alpha(X_j, a_j) + e^{-\alpha \sum_{i=0}^{N} t_i} F(X_N) \mid X_0 = i \right],$$

$$V_\alpha(i,\pi) = E_\pi \left[\sum_{j=0}^{\infty} e^{-\alpha \sum_{i=0}^{j} t_i} C_\alpha(X_j, a_j) \mid X_0 = i \right],$$

$$g(i,\pi)=\limsup_{n\to\infty}\frac{E_\pi\left[\sum_{j=0}^{n-1}C(X_j,a_j)\mid X_0=i\right]}{E_\pi\left[\sum_{j=1}^{n-1}t_j\mid X_0=i\right]}.$$

Η ποσότητα $F(X_N)$ αντιπροσωπεύει ένα τελικό κόστος όταν η διαδικασία σταματά τη χρονική στιγμή N, όπου $F(\cdot)$ είναι μία γνωστή μη-αρνητική συνάρτηση κόστους η οποία ορίζεται στο χώρο καταστάσεων της διαδικασίας.

Στην περίπτωση των ημι-Μαρκοβιανών διαδικασιών αποφάσεων τα Θεωρήματα 2.1-2.7 ισχύουν με τις παρακάτω τροποποιήσεις.

(α) Για κάθε κατάσταση i της διαδικασίας, η εξίσωση βελτιστοποίησης για το πρόβλημα της ελαχιστοποίησης του συνολικού αναμενόμενου αποπληθωρισμένου κόστους όταν απομένουν N βήματα μέχρι το τερματισμό της διαδικασίας είναι (βλέπε σελ. 202 του βιβλίου των Heyman and Sobel (2004)):

$$V_\alpha(i,N)=\min_{a\in A(i)}\left\{C_\alpha(i,a)+\sum_{j=0}^{\infty}p_{ij}(a)\int_0^\infty e^{-\alpha t}V_\alpha(j,N-1)dF_{ij}(t\mid a)\right\}.$$

(β) Για κάθε κατάσταση i της διαδικασίας, η εξίσωση βελτιστοποίησης για το πρόβλημα της ελαχιστοποίησης του συνολικού αναμενόμενου αποπληθωρισμένου κόστους σε άπειρο χρονικό ορίζοντα είναι (βλέπε σελ. 157 του βιβλίου του Ross (1992)):

$$V_\alpha(i)=\min_{a\in A(i)}\left\{C_\alpha(i,a)+\sum_{j=0}^{\infty}p_{ij}(a)\int_0^\infty e^{-\alpha t}V_\alpha(j)dF_{ij}(t\mid a)\right\}.$$

(γ) Για κάθε κατάσταση i της διαδικασίας, η εξίσωση βελτιστοποίησης για το πρόβλημα της ελαχιστοποίησης του μακροπρόθεσμου αναμενόμενου μέσου κόστους ανά μονάδα χρόνου είναι (βλέπε σελ. 162 του βιβλίου του Ross (1992)):

$$h_i = \min_{\alpha \in A(i)} \left\{ C(i,a) - gT(i,a) + \sum_{j=0}^{\infty} p_{ij}(a) h_j \right\}.$$

(δ) Αν ισχύουν οι υποθέσεις του Θεωρήματος 2.5, τα αποτελέσματα (β) και (γ) του θεωρήματος διαμορφώνονται ως εξής (βλέπε σελ. 163 του βιβλίου του Ross (1992)):

Υπάρχει μία ακολουθία αριθμών $\{\alpha_n\}$ τέτοια ώστε $\lim_{n \to \infty} \alpha_n = 0$ για την οποία ισχύει ότι:

$$h_i = \lim_{n \to \infty} \{V_{\alpha_n}(i) - V_{\alpha_n}(0)\} \quad \text{και} \quad \lim_{\alpha \to 0} \alpha V_\alpha(0) = g.$$

(ε) Στις εξισώσεις (2.6) του Θεωρήματος 2.6 η σταθερά g αντικαθίσταται από το γινόμενο $gT(i, f_i)$ και στις ανισότητες (2.7) του Θεωρήματος 2.7 η σταθερά g_f αντικαθίσταται από το γινόμενο $g_f T(i, \tilde{f}_i)$.

Για περισσότερες λεπτομέρειες σχετικά με τις τροποποιήσεις των Θεωρημάτων 2.6 και 2.7 στην περίπτωση των ημι-Μαρκοβιανών διαδικασιών αποφάσεων παραπέμπουμε στις σελ. 220-223 του βιβλίου του Tijms (1994).

Τα αποτελέσματα που παρουσιάζονται στο βιβλίο του Ross (1992) αφορούν την περίπτωση κατά την οποία, για κάθε ενέργεια a και κάθε κατάσταση i μιας ημι-Μαρκοβιανής διαδικασίας αποφάσεων, το κόστος $C(i,a)$ είναι φραγμένο.

Στην περίπτωση κατά την οποία, για κάθε ενέργεια a και κάθε κατάσταση i της διαδικασίας, το κόστος $C(i,a)$ δεν είναι άνω φραγμένο, κατάλληλες υποθέσεις χρειάζεται να εισαχθούν οι οποίες εγγυώνται την ύπαρξη μιας βέλτιστης στάσιμης πολιτικής στο πρόβλημα της ελαχιστοποίησης του μακροπρόθεσμου αναμενόμενου μέσου κόστους ανά μονάδα χρόνου.

Για το σκοπό αυτό η Sennott (1989) κατασκεύασε πέντε υποθέσεις. Οι δύο πρώτες υποθέσεις παρουσιάζονται παρακάτω.

Υπόθεση 1: Για κάθε ενέργεια a και κάθε κατάσταση i της διαδικασίας, υπάρχουν $\delta > 0$ και $\varepsilon > 0$ έτσι ώστε με πιθανότητα τουλάχιστον ε ο χρόνος μετάβασης της διαδικασίας από την κατάσταση i σε μία οποιαδήποτε κατάσταση j είναι μεγαλύτερος από δ.

Υπόθεση 2: Για κάθε ενέργεια a και κάθε κατάσταση i της διαδικασίας, υπάρχει ένας θετικός πραγματικός αριθμός B τέτοιος ώστε: $T(i,a) \le B$.

Οι υπόλοιπες τρεις υποθέσεις σχετίζονται με ιδιότητες της βέλτιστης συνάρτησης $V_\alpha(i),\ i \ge 0$. Σε πολλές περιπτώσεις οι υποθέσεις αυτές είναι δύσκολο να αποδειχθούν. Η Sennott (1989) κατασκεύασε την παρακάτω Συνθήκη SEN η οποία εξασφαλίζει την ισχύ αυτών των τριών υποθέσεων.

Συνθήκη SEN: Έστω S ο χώρος καταστάσεων της διαδικασίας και έστω ότι $c_i = \min_{a \in A(i)} \{c_i(a)\}$, $i \in S$. Για κάθε κατάσταση $i \in S$, υπάρχει μία στάσιμη πολιτική $f \equiv \{f_i\}$, υπό τον έλεγχο της οποίας, το μακροπρόθεσμο αναμενόμενο μέσο κόστος ανά μονάδα χρόνου είναι ίσο με g. Επιπλέον η στάσιμη πολιτική f έχει τις ακόλουθες ιδιότητες:

Ιδιότητα 1: Η πολιτική f ορίζει μία ανάγωγη, απεριοδική, έμμονη θετική Μαρκοβιανή αλυσίδα στο χώρο καταστάσεων S και

$$\sum_{i \in S} \pi_i(f) C(i, f_i) < \infty \quad \text{και} \quad \sum_{i \in S} \pi_i(f) T(i, f_i) < \infty,$$

όπου, $\pi_i(f),\ i \in S$, είναι η στάσιμη κατανομή της Μαρκοβιανής αλυσίδας υπό τον έλεγχο της πολιτικής f.

Ιδιότητα 2: Υπάρχει $\varepsilon > 0$ και ένα πεπερασμένο υποσύνολο G του χώρου καταστάσεων S τέτοιο ώστε $c_i \ge g + \varepsilon$ για κάθε κατάσταση $i \in S - G$. Επιπλέον για κάθε κατάσταση $i \in G$, υπάρχει μία στάσιμη πολιτική $\tilde{f} = \{\tilde{f}_i\}$, τέτοια ώστε $c_{0i}(\tilde{f}_i) < \infty$, όπου $c_{0i}(\tilde{f}_i)$ αναπαριστά το αναμενόμενο κόστος για την πρώτη μετάβαση της διαδικασίας από την κατάσταση 0 στην κατάσταση i υπό τον έλεγχο της πολιτικής $\tilde{f}$.

Μία υποκατηγορία των ημι-Μαρκοβιανών διαδικασιών αποφάσεων είναι οι *Μαρκοβιανές διαδικασίες αποφάσεων σε συνεχή χρόνο*. Σε μία Μαρκοβιανή διαδικασία αποφάσεων σε συνεχή χρόνο, το χρονικό διάστημα t_n που μεσολαβεί μεταξύ των χρονικών στιγμών $n-1$ και n, $n \geq 1$, στις οποίες επιλέγεται μία ενέργεια, είναι μία τυχαία μεταβλητή η οποία ακολουθεί την Εκθετική κατανομή.

Στις περιπτώσεις των ημι-Μαρκοβιανών διαδικασιών αποφάσεων και των Μαρκοβιανών διαδικασιών αποφάσεων σε συνεχή χρόνο, η τεχνική της ομοιομορφοποίησης, η οποία επινοήθηκε από τον Schweitzer (1971), διαμορφώνει κατάλληλα τον αλγόριθμο των διαδοχικών προσεγγίσεων ώστε να είναι εφικτός ο υπολογισμός μιας βέλτιστης στάσιμης πολιτικής για το πρόβλημα της ελαχιστοποίησης του μακροπρόθεσμου αναμενόμενου μέσου κόστους ανά μονάδα χρόνου.

Η τεχνική της ομοιομορφοποίησης μετατρέπει την ημι-Μαρκοβιανή διαδικασία αποφάσεων (ή τη Μαρκοβιανή διαδικασία αποφάσεων σε συνεχή χρόνο) σε μία ισοδύναμη Μαρκοβιανή διαδικασία αποφάσεων σε διακριτό χρόνο έτσι ώστε, υπό τον έλεγχο οποιασδήποτε στάσιμης πολιτικής, το μακροπρόθεσμο αναμενόμενο μέσο κόστος ανά μονάδα χρόνου να είναι το ίδιο στις δύο διαδικασίες. Οι δύο διαδικασίες έχουν τον ίδιο χώρο καταστάσεων και για κάθε κατάσταση i, έχουν τα ίδια σύνολα εναλλακτικών ενεργειών $A(i)$. Διαφέρουν μόνο στα κόστη και στις πιθανότητες μετάβασης.

Έστω ότι υπάρχει ένας αριθμός T τέτοιος ώστε: $0 < T < \inf_{i,a} T(i,a)$. Τα κόστη $\tilde{C}(i,a)$ και οι πιθανότητες μετάβασης $\tilde{p}_{ij}(a)$ της ισοδύναμης Μαρκοβιανής διαδικασίας αποφάσεων σε διακριτό χρόνο, ορίζονται αντίστοιχα, ως εξής:

$$\tilde{C}(i,a) = \frac{k_i(a)}{T(i,a)} + c_i(a), \quad i \geq 0, \quad a \in A(i),$$

$$\tilde{p}_{ij}(a) = \frac{T}{T(i,a)} p_{ij}(a), \quad i \geq 0, \quad a \in A(i), \quad j \neq i,$$

$$\tilde{p}_{ii}(a) = 1 - \frac{T}{T(i,a)}, \quad i \geq 0, \quad a \in A(i).$$

Η τεχνική της ομοιομορφοποίησης παρουσιάζεται λεπτομερώς στις σελ. 221-222 του βιβλίου του Tijms (1994) και στις σελ. 242-248 του βιβλίου της Sennott (1999).

Στα βιβλία των Ross (1992), Puterman (1994), Tijms (1994), Sennott (1999) και Heyman and Sobel (2004) παρουσιάζονται οι ημι-Μαρκοβιανές διαδικασίες αποφάσεων και οι Μαρκοβιανές διαδικασίες αποφάσεων σε συνεχή χρόνο. Αναφέρονται διεξοδικά οι υπολογιστικές τεχνικές για την εύρεση μιας βέλτιστης στάσιμης πολιτικής με αρκετά παραδείγματα.

2.6 Η προσεγγιστική μέθοδος της Sennott

Όταν ο χώρος καταστάσεων μιας διαδικασίας είναι άπειρος, ο αλγόριθμος βελτίωσης των πολιτικών, ο αλγόριθμος των διαδοχικών προσεγγίσεων και η μέθοδος του γραμμικού προγραμματισμού δεν είναι δυνατόν να εφαρμοστούν για να υπολογιστεί μία βέλτιστη στάσιμη πολιτική. Διάφοροι ερευνητές έχουν επινοήσει προσεγγιστικές μεθόδους ώστε να αντιμετωπίσουν τη δυσκολία στην εύρεση μιας βέλτιστης στάσιμης πολιτικής η οποία οφείλεται στον άπειρο χώρο καταστάσεων. Ενδεικτικά αναφέρουμε τις εργασίες των White (1980) και Thomas and Stengos (1985).

Η Sennott (1997) επινόησε μία προσεγγιστική μέθοδο η οποία υπολογίζει μία βέλτιστη στάσιμη πολιτική στην περίπτωση κατά την οποία μία Μαρκοβιανή διαδικασία αποφάσεων έχει άπειρο χώρο καταστάσεων και μη-φραγμένα κόστη. Η μέθοδος παρουσιάζεται αναλυτικά για τα διάφορα κριτήρια βελτιστοποίησης στο βιβλίο της Sennott (1999).

Στο παρόν εδάφιο υποθέτουμε ότι για κάθε ενέργεια a και κάθε κατάσταση i της διαδικασίας, το κόστος $C(i,a)$ δεν είναι άνω φραγμένο. Η μέθοδος της Sennott (1997) παρουσιάζεται παρακάτω με συνοπτικό τρόπο.

Θεωρούμε μία Μαρκοβιανή διαδικασία αποφάσεων με άπειρο χώρο καταστάσεων S η οποία προσεγγίζεται μέσω μιας ακολουθίας Μαρκοβιανών διαδικασιών αποφάσεων με πεπερασμένο χώρο καταστάσεων G_N, όπου $N \geq 1$ ακέραιος αριθμός τέτοιος ώστε: $\bigcup_{N=1}^{\infty} G_N = S$. Για κάθε τιμή του N, κάθε κατάσταση $i \in G_N$ και κάθε ενέργεια a, τα κόστη $C(i,a)$ και τα σύνολα των εναλλακτικών ενεργειών $A(i)$ κάθε Μαρκοβιανής διαδικασίας αποφάσεων της ακολουθίας συμπίπτουν με εκείνα της Μαρκοβιανής διαδικασίας αποφάσεων με τον άπειρο χώρο καταστάσεων. Διαφέρουν μόνο οι πιθανότητες μετάβασης.

Για τις διάφορες τιμές του $N = 1, 2, ...$ ο αλγόριθμος των διαδοχικών προσεγγίσεων χρησιμοποιείται για να υπολογίσει τη βέλτιστη στάσιμη πολιτική κάθε διαδικασίας της ακολουθίας. Κατά την εκτέλεση του αλγορίθμου ο αριθμός N αυξάνεται, παίρνει μεγάλες τιμές και καθώς τείνει στο άπειρο, οι βέλτιστες στάσιμες πολιτικές των Μαρκοβιανών διαδικασιών αποφάσεων της ακολουθίας συγκλίνουν σε μία βέλτιστη στάσιμη πολιτική της Μαρκοβιανής διαδικασίας αποφάσεων με τον άπειρο χώρο καταστάσεων.

Πρέπει να εισαχθούν κατάλληλες συνθήκες οι οποίες εγγυώνται τη σύγκλιση του αλγορίθμου των διαδοχικών προσεγγίσεων σε μία βέλτιστη στάσιμη πολιτική της Μαρκοβιανής διαδικασίας αποφάσεων με τον άπειρο χώρο καταστάσεων. Στην εργασία και στο βιβλίο της Sennott (1997, 1999) παρουσιάζονται αναλυτικά αυτές οι συνθήκες με διάφορες εφαρμογές της μεθόδου.

Στα Εδάφια 6.5 και 7.2 της παρούσας διατριβής, η μέθοδος της Sennott (1997) εφαρμόζεται κατάλληλα σε δύο διαφορετικά προβλήματα βέλτιστου ελέγχου ενός πληθυσμού παρασίτων.

ΚΕΦΑΛΑΙΟ 3

Βέλτιστος έλεγχος δύο ανταγωνιζομένων ασθενειών ή ειδών

3.1 Εισαγωγή

Θεωρούμε έναν πληθυσμό ατόμων τα οποία είναι δυνατόν να προσβληθούν από δύο μεταδοτικές ασθένειες. Υποθέτουμε ότι το συνολικό μέγεθος του πληθυσμού είναι ίσο με N και ότι σε κάθε χρονική στιγμή $t \geq 0$ το πολύ ένα άτομο μπορεί να προσβληθεί από μία από τις δύο ασθένειες. Θεωρούμε ότι οι δύο ασθένειες είναι ανταγωνιζόμενες υπό την έννοια ότι αν ένα άτομο προσβληθεί από την ασθένεια r $(r = 1, 2),$ παραμένει προσβεβλημένο από αυτήν την ασθένεια και δεν μπορεί να προσβληθεί από την άλλη.

Έστω η διδιάστατη Μαρκοβιανή διαδικασία σε συνεχή χρόνο $\{(X(t), Y(t)), t \geq 0\}.$ Οι τυχαίες μεταβλητές $X(t)$ και $Y(t),$ $0 \leq X(t) + Y(t) \leq N,$ αναπαριστούν την κατάσταση της επιδημικής διαδικασίας τη χρονική στιγμή $t \geq 0.$ Η τυχαία μεταβλητή $X(t)$ αναπαριστά τον αριθμό των ατόμων που έχουν προσβληθεί από την ασθένεια 1 και η τυχαία μεταβλητή $Y(t)$ αναπαριστά τον αριθμό των ατόμων που έχουν προσβληθεί από την ασθένεια 2, αντίστοιχα, τη χρονική στιγμή $t \geq 0.$ Υποθέτουμε ότι οι πιθανότητες ένα άτομο να προσβληθεί από τις ασθένειες 1 και 2 σε ένα χρονικό διάστημα $(t, t + \delta t),$ καθώς $\delta t \to 0,$ δοθέντος ότι $X(t) = x$ και $Y(t) = y,$ είναι ίσες με $c_1 x^\alpha (N - x - y)\delta t + o(\delta t)$ και $c_2 y^\beta (N - x - y)\delta t + o(\delta t),$ αντίστοιχα, όπου $c_1, c_2, \alpha, \beta > 0.$ Η συνάρτηση $o(\cdot)$ είναι τέτοια ώστε $o(h)/h \to 0,$ καθώς $h \to 0.$ Όλες οι υπόλοιπες μεταβάσεις της διαδικασίας έχουν πιθανότητα ίση με $o(\delta t),$ καθώς $\delta t \to 0.$ Η διαδικασία σταματά όταν ο συνολικός αριθμός των ατόμων που έχουν προσβληθεί από τις ασθένειες 1 και 2 γίνει ίσος με $N,$ το οποίο θεωρούμε ότι σχεδόν σίγουρα θα συμβεί σε πεπερασμένο χρόνο. Οι μεταβάσεις του τυχαίου περίπατου της επιδημικής διαδικασίας (βλέπε π.χ. σελ. 68 του βιβλίου του Ross (1992)) είναι:

$$(x, y) \to (x + 1, y) \quad \text{με πιθανότητα} \quad \frac{c_1 x^\alpha}{c_1 x^\alpha + c_2 y^\beta}, \tag{3.1}$$

$$(x, y) \to (x, y+1) \quad \text{με πιθανότητα} \quad \frac{c_2 y^{\beta}}{c_1 x^{\alpha} + c_2 y^{\beta}}. \tag{3.2}$$

Αν $\alpha = \beta = 1$, η διαδικασία συμπίπτει με την απλή διδιάστατη επιδημική διαδικασία την οποία εισήγαγαν οι Billard et al. (1979). Γενίκευσαν την απλή μονοδιάστατη επιδημική διαδικασία η οποία επινοήθηκε από τον Bailey (1950).

Στους θετικούς πραγματικούς αριθμούς α και β μπορούμε να αποδώσουμε τον όρο «μολυσματική ισχύς» των ασθενειών 1 και 2, αντίστοιχα, και να τον δικαιολογήσουμε ως εξής. Αν η εξάπλωση μιας μεταδοτικής ασθένειας σε έναν πληθυσμό επιδεκτικών ατόμων εξαρτάται περισσότερο από το κατά πόσο ένα επιδεκτικό άτομο είναι επιρρεπές στην ασθένεια και λιγότερο από το κατά πόσο ένα άτομο που έχει προσβληθεί από την ασθένεια μπορεί να τη μεταδώσει στον υπόλοιπο πληθυσμό, τότε ο ρυθμός με τον οποίο νέα άτομα θα προσβληθούν από την ασθένεια δεν εξαρτάται ιδιαίτερα από τον αριθμό των ατόμων που ήδη έχουν προσβληθεί από την ασθένεια. Σε αυτή την περίπτωση μπορούμε να θεωρήσουμε ότι η μολυσματική ισχύς των ασθενειών 1 και 2 είναι μικρή και οι θετικοί πραγματικοί αριθμοί α και β παίρνουν τιμές κοντά στο μηδέν. Στην αντίθετη περίπτωση κατά την οποία, η μολυσματική ισχύς των ασθενειών 1 και 2 είναι μεγάλη, μπορούμε να θεωρήσουμε ότι οι παράμετροι α και β παίρνουν τιμές μεγαλύτερες της μονάδας. Σε αυτή την περίπτωση οι επιδημίες εξαπλώνονται στον πληθυσμό με πολύ γρήγορους ρυθμούς.

Η έννοια της μολυσματικής ισχύος επινοήθηκε από τον Severo (1969) ο οποίος γενίκευσε την απλή επιδημική διαδικασία και υπολόγισε τις πιθανότητες μετάβασης της.

Στις εργασίες των Saunders (1980a, b), Gleissner (1988), Ball and O’ Neil (1993), O’ Neil (1997) και Clancy (1999a, b) μελετώνται διάφορες επιδημικές διαδικασίες στις οποίες ο ρυθμός προσβολής των επιδεκτικών από μία μεταδοτική ασθένεια δεν είναι σταθερός αλλά εξαρτάται από το πλήθος των επιδεκτικών και των προσβληθέντων ατόμων.

Υποθέτουμε ότι η ασθένεια 1 προξενεί σοβαρά συμπτώματα σε ένα άτομο που έχει προσβληθεί από αυτήν και μειώνει την παραγωγικότητά του. Η παρουσία ενός ατόμου που έχει προσβληθεί από την ασθένεια 1 επιφέρει κάποιο κόστος στην κοινωνία το οποίο θεωρούμε ότι είναι σταθερό και ίσο με τη μονάδα. Υποθέτουμε ότι η ασθένεια 2, σε σύγκριση με την ασθένεια 1, είναι λιγότερο επιβλαβής για ένα άτομο που έχει προσβληθεί από αυτήν. Θεωρούμε ότι η παρουσία ενός ατόμου που έχει προσβληθεί από την ασθένεια 2 δεν επιφέρει κανένα κόστος στην κοινωνία.

Ο έλεγχος της επιδημικής διαδικασίας σε κάθε χρονική στιγμή μπορεί να πραγματοποιηθεί με την επιλογή μιας ενέργειας. Θεωρούμε ότι μία ενέργεια, η οποία μπορεί να ελέγξει τη διαδικασία σε κάθε χρονική στιγμή, είναι ο εμβολιασμός με την ήπια ασθένεια 2 οποιουδήποτε αριθμού επιδεκτικών ατόμων έχουν απομείνει στον πληθυσμό και δεν έχουν προσβληθεί από καμία από τις δύο ασθένειες. Θεωρούμε ότι ο εμβολιασμός ενός ατόμου με την ήπια ασθένεια 2 επιφέρει ένα κόστος το οποίο είναι ίσο με $K > 0$.

Μία άλλη ενέργεια η οποία επίσης θεωρούμε ότι μπορεί να ελέγξει την επιδημική διαδικασία σε κάθε χρονική στιγμή είναι η απομόνωση κάποιων ή όλων των ατόμων που έχουν προσβληθεί από τη σοβαρή ασθένεια 1. Υποθέτουμε ότι η απομόνωση ενός ατόμου που έχει προσβληθεί από τη σοβαρή ασθένεια 1 επιφέρει ένα κόστος το οποίο είναι ίσο με $L > 0$.

Μία πολιτική είναι ένας κανόνας ο οποίος σε κάθε χρονική στιγμή καθορίζει την ενέργεια που επιλέγεται για τον έλεγχο της διαδικασίας. Στο παρόν κεφάλαιο θα μας απασχολήσει το πρόβλημα της εύρεσης εκείνης της πολιτικής η οποία, για οποιαδήποτε αρχική κατάσταση της επιδημικής διαδικασίας, ελαχιστοποιεί το συνολικό αναμενόμενο κόστος. Επειδή η διαδικασία θεωρούμε ότι σταματά όταν ο συνολικός αριθμός των ατόμων που έχουν προσβληθεί από τις ασθένειες 1 και 2 γίνει ίσος με $N,$ το πρόβλημα της εύρεσης της βέλτιστης πολιτικής είναι ένα πρόβλημα πεπερασμένου χρονικού ορίζοντα.

Όπως αναφέρεται στην εργασία του Kyriakidis (1995) η επιδημική διαδικασία που περιγράψαμε στο παρόν εδάφιο βρίσκει πιθανή εφαρμογή στην περίπτωση της γνωστής ασθένειας του νωτιαίου μυελού, πολιομυελίτιδας. Η ασθένεια 1 μπορεί να θεωρηθεί ότι είναι η σοβαρή μορφή της πολιομυελίτιδας ενώ η ασθένεια 2 μπορεί να θεωρηθεί ότι είναι η ήπια μορφή της.

Σύμφωνα με τον Kyriakidis (1995) στην επιδημική διαδικασία μπορεί επίσης να αποδοθεί η ακόλουθη οικολογική ερμηνεία. Θεωρούμε δύο είδη ζωντανών οργανισμών τα οποία αναπτύσσονται σε ένα βιότοπο που έχει μέγιστη χωρητικότητα ίση με $N.$ Το είδος 1 θεωρούμε ότι είναι ένα παράσιτο, η παρουσία του οποίου είναι βλαβερή. Η παρουσία ενός παρασίτου επιφέρει κάποιο κόστος το οποίο είναι σταθερό και ίσο με τη μονάδα. Το είδος 2 θεωρούμε ότι είναι ένα ήπιο είδος, η παρουσία του οποίου είναι ακίνδυνη. Η παρουσία ενός ήπιου είδους δεν επιφέρει κανένα κόστος. Θεωρούμε πολιτικές οι οποίες σε κάθε χρονική στιγμή ελέγχουν την ανάπτυξη των ζωντανών οργανισμών στο βιότοπο είτε με τη σκόπιμη εισαγωγή ήπιων ειδών είτε με την απομόνωση ή την απομάκρυνση από το βιότοπο οποιουδήποτε αριθμού παρασίτων. Η σκόπιμη εισαγωγή ενός ήπιου είδους επιφέρει ένα κόστος ίσο με $K > 0$ ενώ η απομόνωση ή η απομάκρυνση ενός παρασίτου επιφέρει ένα κόστος ίσο με $L > 0$.

Για την επιδημική διαδικασία θεωρούμε τα παρακάτω τέσσερα προβλήματα βελτιστοποίησης.

Πρόβλημα 1. Εύρεση εκείνης της πολιτικής η οποία, για οποιαδήποτε αρχική κατάσταση της διαδικασίας, ελαχιστοποιεί το συνολικό αναμενόμενο κόστος, αν η διαδικασία σε κάθε χρονική στιγμή είναι δυνατόν να ελεγχθεί μέσω του εμβολιασμού με την ήπια ασθένεια 2 οποιουδήποτε αριθμού επιδεκτικών ατόμων έχουν απομείνει στον πληθυσμό και δεν έχουν προσβληθεί από καμία από τις δύο ασθένειες.

Πρόβλημα 2. Εύρεση εκείνης της πολιτικής η οποία, για οποιαδήποτε αρχική κατάσταση της διαδικασίας, ελαχιστοποιεί το συνολικό αναμενόμενο κόστος, αν η διαδικασία σε κάθε χρονική στιγμή είναι δυνατόν να ελεγχθεί μέσω της απομόνωσης οποιουδήποτε αριθμού ατόμων που έχουν προσβληθεί από τη σοβαρή ασθένεια 1.

Πρόβλημα 3. Εύρεση εκείνης της πολιτικής η οποία, για οποιαδήποτε αρχική κατάσταση της διαδικασίας, ελαχιστοποιεί το συνολικό αναμενόμενο κόστος, αν η διαδικασία σε κάθε χρονική στιγμή είναι δυνατόν να ελεγχθεί μέσω της απομόνωσης κανενός ή όλων των ατόμων που έχουν προσβληθεί από τη σοβαρή ασθένεια 1.

Πρόβλημα 4. Εύρεση εκείνης της πολιτικής η οποία, για οποιαδήποτε αρχική κατάσταση της διαδικασίας, ελαχιστοποιεί το συνολικό αναμενόμενο κόστος, αν η διαδικασία σε κάθε χρονική στιγμή είναι δυνατόν να ελεγχθεί είτε μέσω του εμβολιασμού με την ήπια ασθένεια 2 οποιουδήποτε αριθμού επιδεκτικών ατόμων έχουν απομείνει στον πληθυσμό και δεν έχουν προσβληθεί από καμία από τις δύο ασθένειες είτε μέσω της απομόνωσης κανενός ή όλων των ατόμων που έχουν προσβληθεί από τη σοβαρή ασθένεια 1.

Σε δύο προηγούμενες εργασίες του Kyriakidis (1995, 1999c) κατάλληλοι αλγόριθμοι του δυναμικού προγραμματισμού έχουν κατασκευαστεί για τα Προβλήματα 1 και 3 στην περίπτωση κατά την οποία $\alpha = \beta = 1$. Οι αντίστοιχες ντετερμινιστικές επιδημικές διαδικασίες έχουν μελετηθεί και οι βέλτιστες πολιτικές τους έχουν βρεθεί αναλυτικά και έχουν συγκριθεί αριθμητικά με τις βέλτιστες πολιτικές των αντίστοιχων στοχαστικών επιδημικών διαδικασιών.

Τα παραπάνω τέσσερα προβλήματα βελτιστοποίησης σχετίζονται με δύο παρόμοια προβλήματα που μελετήθηκαν από τον Abakuks (1973, 1974). Στα προβλήματα αυτά μελετήθηκε ο έλεγχος μιας γενικής επιδημικής διαδικασίας. Ο Abakuks θεώρησε ότι η εξάπλωση μιας επιδημίας σε έναν πληθυσμό μπορεί να ελεγχθεί σε κάθε χρονική στιγμή με την επιλογή μιας ενέργειας η οποία είτε απομονώνει τα άτομα που έχουν προσβληθεί από την ασθένεια είτε ανοσοποιεί τα άτομα που δεν έχουν προσβληθεί από την ασθένεια, αντίστοιχα. Δοθέντος ενός αριθμού επιδεκτικών ατόμων αποδείχθηκε ότι η βέλτιστη πολιτική απομονώνει τους προσβληθέντες ή ανοσοποιεί τους επιδεκτικούς αν και μόνο αν ο αριθμός των προσβληθέντων από την ασθένεια είναι μικρότερος μιας κρίσιμης τιμής ή υπερβαίνει μία κρίσιμη τιμή, αντίστοιχα. Ο Clancy (1999a) γενίκευσε τα αποτελέσματα του Abakuks για τη γενική επιδημική διαδικασία στην περίπτωση κατά την οποία ο ρυθμός προσβολής των επιδεκτικών από μία μεταδοτική ασθένεια εξαρτάται από τους αριθμούς των επιδεκτικών και των προσβληθέντων ατόμων.

Στο επόμενο εδάφιο κατάλληλοι αλγόριθμοι του δυναμικού προγραμματισμού κατασκευάζονται για τα Προβλήματα 1, 3 και 4. Τα αριθμητικά αποτελέσματα των αλγορίθμων οδηγούν σε εικασίες σχετικά με τη μορφή της βέλτιστης πολιτικής. Στο Εδάφιο 3.3 μελετούμε την αντίστοιχη ντετερμινιστική επιδημική διαδικασία. Η μορφή της βέλτιστης πολιτικής αποδεικνύεται αναλυτικά για το Πρόβλημα 1 στην περίπτωση κατά την οποία $\alpha \neq 1$ και $\beta = 1$ καθώς επίσης για το Πρόβλημα 2 στην περίπτωση κατά την οποία $\alpha = 1$ και $\beta \neq 1$. Οι βέλτιστες πολιτικές των ντετερμινιστικών επιδημικών διαδικασιών συγκρίνονται αριθμητικά με τις βέλτιστες πολιτικές των αντίστοιχων στοχαστικών επιδημικών διαδικασιών. Στο Εδάφιο 3.4, σε αντίθεση με το Εδάφιο 3.2, θεωρούμε ότι η ποσότητα $c_1^{-1}c_2$ δεν είναι ένας σταθερός θετικός πραγματικός αριθμός αλλά μία τυχαία μεταβλητή η οποία ακολουθεί μία γνωστή κατανομή. Κατάλληλοι αλγόριθμοι του δυναμικού προγραμματισμού κατασκευάζονται για τα Προβλήματα 1 και 3. Οι βέλτιστες πολιτικές συγκρίνονται αριθμητικά με τις βέλτιστες πολιτικές των αντίστοιχων επιδημικών διαδικασιών του Εδαφίου 3.2.

Τα αποτελέσματα των Εδαφίων 3.2 και 3.3 έχουν δημοσιευτεί στην εργασία των Kyriakidis and Dimitrakos (2002).

3.2 Αλγόριθμοι δυναμικού προγραμματισμού για τα Προβλήματα 1, 3 και 4

Η κατάσταση της διδιάστατης στοχαστικής επιδημικής διαδικασίας που περιγράψαμε στο προηγούμενο εδάφιο μπορεί να αναπαρασταθεί σε κάθε χρονική στιγμή με το ζεύγος των μεταβλητών (x, y), όπου $0 \leq x, y \leq N$ και $0 < x + y \leq N$.

Κατάλληλοι αλγόριθμοι του δυναμικού προγραμματισμού κατασκευάζονται για κάθε ένα από τα Προβλήματα 1, 3 και 4 ώστε να βρεθεί η μορφή της βέλτιστης πολιτικής.

Πρόβλημα 1. Σε κάθε κατάσταση (x, y) της διαδικασίας τέτοια ώστε $x + y < N$ μπορούμε να επιλέξουμε μία από τις δύο παρακάτω ενέργειες:

(i) να αφήσουμε τη διαδικασία να μεταβεί στις καταστάσεις $(x+1, y)$ και $(x, y+1)$, σύμφωνα με τον τυχαίο περίπατο με πιθανότητες οι οποίες δίνονται από τις (3.1), (3.2) και κόστος ίσο με 1 και 0, αντίστοιχα, ή
(ii) να εμβολιάσουμε με την ήπια ασθένεια 2 και με κόστος ίσο με K ένα από τα επιδεκτικά άτομα που έχουν απομείνει στον πληθυσμό και δεν έχουν προσβληθεί από καμία από τις δύο ασθένειες.

Αν $y \leq N - n - x$, μία ακολουθία από n επιλογές της ενέργειας (ii) θα έχει ως επακόλουθο τον εμβολιασμό n επιδεκτικών ατόμων. Κατά αυτόν τον τρόπο έχουμε την ακαριαία μετάβαση $(x, y) \rightarrow (x, y+n)$.

Για κάθε κατάσταση (x, y) της επιδημικής διαδικασίας, έστω $V(x, y)$ το ελάχιστο αναμενόμενο κόστος και $W(x, y)$ το ελάχιστο αναμενόμενο κόστος όταν η διαδικασία πραγματοποιήσει μία μετάβαση σύμφωνα με τον τυχαίο περίπατο με πιθανότητες οι οποίες δίνονται από τις (3.1), (3.2) και στη συνέχεια υιοθετηθεί η βέλτιστη πολιτική. Επειδή η διαδικασία σταματά όταν $x + y = N$, η εξίσωση βελτιστοποίησης (2.1) στη σελ. 18 της παρούσας διατριβής, για το πρόβλημα της ελαχιστοποίησης του συνολικού αναμενόμενου κόστους σε πεπερασμένο χρονικό ορίζοντα, παίρνει την ακόλουθη μορφή:

$$V(x, y) = \min\{K + V(x, y+1), W(x, y)\}, \quad 0 < x + y < N, \tag{3.3}$$

όπου,

$$W(x,y)=\frac{c_1x^{\alpha}}{c_1x^{\alpha}+c_2y^{\beta}}[1+V(x+1,y)]+\frac{c_2y^{\beta}}{c_1x^{\alpha}+c_2y^{\beta}}V(x,y+1),\quad 0<x+y<N, \tag{3.4}$$

και

$$V(x,N-x)=0,\quad 0\le x\le N. \tag{3.5}$$

Όταν η διαδικασία βρίσκεται στην κατάσταση (x,y) και ισχύει η ανισότητα $K+V(x,y+1)<W(x,y),$ τότε η βέλτιστη πολιτική επιλέγει την ενέργεια (ii), δηλαδή εμβολιάζει με την ήπια ασθένεια 2 ένα από τα επιδεκτικά άτομα που έχουν απομείνει στον πληθυσμό και δεν έχουν προσβληθεί από καμία από τις δύο ασθένειες. Στην περίπτωση αυτή η διαδικασία μεταβαίνει στην κατάσταση $(x,y+1).$

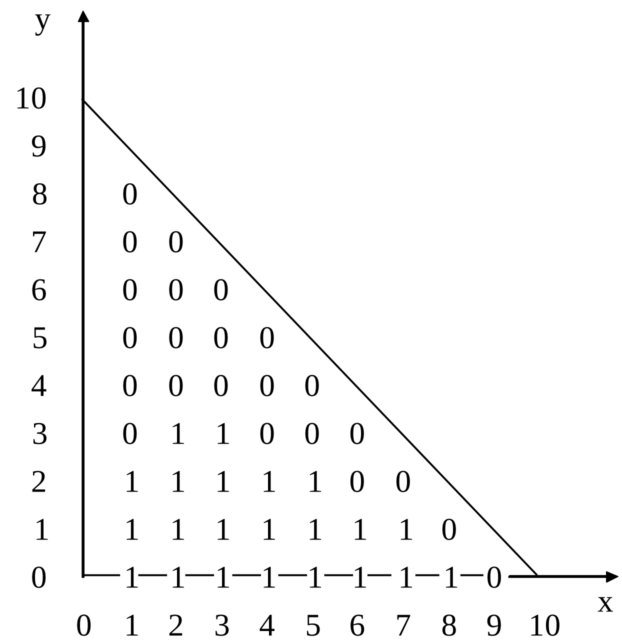
Όταν η διαδικασία βρίσκεται στην κατάσταση (x,y) και ισχύει η ανισότητα $W(x,y)\le K+V(x,y+1),$ τότε η βέλτιστη πολιτική επιλέγει την ενέργεια (i), δηλαδή δεν επεμβαίνει στην εξέλιξη της επιδημικής διαδικασίας.

Οι εξισώσεις (3.3)-(3.5) μας επιτρέπουν να υπολογίσουμε αριθμητικά το ελάχιστο αναμενόμενο κόστος $V(x,y)$ για κάθε κατάσταση (x,y) της διαδικασίας για την οποία ισχύει ότι $0<x+y<N.$ Επιπλέον προσδιορίζουν την ενέργεια που επιλέγεται από τη βέλτιστη πολιτική για κάθε κατάσταση $(x,y), 0<x+y<N.$

Το ελάχιστο αναμενόμενο κόστος υπολογίζεται διαδοχικά για τις καταστάσεις $(1,N-2),$ $(2,N-3),$ ..., $(N-1,0),$ $(1,N-3),$ ..., $(N-2,0),$ ..., $(1,0),$ από τις εξισώσεις (3.3)-(3.5) με αναδρομικό τρόπο.

Ένα αριθμητικό παράδειγμα παρουσιάζεται παρακάτω. Θεωρούμε την περίπτωση κατά την οποία $N=10, K=1, \alpha=2, \beta=1, c_1=1.5, c_2=1.$ Η βέλτιστη πολιτική για αυτές τις τιμές των παραμέτρων παρουσιάζεται στο ακόλουθο Σχήμα 3.1, όπου για κάθε κατάσταση (x,y) της διαδικασίας για την οποία ισχύει ότι $0<x+y\le 9,$ η ενέργεια (i) αναπαριστάται με "0" και η

ενέργεια (ii) αναπαριστάται με “1”. Στο σχήμα παρατηρούμε ότι δεν υπάρχει κανένα “1” το οποίο βρίσκεται πάνω από ένα “0”.



Σχήμα 3.1. *Η βέλτιστη πολιτική όταν* $(N, K, \alpha, \beta, c_1, c_2) = (10, 1, 2, 1, 1.5, 1)$.

Πολλά αριθμητικά παραδείγματα για διάφορες τιμές των παραμέτρων παρέχουν ισχυρή ένδειξη ότι η βέλτιστη πολιτική έχει πάντα αυτήν την ιδιότητα. Έτσι οδηγούμαστε στην ακόλουθη εικασία σχετικά με τη μορφή της βέλτιστης πολιτικής.

Εικασία για το Πρόβλημα 1. Για κάθε ακέραιο αριθμό $x,\ 0 < x < N,$ υπάρχουν δύο περιπτώσεις.

Περίπτωση 1. Η βέλτιστη πολιτική δεν επεμβαίνει στην εξέλιξη της διαδικασίας σε όλες τις καταστάσεις (x, y) για τις οποίες ισχύει ότι $0 \le y < N - x.$

Περίπτωση 2. Υπάρχει ένας ακέραιος αριθμός $\tilde{y},\ 0 \le \tilde{y} < N - x,$ τέτοιος ώστε η βέλτιστη πολιτική επιλέγει την ενέργεια (ii) σε όλες τις καταστάσεις (x, y) για τις οποίες ισχύει ότι $0 \le y \le \tilde{y}$ και επιλέγει την ενέργεια (i), δηλαδή δεν επεμβαίνει στην εξέλιξη της διαδικασίας, σε όλες τις καταστάσεις (x, y) για τις οποίες ισχύει ότι $\tilde{y} < y < N - x.$

Στην Περίπτωση 2 της παραπάνω εικασίας μπορούμε να θεωρήσουμε ότι η βέλτιστη πολιτική είναι μονότονη. Για κάθε κατάσταση (x, y) της διαδικασίας η κρίσιμη τιμή $\widetilde{y}$ χαρακτηρίζει τη μονότονη μορφή της βέλτιστης πολιτικής.

Πρόβλημα 3. Σε κάθε κατάσταση (x, y) της διαδικασίας τέτοια ώστε $x + y < N$ μπορούμε να επιλέξουμε μία από τις δύο παρακάτω ενέργειες:

(i) να αφήσουμε τη διαδικασία να μεταβεί στις καταστάσεις $(x+1, y)$ και $(x, y+1)$, σύμφωνα με τον τυχαίο περίπατο με πιθανότητες οι οποίες δίνονται από τις (3.1), (3.2) και κόστος ίσο με 1 και 0, αντίστοιχα, ή
(ii) να απομονώσουμε όλα τα άτομα που έχουν προσβληθεί από τη σοβαρή ασθένεια 1 με κόστος ίσο με $Lx.$

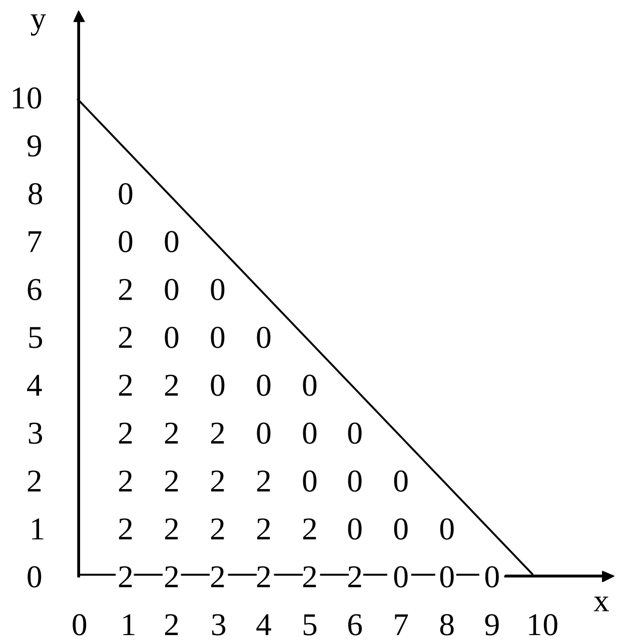
Η εξίσωση βελτιστοποίησης (2.1) παίρνει τώρα την ακόλουθη μορφή:

$$V(x, y) = \min\{Lx, W(x, y)\}, \quad 0 < x + y < N. \tag{3.6}$$

Όταν η διαδικασία βρίσκεται στην κατάσταση (x, y) και ισχύει η ανισότητα $Lx < W(x, y),$ τότε η βέλτιστη πολιτική επιλέγει την ενέργεια (ii), δηλαδή απομονώνει τα x άτομα που έχουν προσβληθεί από τη σοβαρή ασθένεια 1. Στην περίπτωση αυτή η διαδικασία μεταβαίνει στην κατάσταση $(0, y).$

Όταν η διαδικασία βρίσκεται στην κατάσταση (x, y) και ισχύει η ανισότητα $W(x, y) \leq Lx,$ τότε η βέλτιστη πολιτική επιλέγει την ενέργεια (i), δηλαδή δεν επεμβαίνει στην εξέλιξη της επιδημικής διαδικασίας.

Ένα αριθμητικό παράδειγμα παρουσιάζεται παρακάτω. Θεωρούμε την περίπτωση κατά την οποία $N = 10, L = 0.6, \alpha = 1, \beta = 0.5, c_1 = 0.8, c_2 = 1.2.$ Η βέλτιστη πολιτική για αυτές τις τιμές των παραμέτρων παρουσιάζεται στο ακόλουθο Σχήμα 3.2, όπου για κάθε κατάσταση (x, y) της διαδικασίας για την οποία ισχύει ότι $0 < x + y \leq 9,$ η ενέργεια (i) αναπαριστάται με “0” και η ενέργεια (ii) αναπαριστάται με "2". Στο σχήμα παρατηρούμε ότι δεν υπάρχει κανένα "2" το οποίο βρίσκεται δεξιά από ένα "0".



Σχήμα 3.2. *Η βέλτιστη πολιτική όταν* $(N, L, \alpha, \beta, c_1, c_2) = (10, 0.6, 1, 0.5, 0.8, 1.2)$.

Πολλά αριθμητικά παραδείγματα για διάφορες τιμές των παραμέτρων παρέχουν ισχυρή ένδειξη ότι η βέλτιστη πολιτική έχει πάντα αυτήν την ιδιότητα. Έτσι οδηγούμαστε στην ακόλουθη εικασία σχετικά με τη μορφή της βέλτιστης πολιτικής.

Εικασία για το Πρόβλημα 3. Για κάθε ακέραιο αριθμό y, $0 \le y < N$, υπάρχουν δύο περιπτώσεις.

Περίπτωση 1. Η βέλτιστη πολιτική δεν επεμβαίνει στην εξέλιξη της διαδικασίας σε όλες τις καταστάσεις (x, y) για τις οποίες ισχύει ότι $0 \le x < N - y$.

Περίπτωση 2. Υπάρχει ένας ακέραιος αριθμός $\tilde{x}$, $1 \le \tilde{x} < N - y$, τέτοιος ώστε η βέλτιστη πολιτική επιλέγει την ενέργεια (ii) σε όλες τις καταστάσεις (x, y) για τις οποίες ισχύει ότι $1 \le x \le \tilde{x}$ και επιλέγει την ενέργεια (i), δηλαδή δεν επεμβαίνει στην εξέλιξη της διαδικασίας, σε όλες τις καταστάσεις (x, y) για τις οποίες ισχύει ότι $\tilde{x} < x < N - y$.

Στην Περίπτωση 2 της παραπάνω εικασίας μπορούμε να θεωρήσουμε ότι η βέλτιστη πολιτική είναι μονότονη. Για κάθε κατάσταση (x, y) της διαδικασίας η κρίσιμη τιμή $\widetilde{x}$ χαρακτηρίζει τη μονότονη μορφή της βέλτιστης πολιτικής.

Πρόβλημα 4. Σε κάθε κατάσταση (x, y) της διαδικασίας τέτοια ώστε $x + y < N$ μπορούμε να επιλέξουμε μία από τις τρεις παρακάτω ενέργειες:

(i) να αφήσουμε τη διαδικασία να μεταβεί στις καταστάσεις $(x+1, y)$ και $(x, y+1)$, σύμφωνα με τον τυχαίο περίπατο με πιθανότητες οι οποίες δίνονται από τις (3.1), (3.2) και κόστος ίσο με 1 και 0, αντίστοιχα, ή
(ii) να εμβολιάσουμε με την ήπια ασθένεια 2 και με κόστος ίσο με K ένα από τα επιδεκτικά άτομα που έχουν απομείνει στον πληθυσμό και δεν έχουν προσβληθεί από καμία από τις δύο ασθένειες, ή
(iii) να απομονώσουμε όλα τα άτομα που έχουν προσβληθεί από τη σοβαρή ασθένεια 1 με κόστος ίσο με Lx.

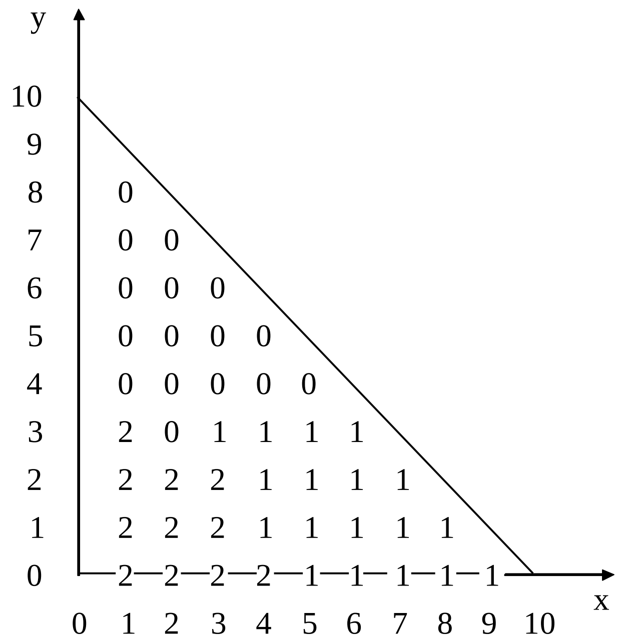
Η εξίσωση βελτιστοποίησης (2.1) παίρνει τώρα την ακόλουθη μορφή:

$$V(x, y) = \min\{K + V(x, y+1), Lx, W(x, y)\}, \quad 0 < x + y < N.$$

Όταν η διαδικασία βρίσκεται στην κατάσταση (x, y) και ισχύουν οι ανισότητες $\min\{K + V(x, y+1), Lx\} < W(x, y)$ και $K + V(x, y+1) \leq Lx,$ τότε η βέλτιστη πολιτική επιλέγει την ενέργεια (ii), δηλαδή εμβολιάζει με την ήπια ασθένεια 2 ένα από τα επιδεκτικά άτομα που έχουν απομείνει στον πληθυσμό και δεν έχουν προσβληθεί από καμία από τις δύο ασθένειες. Στην περίπτωση αυτή η διαδικασία μεταβαίνει στην κατάσταση $(x, y+1)$.

Όταν η διαδικασία βρίσκεται στην κατάσταση (x, y) και ισχύουν οι ανισότητες $\min\{K + V(x, y+1), Lx\} < W(x, y)$ και $Lx < K + V(x, y+1),$ τότε η βέλτιστη πολιτική επιλέγει την ενέργεια (iii), δηλαδή απομονώνει τα x άτομα που έχουν προσβληθεί από τη σοβαρή ασθένεια 1. Στην περίπτωση αυτή η διαδικασία μεταβαίνει στην κατάσταση $(0, y)$.

Όταν η διαδικασία βρίσκεται στην κατάσταση (x, y) και ισχύει η ανισότητα $W(x, y) \le \min\{K + V(x, y+1), Lx\},$ τότε η βέλτιστη πολιτική επιλέγει την ενέργεια (i), δηλαδή δεν επεμβαίνει στην εξέλιξη της επιδημικής διαδικασίας.



Σχήμα 3.3. *Η βέλτιστη πολιτική όταν* $(N, K, L, \alpha, \beta, c_1, c_2) = (10, 0.6, 0.7, 1.1, 1.3, 0.8, 0.9).$

Ένα αριθμητικό παράδειγμα παρουσιάζεται παρακάτω. Θεωρούμε την περίπτωση κατά την οποία $N = 10, K = 0.6, L = 0.7, \alpha = 1.1, \beta = 1.3, c_1 = 0.8, c_2 = 0.9.$ Η βέλτιστη πολιτική για αυτές τις τιμές των παραμέτρων παρουσιάζεται στο παραπάνω Σχήμα 3.3, όπου για κάθε κατάσταση (x, y) της διαδικασίας για την οποία ισχύει ότι $0 < x + y \le 9,$ η ενέργεια (i) αναπαριστάται με “0”, η ενέργεια (ii) αναπαριστάται με “1” και η ενέργεια (iii) αναπαριστάται με “2”. Στο σχήμα παρατηρούμε ότι εμφανίζονται τρεις ξεχωριστές περιοχές οι οποίες αποτελούνται από “0”, “1” και “2”. Δεν υπάρχει κανένα “2” το οποίο βρίσκεται δεξιά από ένα “0” και δεν υπάρχει κανένα “1” ή κανένα “2” το οποίο βρίσκεται πάνω από ένα “0”.

Πολλά αριθμητικά παραδείγματα για διάφορες τιμές των παραμέτρων παρέχουν ισχυρή ένδειξη ότι η βέλτιστη πολιτική έχει πάντα αυτές τις ιδιότητες. Μία αναλυτική απόδειξη φαίνεται δύσκολη.

Σημειώνεται ότι η κατασκευή κατάλληλου αλγορίθμου του δυναμικού προγραμματισμού για το Πρόβλημα 2 δεν φαίνεται δυνατή.

3.3 Ντετερμινιστική επιδημική διαδικασία

Στην αντίστοιχη ντετερμινιστική επιδημική διαδικασία αναπαριστούμε με x και y, $0 \le x, y \le N$, τον αριθμό των ατόμων που έχουν προσβληθεί από τις ασθένειες 1 και 2, αντίστοιχα, τη χρονική στιγμή $t \ge 0$. Σε αντίθεση με το προηγούμενο εδάφιο στο οποίο οι αριθμοί x και y είναι δυνατόν να πάρουν μόνο ακέραιες τιμές, στο παρόν εδάφιο θεωρούμε ότι είναι συνεχείς μεταβλητές που ικανοποιούν τις ακόλουθες διαφορικές εξισώσεις:

$$\frac{dx}{dt} = c_1 x^{\alpha} (N - x - y), \quad \frac{dy}{dt} = c_2 y^{\beta} (N - x - y). \tag{3.7}$$

Η επιδημική διαδικασία σταματά όταν $x + y = N$. Όπως στην περίπτωση της στοχαστικής επιδημικής διαδικασίας του Εδαφίου 3.1, υποθέτουμε ότι η παρουσία ενός ατόμου που έχει προσβληθεί από τη σοβαρή ασθένεια 1 επιφέρει κάποιο κόστος στην κοινωνία το οποίο είναι σταθερό και ίσο με τη μονάδα.

Θεωρούμε πολιτικές οι οποίες σε κάθε χρονική στιγμή $t \ge 0$ μπορούν να ελέγξουν τη ντετερμινιστική διαδικασία είτε μέσω του εμβολιασμού με την ήπια ασθένεια 2 οποιουδήποτε αριθμού επιδεκτικών ατόμων έχουν απομείνει στον πληθυσμό και δεν έχουν προσβληθεί από καμία από τις δύο ασθένειες είτε μέσω της απομόνωσης οποιουδήποτε αριθμού ατόμων που έχουν προσβληθεί από τη σοβαρή ασθένεια 1. Ο εμβολιασμός ενός επιδεκτικού ατόμου με την ήπια ασθένεια 2 επιφέρει ένα κόστος ίσο με $K > 0$ και η απομόνωση ενός ατόμου που έχει προσβληθεί από τη σοβαρή ασθένεια 1 επιφέρει ένα κόστος ίσο με $L > 0$.

Στη συνέχεια θα παρουσιάσουμε αναλυτικές λύσεις για το Πρόβλήμα 1 στην περίπτωση κατά την οποία $\alpha \ne 1$ και $\beta = 1$ καθώς επίσης για το Πρόβλημα 2 στην περίπτωση κατά την οποία $\alpha = 1$ και $\beta \ne 1$.

Πρόβλημα 1, αν $\alpha \neq 1, \beta = 1.$

Υποθέτουμε ότι η κατάσταση (x_0, y_0) είναι η αρχική κατάσταση της διαδικασίας. Από τις εξισώσεις (3.7) προκύπτει ότι η διδιάστατη ντετερμινιστική επιδημική διαδικασία βρίσκεται στην καμπύλη:

$$y = y_0 \exp[c(1-\alpha)^{-1}(x^{1-\alpha} - x_0^{1-\alpha})], \quad x_0 \leq x \leq \xi(x_0, y_0),$$

όπου, $c = c_1^{-1}c_2$ και $\xi(x_0, y_0)$ είναι η μοναδική ρίζα της ακόλουθης εξίσωσης ως προς x στο διάστημα (x_0, N):

$$y_0 \exp[c(1-\alpha)^{-1}(x^{1-\alpha} - x_0^{1-\alpha})] + x - N = 0. \tag{3.8}$$

Αν η αρχική κατάσταση της διαδικασίας είναι η κατάσταση $(0, y_0)$ τέτοια ώστε $y_0 \neq 0,$ η επιδημική διαδικασία βρίσκεται στην ευθεία $x = 0, \; y_0 \leq y \leq N.$ Στην περίπτωση αυτή δεν χρειάζεται να επέμβουμε στην εξέλιξη της διαδικασίας.

Αν η αρχική κατάσταση της διαδικασίας είναι η κατάσταση $(x_0, 0)$ τέτοια ώστε $x_0 \neq 0,$ το κόστος $C(x_0, y_0)$ μέχρι το τερματισμό της διαδικασίας, αν επιλέξουμε να μην επέμβουμε ποτέ στην εξέλιξη της, θα είναι ίσο με τον αριθμό των ατόμων που έχουν προσβληθεί από τη σοβαρή ασθένεια 1. Επομένως θα ισχύει ότι:

$$C(x_0, y_0) = \xi(x_0, y_0) - x_0.$$

Υποθέτουμε ότι όταν η επιδημική διαδικασία βρίσκεται στην κατάσταση $(x_0, y_0),$ επιλέγουμε εκείνη την πολιτική σύμφωνα με την οποία, εμβολιάζουμε Δy_0 επιδεκτικά άτομα με την ήπια ασθένεια 2 και στη συνέχεια δεν επεμβαίνουμε περαιτέρω στην εξέλιξη της διαδικασίας.

Το κόστος αυτής της πολιτικής είναι ίσο με $K\Delta y_0 + C(x_0, y_0 + \Delta y_0),$ το οποίο είναι μικρότερο από $C(x_0, y_0)$ αν:

$$\frac{C(x_0, y_0 + \Delta y_0) - C(x_0, y_0)}{\Delta y_0} < -K.$$

Αν θεωρήσουμε ότι ο αριθμός Δy_0 είναι πολύ μικρός, η παραπάνω ανισότητα είναι κατά προσέγγιση ισοδύναμη με την ανισότητα:

$$\frac{\partial \xi(x_0, y_0)}{\partial y_0} < -K.$$

Τα παραπάνω μας οδηγούν να εικάσουμε ότι, όταν η διαδικασία βρίσκεται στην κατάσταση (x_0, y_0) και ισχύει ότι $\partial \xi(x_0, y_0)/\partial y_0 < -K,$ η βέλτιστη πολιτική εμβολιάζει με την ήπια ασθένεια 2 κάποια από τα επιδεκτικά άτομα που έχουν απομείνει στον πληθυσμό και δεν έχουν προσβληθεί από καμία από τις δύο ασθένειες, ενώ όταν η διαδικασία βρίσκεται στην κατάσταση (x_0, y_0) και ισχύει ότι $\partial \xi(x_0, y_0)/\partial y_0 \geq -K,$ η βέλτιστη πολιτική δεν επεμβαίνει στην εξέλιξη της διαδικασίας.

Η εικασία αποδεικνύεται με παρόμοιο τρόπο όπως στην εργασία του Kyriakidis (1995) ο οποίος απέδειξε την εικασία για την περίπτωση κατά την οποία $\alpha = \beta = 1.$ Στη συνέχεια θα παρουσιάσουμε με συνοπτικό τρόπο ένα βασικό σημείο της απόδειξης.

Έστω $P_\theta,\ \ 0 < \theta \leq N - x_0 - y_0,$ η πολιτική σύμφωνα με την οποία, όταν η επιδημική διαδικασία βρίσκεται στην κατάσταση (x_0, y_0) εμβολιάζουμε αμέσως θ επιδεκτικά άτομα με την ήπια ασθένεια 2 και δεν επεμβαίνουμε περαιτέρω στην εξέλιξη της διαδικασίας. Ορίζουμε τη συνάρτηση:

$$A(x) \equiv \xi(x, y_0 \exp[c(1-\alpha)^{-1}(x^{1-\alpha} - x_0^{1-\alpha})] + \theta), \quad x_0 \leq x \leq x_1,$$

όπου, (x_1, y_1) είναι μία κατάσταση της διαδικασίας τέτοια ώστε $x_1 > x_0$ και

$$y_1 = y_0 \exp[c(1-\alpha)^{-1}(x_1^{1-\alpha} - x_0^{1-\alpha})].$$

Η συνάρτηση $A(x)$ ικανοποιεί την εξίσωση:

$$\left(y_0 \exp[-c(1-\alpha)^{-1} x_0^{1-\alpha}] + \theta \exp[-c(1-\alpha)^{-1} x^{1-\alpha}]\right)\exp[c(1-\alpha)^{-1}(A(x))^{1-\alpha}] + A(x) - N = 0.$$

Αν παραγωγίσουμε την παραπάνω εξίσωση ως προς x παίρνουμε μία θετική έκφραση για την ποσότητα $dA(x)/dx.$ Επομένως η συνάρτηση $A(x)$ είναι αύξουσα ως προς x και ένα παρόμοιο αποτέλεσμα με αυτό του Λήμματος 4.1 στην εργασία του Kyriakidis (1995) μπορεί να αποδειχθεί. Σύμφωνα με το αποτέλεσμα αυτό, αν η αρχική κατάσταση της διαδικασίας είναι η κατάσταση $(x_0, y_0),\ x_0 \neq 0,$ ανάμεσα σε όλες τις πολιτικές οι οποίες επιλέγουν τον εμβολιασμό θ συνολικά επιδεκτικών ατόμων με την ήπια ασθένεια 2, η πολιτική P_θ είναι η βέλτιστη.

Το θεώρημα που παρουσιάζουμε παρακάτω αντιστοιχεί στην εικασία του Προβλήματος 1 για τη στοχαστική επιδημική διαδικασία του Εδαφίου 3.2 και είναι παρόμοιο με το Θεώρημα 4.3 στην εργασία του Kyriakidis (1995).

Θεώρημα 3.1. (α) Αν η αρχική κατάσταση της διαδικασίας είναι η κατάσταση $(x_0, y_0),\ x_0 \neq 0$ και ισχύει ότι $\partial\xi(x_0, y_0)/\partial y_0 \geq -K,$ τότε η βέλτιστη πολιτική δεν επεμβαίνει στην εξέλιξη της διαδικασίας.

(β) Υποθέτουμε ότι η αρχική κατάσταση της διαδικασίας είναι η κατάσταση $(x_0, y_0),\ x_0 \neq 0$ και ισχύει ότι $\partial\xi(x_0, y_0)/\partial y_0 < -K.$ Διακρίνουμε δύο περιπτώσεις:

Περίπτωση 1. Αν υπάρχει μία κατάσταση (x_0, y^*) τέτοια ώστε $[\partial\xi(x_0, y_0)/\partial y_0]_{(x_0, y^*)} = -K,$ τότε η πολιτική $P_{y^*-y_0}$ είναι βέλτιστη.

Περίπτωση 2. Αν δεν υπάρχει κατάσταση (x_0, y^*) τέτοια ώστε $[\partial\xi(x_0, y_0)/\partial y_0]_{(x_0, y^*)} = -K,$ τότε η πολιτική $P_{N-x_0-y_0}$ είναι βέλτιστη.

Αν θέσουμε $x = \xi(x_0, y_0)$ και παραγωγίσουμε την εξίσωση (3.8) ως προς y_0, παίρνουμε την ακόλουθη έκφραση για την ποσότητα $\partial\xi(x_0, y_0)/\partial y_0.$

$$\frac{\partial \xi(x_0, y_0)}{\partial y_0} = -\frac{\exp[c(1-\alpha)^{-1}(\xi(x_0, y_0))^{1-\alpha}]}{\exp[c(1-\alpha)^{-1}x_0^{1-\alpha}] + cy_0[\xi(x_0, y_0)]^{-\alpha}\exp[c(1-\alpha)^{-1}(\xi(x_0, y_0))^{1-\alpha}]} \quad (3.9)$$

Αν παραγωγίσουμε την εξίσωση (3.9) ως προς y_0 μπορούμε να επαληθεύσουμε ότι για κάθε x_0, $0 < x_0 < N$, η μερική παράγωγος $\partial \xi(x_0, y_0)/\partial y_0$, $0 \le y_0 \le N - x_0$, είναι αύξουσα ως προς y_0. Αν $[\partial \xi(x_0, y_0)/\partial y_0]_{(x_0,0)} \ge -K$, ή ισοδύναμα,

$$K \ge \exp[c(1-\alpha)^{-1}(N^{1-\alpha} - x_0^{1-\alpha})], \quad (3.10)$$

από τη μονοτονία της $\partial \xi(x_0, y_0)/\partial y_0$, προκύπτει ότι $\partial \xi(x_0, y_0)/\partial y_0 \ge -K$, για κάθε y_0 τέτοιο ώστε $0 \le y_0 \le N - x_0$. Έτσι, σύμφωνα με το Θεώρημα 3.1(α), αν η αρχική κατάσταση της διαδικασίας είναι η κατάσταση (x_0, y_0), $0 \le y_0 \le N - x_0$, η συνθήκη (3.10) συνεπάγεται ότι η βέλτιστη πολιτική δεν επεμβαίνει στην εξέλιξη της διαδικασίας.

Αν $[\partial \xi(x_0, y_0)/\partial y_0]_{(x_0,N-x_0)} < -K$, ή ισοδύναμα,

$$K < [1 + c(N - x_0)x_0^{-\alpha}]^{-1} \quad (3.11)$$

πάλι από τη μονοτονία της $\partial \xi(x_0, y_0)/\partial y_0$, προκύπτει ότι $\partial \xi(x_0, y_0)/\partial y_0 < -K$, για κάθε y_0 τέτοιο ώστε $0 \le y_0 \le N - x_0$. Έτσι, σύμφωνα με το Θεώρημα 3.1(β), αν η αρχική κατάσταση της διαδικασίας είναι η κατάσταση (x_0, y_0), $0 \le y_0 \le N - x_0$, η συνθήκη (3.11) συνεπάγεται ότι η πολιτική $P_{N-x_0-y_0}$ είναι βέλτιστη.

Στην περίπτωση κατά την οποία οι συνθήκες (3.10) και (3.11) δεν ισχύουν, η κρίσιμη τιμή y^* η οποία αναφέρεται στο Θεώρημα 3.1(β), ικανοποιεί την εξίσωση $[\partial \xi(x_0, y_0)/\partial y_0]_{(x_0,y^*)} = -K$ και μπορεί να υπολογιστεί αριθμητικά για κάθε x_0, $0 < x_0 < N$, με τη βοήθεια του παρακάτω αλγορίθμου.

Αλγόριθμος για τον υπολογισμό της κρίσιμης τιμής y^*

Βήμα 1. Θέτουμε $y = 0$.

Βήμα 2. Θέτουμε $y = y + \varepsilon,$ όπου ε είναι ένας μικρός θετικός αριθμός (π.χ. $\varepsilon = 0.001$).

Βήμα 3. Υπολογίζουμε τον αριθμό $\xi(x_0, y)$ λύνοντας αριθμητικά την εξίσωση (3.8) ως προς x.

Βήμα 4. Αν $[\partial\xi(x_0, y_0)/\partial y_0]_{(x_0,y)} \geq -K,$ τότε $y^* = y$. Διαφορετικά, επιστρέφουμε στο Βήμα 2.

Στο Βήμα 3 του αλγορίθμου, η εξίσωση (3.8) επιλύθηκε αριθμητικά χρησιμοποιώντας τη μέθοδο της διχοτόμησης.

Στην περίπτωση κατά την οποία η συνθήκη (3.10) ισχύει, πολλά αριθμητικά παραδείγματα παρέχουν ισχυρή ένδειξη ότι η βέλτιστη πολιτική της αντίστοιχης στοχαστικής επιδημικής διαδικασίας του Εδαφίου 3.2 δεν επεμβαίνει στην εξέλιξη της διαδικασίας.

Στην περίπτωση κατά την οποία η συνθήκη (3.11) ισχύει, πολλά αριθμητικά παραδείγματα παρέχουν ισχυρή ένδειξη ότι η κρίσιμη τιμή $\tilde{y}$ της αντίστοιχης στοχαστικής επιδημικής διαδικασίας του Εδαφίου 3.2 είναι ίση με $N - x_0 - 1$.

Στην περίπτωση κατά την οποία και οι δύο συνθήκες (3.10) και (3.11) δεν ισχύουν, πολλά αριθμητικά παραδείγματα παρέχουν ισχυρή ένδειξη ότι $\tilde{y} < y^*$. Έτσι οδηγούμαστε στην ενδιαφέρουσα εικασία σύμφωνα με την οποία η κρίσιμη τιμή $\tilde{y}$ της στοχαστικής επιδημικής διαδικασίας είναι πάντοτε μικρότερη από την αντίστοιχη κρίσιμη τιμή y^* της ντετερμινιστικής επιδημικής διαδικασίας. Η ίδια εικασία είχε εμφανιστεί και στην εργασία του Kyriakidis (1995) στην περίπτωση κατά την οποία $\alpha = \beta = 1$.

Θα παρουσιάσουμε ένα αριθμητικό παράδειγμα. Στον Πίνακα 3.1 παρουσιάζονται οι κρίσιμες τιμές $\tilde{y}$ και y^* όταν $N = 10, K = 1, \alpha = 2, \beta = 1, c_1 = 1.5, c_2 = 1,$ για κάθε x τέτοιο ώστε $0 < x \leq 9$. Το σύμβολο “$\times$” υποδηλώνει ότι ισχύει η Περίπτωση 1 της εικασίας του Προβλήματος 1 για τη στοχαστική επιδημική διαδικασία.

Πίνακας 3.1. Τιμές των $\tilde{y}$ και y^*

x	1	2	3	4	5	6	7	8	9
$\tilde{y}$	2	3	3	2	2	2	1	0	$\times$
y^*	4.56	4.71	4.15	3.49	2.83	2.21	1.61	1.05	0.51

Επισημαίνουμε ότι φαίνεται δύσκολο να επιλυθεί αναλυτικά το Πρόβλημα 1 για τη ντετερμινιστική επιδημική διαδικασία στην περίπτωση κατά την οποία $\alpha = 1$ και $\beta \neq 1$ ή στην περίπτωση κατά την οποία $\alpha \neq 1$ και $\beta \neq 1$.

Πρόβλημα 2, αν $\alpha = 1, \beta \neq 1$.

Υποθέτουμε ότι η κατάσταση (x_0, y_0) είναι η αρχική κατάσταση της διαδικασίας. Στο πρόβλημα αυτό είναι προτιμότερο να εκφράσουμε τη μεταβλητή x ως συνάρτηση της μεταβλητής y. Από τις εξισώσεις (3.7) προκύπτει ότι η διδιάστατη ντετερμινιστική επιδημική διαδικασία βρίσκεται στην καμπύλη:

$$x = x_0 \exp[\tilde{c}(1-\beta)^{-1}(y^{1-\beta} - y_0^{1-\beta})], \quad y_0 \leq y \leq \tilde{\xi}(x_0, y_0),$$

όπου, $\tilde{c} = c_1 c_2^{-1}$ και $\tilde{\xi}(x_0, y_0)$ είναι η μοναδική ρίζα της ακόλουθης εξίσωσης ως προς y στο διάστημα (y_0, N):

$$x_0 \exp[\tilde{c}(1-\beta)^{-1}(y^{1-\beta} - y_0^{1-\beta})] + y - N = 0. \tag{3.12}$$

Αρχικά θεωρούμε δύο ειδικές περιπτώσεις. Αν η αρχική κατάσταση της διαδικασίας είναι η κατάσταση $(0, y_0)$, $0 < y_0 < N$, τότε η επιδημική διαδικασία βρίσκεται στην ευθεία $x = 0$, $y_0 \leq y \leq N$. Στην περίπτωση αυτή το κόστος της διαδικασίας θα είναι ίσο με μηδέν διότι η παρουσία ενός ατόμου που έχει προσβληθεί από την ήπια ασθένεια 2 δεν επιφέρει κάποιο κόστος στην κοινωνία.

Αν η αρχική κατάσταση της διαδικασίας είναι η κατάσταση $(x_0, 0)$, $0 < x_0 < N$, τότε η επιδημική διαδικασία βρίσκεται στην ευθεία $y = 0$, $x_0 \leq x \leq N$. Στην περίπτωση αυτή μπορούμε να επαληθεύσουμε εύκολα ότι αν $L < N x_0^{-1} - 1$, τότε η βέλτιστη πολιτική επιλέγει στην αρχική κατάσταση της διαδικασίας και μέχρι το τερματισμό της, την απομόνωση όλων των ατόμων που έχουν προσβληθεί από τη σοβαρή ασθένεια 1. Αν αντιθέτως ισχύει ότι $L \geq N x_0^{-1} - 1$, τότε η βέλτιστη πολιτική δεν επεμβαίνει στην εξέλιξη της διαδικασίας.

Αν η αρχική κατάσταση της επιδημικής διαδικασίας είναι η κατάσταση (x_0, y_0), τέτοια ώστε $x_0 \neq 0$ και $y_0 \neq 0$, τότε το κόστος $\widetilde{C}(x_0, y_0)$ μέχρι το τερματισμό της διαδικασίας, αν επιλέξουμε να μην επέμβουμε ποτέ στην εξέλιξη της, είναι:

$$\widetilde{C}(x_0, y_0) = N - \widetilde{\xi}(x_0, y_0) - x_0.$$

Υποθέτουμε ότι όταν η επιδημική διαδικασία βρίσκεται στην κατάσταση (x_0, y_0), επιλέγουμε εκείνη την πολιτική σύμφωνα με την οποία, απομονώνουμε Δx_0 άτομα που έχουν προσβληθεί από τη σοβαρή ασθένεια 1 και στη συνέχεια δεν επεμβαίνουμε περαιτέρω στην εξέλιξη της διαδικασίας.

Το κόστος αυτής της πολιτικής είναι ίσο με $L\Delta x_0 + \widetilde{C}(x_0 - \Delta x_0, y_0)$, το οποίο είναι μικρότερο από $\widetilde{C}(x_0, y_0)$ αν:

$$\frac{\widetilde{C}(x_0 - \Delta x_0, y_0) - \widetilde{C}(x_0, y_0)}{\Delta x_0} < -L.$$

Αν θεωρήσουμε ότι ο αριθμός Δx_0 είναι πολύ μικρός, η παραπάνω ανισότητα είναι κατά προσέγγιση ισοδύναμη με την ανισότητα:

$$\frac{\partial \widetilde{C}(x_0, y_0)}{\partial x_0} > L.$$

Τα παραπάνω μας οδηγούν να εικάσουμε ότι, όταν η διαδικασία βρίσκεται στην κατάσταση (x_0, y_0) και ισχύει ότι $\partial \widetilde{C}(x_0, y_0) / \partial x_0 > L$, η βέλτιστη πολιτική απομονώνει κάποια ή όλα τα άτομα που έχουν προσβληθεί από τη σοβαρή ασθένεια 1.

Η εικασία αποδεικνύεται με παρόμοιο τρόπο όπως στην εργασία του Kyriakidis (1999c) ο οποίος απέδειξε την εικασία για την περίπτωση κατά την οποία $\alpha = \beta = 1$. Στη συνέχεια θα παρουσιάσουμε με συνοπτικό τρόπο ένα βασικό σημείο της απόδειξης.

Έστω P_θ, $0 < \theta \le N - x_0 - y_0$, η πολιτική σύμφωνα με την οποία, όταν η επιδημική διαδικασία βρίσκεται στην κατάσταση (x_0, y_0) απομονώνουμε αμέσως θ άτομα που έχουν προσβληθεί από τη σοβαρή ασθένεια 1 και δεν επεμβαίνουμε περαιτέρω στην εξέλιξη της διαδικασίας.

Ορίζουμε τη συνάρτηση:

$$A(y) \equiv \tilde{\xi}(x_0 \exp[\tilde{c}(1-\beta)^{-1}(y^{1-\beta} - y_0^{1-\beta})] - \theta, y), \quad y_0 \le y \le y_1,$$

όπου, (x_1, y_1) είναι μία κατάσταση της διαδικασίας τέτοια ώστε $y_1 > y_0$ και

$$x_1 = x_0 \exp[\tilde{c}(1-\beta)^{-1}(y_1^{1-\beta} - y_0^{1-\beta})].$$

Η συνάρτηση $A(y)$ ικανοποιεί την εξίσωση:

$$\left(x_0 \exp[-\tilde{c}(1-\beta)^{-1} y_0^{1-\beta}] - \theta \exp[-\tilde{c}(1-\beta)^{-1} y^{1-\beta}]\right) \exp[\tilde{c}(1-\beta)^{-1}(A(y))^{1-\beta}] + A(y) - N = 0.$$

Αν παραγωγίσουμε την παραπάνω εξίσωση ως προς y παίρνουμε μία αρνητική έκφραση για την ποσότητα $dA(y)/dy$. Επομένως η συνάρτηση $A(y)$ είναι φθίνουσα ως προς y και ένα παρόμοιο αποτέλεσμα με αυτό του Λήμματος 1 στην εργασία του Kyriakidis (1999c) μπορεί να αποδειχθεί. Σύμφωνα με το αποτέλεσμα αυτό, αν η αρχική κατάσταση της διαδικασίας είναι η κατάσταση (x_0, y_0), $x_0 \ne 0$, $y_0 \ne 0$, ανάμεσα σε όλες τις πολιτικές οι οποίες απομονώνουν θ συνολικά άτομα που έχουν προσβληθεί από τη σοβαρή ασθένεια 1, η πολιτική P_θ είναι η βέλτιστη.

Το θεώρημα που παρουσιάζουμε παρακάτω αντιστοιχεί στην εικασία του Προβλήματος 3 για τη στοχαστική επιδημική διαδικασία του Εδαφίου 3.2 και είναι παρόμοιο με το Θεώρημα 5 στην εργασία του Kyriakidis (1999c).

Θεώρημα 3.2. Έστω ότι η αρχική κατάσταση της διαδικασίας είναι η κατάσταση (x_0, y_0), $0 < y_0 < N.$ Αν

$$L + 1 \geq \exp[\widetilde{c}(1-\beta)^{-1}(N^{1-\beta} - y_0^{1-\beta})] \tag{3.13}$$

τότε για κάθε $x_0, 0 < x_0 < N - y_0$, η βέλτιστη πολιτική δεν επεμβαίνει στην εξέλιξη της διαδικασίας. Αν η συνθήκη (3.13) δεν ισχύει, τότε υπάρχει μία κρίσιμη τιμή x^*, $0 < x^* < N - y_0$, τέτοια ώστε όταν $0 < x_0 \leq x^*$, η βέλτιστη πολιτική είναι η πολιτική P_{x_0} και όταν $x^* < x_0 < N - y_0$, η βέλτιστη πολιτική δεν επεμβαίνει στην εξέλιξη της διαδικασίας. Η κρίσιμη τιμή x^*, $0 < x^* < N - y_0$, ικανοποιεί την εξίσωση: $\widetilde{C}(x^*, y_0)/x^* = L.$

Για κάθε y_0, $0 < y_0 < N$, μπορούμε να επαληθεύσουμε ότι η μερική παράγωγος του πηλίκου $\widetilde{C}(x_0, y_0)/x_0$ ως προς x_0 είναι αρνητική. Συνεπώς, για κάθε y_0, $0 < y_0 < N$, έπεται ότι η ποσότητα $\widetilde{C}(x_0, y_0)/x_0$ είναι φθίνουσα ως προς x_0, $0 < x_0 < N - y_0$.

Στην περίπτωση κατά την οποία η συνθήκη (3.13) δεν ισχύει, η κρίσιμη τιμή x^*, $0 < x^* < N - y_0$, ικανοποιεί την ακόλουθη εξίσωση:

$$\frac{\widetilde{C}(x^*, y_0)}{x^*} = \frac{N - \widetilde{\xi}(x^*, y_0)}{x^*} - 1 = L,$$

και μπορεί να υπολογιστεί αριθμητικά για κάθε y_0, $0 < y_0 < N$, με τη βοήθεια του παρακάτω αλγορίθμου.

Αλγόριθμος για τον υπολογισμό της κρίσιμης τιμής x^*

Βήμα 1. Θέτουμε $x = 0.$

Βήμα 2. Θέτουμε $x = x + \varepsilon$, όπου ε είναι ένας μικρός θετικός αριθμός (π.χ. $\varepsilon = 0.001$).

Βήμα 3. Υπολογίζουμε τον αριθμό $\tilde{\xi}(x, y_0)$ λύνοντας αριθμητικά την εξίσωση (3.12) ως προς y.

Βήμα 4. Αν $\dfrac{\tilde{C}(x, y_0)}{x} = \dfrac{N - \tilde{\xi}(x, y_0)}{x} - 1 \leq L,$ τότε $x^* = x.$ Διαφορετικά, επιστρέφουμε στο Βήμα 2.

Στο Βήμα 3 του αλγορίθμου, η εξίσωση (3.12) επιλύθηκε αριθμητικά χρησιμοποιώντας τη μέθοδο της διχοτόμησης.

Στην περίπτωση κατά την οποία η συνθήκη (3.13) ισχύει και η αρχική κατάσταση της διαδικασίας είναι η κατάσταση (x_0, y_0), $0 < x_0 \leq N - y_0$, πολλά αριθμητικά παραδείγματα παρέχουν ισχυρή ένδειξη ότι η βέλτιστη πολιτική της αντίστοιχης στοχαστικής επιδημικής διαδικασίας του Εδαφίου 3.2 δεν επεμβαίνει στην εξέλιξη της διαδικασίας.

Στην περίπτωση κατά την οποία η συνθήκη (3.13) δεν ισχύει, πολλά αριθμητικά παραδείγματα παρέχουν ισχυρή ένδειξη ότι η κρίσιμη τιμή $\tilde{x}$ της στοχαστικής επιδημικής διαδικασίας είναι πάντοτε μικρότερη από την αντίστοιχη κρίσιμη τιμή x^* της ντετερμινιστικής επιδημικής διαδικασίας. Η ίδια εικασία είχε εμφανιστεί και στην εργασία του Kyriakidis (1999c) στην περίπτωση κατά την οποία $\alpha = \beta = 1$.

Θα παρουσιάσουμε ένα αριθμητικό παράδειγμα. Στον Πίνακα 3.2 παρουσιάζονται οι κρίσιμες τιμές $\tilde{x}$ και x^* όταν $N = 10, L = 0.6, \alpha = 1, \beta = 0.5, c_1 = 0.8, c_2 = 1.2,$ για κάθε y τέτοιο ώστε $0 \leq y < 9$. Το σύμβολο "$\times$" υποδηλώνει ότι ισχύει η Περίπτωση 1 της εικασίας του Προβλήματος 3 για τη στοχαστική επιδημική διαδικασία και το σύμβολο "–" υποδηλώνει ότι ισχύει η συνθήκη (3.13).

Πίνακας 3.2. Τιμές των $\tilde{x}$ και x^*

y	0	1	2	3	4	5	6	7	8
$\tilde{x}$	6	5	4	3	2	1	1	$\times$	$\times$
x^*	6.25	5.1	4.3	3.53	2.79	2.06	1.34	0.63	–

Επισημαίνουμε ότι φαίνεται δύσκολο να επιλυθεί αναλυτικά το Πρόβλημα 2 για τη ντετερμινιστική επιδημική διαδικασία στην περίπτωση κατά την οποία $\alpha \neq 1$ και $\beta = 1$ ή στην περίπτωση κατά την οποία $\alpha \neq 1$ και $\beta \neq 1$.

3.4 Αβεβαιότητα στις τιμές των παραμέτρων

Στο παρόν εδάφιο για τη στοχαστική επιδημική διαδικασία που περιγράψαμε στο Εδάφιο 3.1, θεωρούμε ότι η ποσότητα $c_1^{-1}c_2$ δεν είναι ένας σταθερός θετικός πραγματικός αριθμός αλλά μία τυχαία μεταβλητή που ακολουθεί μία γνωστή κατανομή. Η κατανομή χρησιμοποιείται για να αναπαραστήσει την «αβεβαιότητα» στις τιμές των παραμέτρων c_1 και c_2. Οι κατανομές που έχουν χρησιμοποιηθεί είναι η κατανομή Βήτα, η κατανομή Γάμα καθώς επίσης η Εκθετική κατανομή.

Θεωρούμε τα Προβλήματα 1 και 3 που περιγράψαμε στο Εδάφιο 3.2 με τις παρακάτω τροποποιήσεις.

Πρόβλημα 1. Όταν επιλέγουμε την ενέργεια (i) σε κάθε κατάσταση (x, y) της διαδικασίας τέτοια ώστε $x + y < N$:

(i) αφήνουμε τη διαδικασία να μεταβεί στις καταστάσεις $(x+1, y)$ και $(x, y+1)$, με πιθανότητες

$$p_{xy} = E\left(\frac{x^{\alpha}}{x^{\alpha} + c_1^{-1}c_2 y^{\beta}}\right), \tag{3.14}$$

$$q_{xy} = E\left(\frac{y^{\beta}}{c_1 c_2^{-1} x^{\alpha} + y^{\beta}}\right), \tag{3.15}$$

και κόστος ίσο με 1 και 0, αντίστοιχα, όπου E αναπαριστά την αναμενόμενη τιμή.

Η εξίσωση (3.4) παίρνει τώρα τη μορφή:

$$W(x, y) = p_{xy}[1 + V(x+1, y)] + q_{xy}V(x, y+1), \quad 0 < x + y < N. \tag{3.16}$$

Οι εξισώσεις (3.3), (3.16) και (3.5) μας επιτρέπουν να υπολογίσουμε αριθμητικά το ελάχιστο αναμενόμενο κόστος $V(x, y)$ για κάθε κατάσταση (x, y) της διαδικασίας για την οποία ισχύει ότι $0 < x + y < N.$ Επιπλέον προσδιορίζουν την ενέργεια που επιλέγεται από τη βέλτιστη πολιτική για κάθε κατάσταση $(x, y), 0 < x + y < N.$

Στο αριθμητικό παράδειγμα του Εδαφίου 3.2 για το Πρόβλημα 1 στο οποίο $c_1 = 1.5,$ $c_2 = 1$ και επομένως $c_1^{-1}c_2 \cong 0.6,$ θεωρούμε ότι η τυχαία μεταβλητή $c_1^{-1}c_2$ ακολουθεί την κατανομή Βήτα με αναμενόμενη τιμή κατά προσέγγιση ίση με 0.6. Για κάθε κατάσταση (x, y) της διαδικασίας για την οποία ισχύει ότι $0 < x + y \leq 9,$ βρέθηκε ότι η βέλτιστη πολιτική έχει την ίδια μορφή με αυτή του Σχήματος 3.1.

Ένα ακόμη αριθμητικό παράδειγμα παρουσιάζεται παρακάτω. Θεωρούμε την περίπτωση κατά την οποία $N = 10, K = 0.5, \alpha = 3, \beta = 2.$ Στον Πίνακα 3.3 παρουσιάζονται, για κάθε $x,$ τέτοιο ώστε $0 < x \leq 9,$ η κρίσιμη τιμή $\tilde{y}$ της επιδημικής διαδικασίας του Εδαφίου 3.2 στην περίπτωση κατά την οποία $c_1 = 0.7,$ $c_2 = 1.3$ και επομένως $c_1^{-1}c_2 \cong 1.86$ και η κρίσιμη τιμή $\hat{y}$ της επιδημικής διαδικασίας στην περίπτωση κατά την οποία η τυχαία μεταβλητή $c_1^{-1}c_2$ ακολουθεί την Εκθετική κατανομή με αναμενόμενη τιμή κατά προσέγγιση ίση με 1.86.

Πίνακας 3.3. Τιμές των $\tilde{y}$ και $\hat{y}$

x	1	2	3	4	5	6	7	8	9
$\tilde{y}$	1	3	4	5	4	3	2	1	0
$\hat{y}$	2	4	5	5	4	3	2	1	0

Στον παραπάνω πίνακα παρατηρούμε ότι για κάθε $x,$ ισχύει ότι $\hat{y} \geq \tilde{y}.$ Πολλά αριθμητικά παραδείγματα στα οποία, η τυχαία μεταβλητή $c_1^{-1}c_2$ ακολουθεί την κατανομή Βήτα, την κατανομή Γάμα και την Εκθετική κατανομή, παρέχουν ισχυρή ένδειξη ότι για κάθε $x,$ η κρίσιμη τιμή $\hat{y}$ της επιδημικής διαδικασίας είναι πάντοτε μεγαλύτερη ή ίση με την κρίσιμη τιμή $\tilde{y}$ της αντίστοιχης επιδημικής διαδικασίας του Εδαφίου 3.2.

Πρόβλημα 3. Όταν επιλέγουμε την ενέργεια (i) σε κάθε κατάσταση (x, y) της διαδικασίας τέτοια ώστε $x + y < N$:

(i) αφήνουμε τη διαδικασία να μεταβεί στις καταστάσεις $(x+1, y)$ και $(x, y+1)$, με πιθανότητες οι οποίες δίνονται από τις σχέσεις (3.14) και (3.15) και κόστος ίσο με 1 και 0, αντίστοιχα.

Οι εξισώσεις (3.6), (3.16) και (3.5) μας επιτρέπουν να υπολογίσουμε αριθμητικά το ελάχιστο αναμενόμενο κόστος $V(x, y)$ για κάθε κατάσταση (x, y) της διαδικασίας για την οποία ισχύει ότι $0 < x + y < N$. Επιπλέον προσδιορίζουν την ενέργεια που επιλέγεται από τη βέλτιστη πολιτική για κάθε κατάσταση $(x, y), 0 < x + y < N$.

Στο αριθμητικό παράδειγμα του Εδαφίου 3.2 για το Πρόβλημα 3 στο οποίο $c_1 = 0.8,\ c_2 = 1.2$ και επομένως $c_1^{-1}c_2 = 1.5,$ θεωρούμε ότι η τυχαία μεταβλητή $c_1^{-1}c_2$ ακολουθεί την κατανομή Γάμα με αναμενόμενη τιμή ίση με 1.5. Για κάθε κατάσταση (x, y) της διαδικασίας για την οποία ισχύει ότι $0 < x + y \leq 9,$ βρέθηκε ότι η βέλτιστη πολιτική έχει την ίδια μορφή με αυτή του Σχήματος 3.2.

Ένα ακόμη αριθμητικό παράδειγμα παρουσιάζεται παρακάτω. Θεωρούμε την περίπτωση κατά την οποία $N = 10, L = 0.3, \alpha = 5, \beta = 3.$ Στον Πίνακα 3.4 παρουσιάζονται, για κάθε $y,$ τέτοιο ώστε $0 \leq y < 9,$ η κρίσιμη τιμή $\tilde{x}$ της επιδημικής διαδικασίας του Εδαφίου 3.2 στην περίπτωση κατά την οποία $c_1 = 0.7,\ c_2 = 1.1$ και επομένως $c_1^{-1}c_2 \cong 1.57$ και η κρίσιμη τιμή $\hat{x}$ της επιδημικής διαδικασίας στην περίπτωση κατά την οποία η τυχαία μεταβλητή $c_1^{-1}c_2$ ακολουθεί την Εκθετική κατανομή με αναμενόμενη τιμή κατά προσέγγιση ίση με 1.57. Το σύμβολο “×” υποδηλώνει ότι ισχύει η Περίπτωση 1 της εικασίας του Προβλήματος 3.

Πίνακας 3.4. Τιμές των $\tilde{x}$ και $\hat{x}$

y	0	1	2	3	4	5	6	7	8
$\tilde{x}$	7	6	5	4	3	1	×	×	×
$\hat{x}$	7	6	6	4	3	2	1	×	×

Στον παραπάνω πίνακα παρατηρούμε ότι για κάθε y, ισχύει ότι $\hat{x} \geq \tilde{x}$. Πολλά αριθμητικά παραδείγματα στα οποία η τυχαία μεταβλητή $c_1^{-1}c_2$ ακολουθεί την κατανομή Βήτα, την κατανομή Γάμα και την Εκθετική κατανομή παρέχουν ισχυρή ένδειξη ότι για κάθε y, η κρίσιμη τιμή $\hat{x}$ της επιδημικής διαδικασίας είναι πάντοτε μεγαλύτερη ή ίση με την κρίσιμη τιμή $\tilde{x}$ της αντίστοιχης επιδημικής διαδικασίας του Εδαφίου 3.2.

ΚΕΦΑΛΑΙΟ 4

Βέλτιστη προληπτική συντήρηση ενός συστήματος παραγωγής

4.1 Εισαγωγή

Κατά τη διάρκεια των τελευταίων τριών δεκαετιών πολλοί ερευνητές έχουν ασχοληθεί με το πρόβλημα της βέλτιστης συντήρησης ή της αντικατάστασης ενός συστήματος του οποίου η κατάσταση επιδεινώνεται με την πάροδο του χρόνου. Σε αρκετές σχετικές εργασίες κατάλληλα μοντέλα έχουν κατασκευαστεί και έχουν αναλυθεί με τη βοήθεια της θεωρίας των Μαρκοβιανών διαδικασιών αποφάσεων. Οι εργασίες των Valdez-Flores and Feldman (1989), Scarf (1997) και Wang (2002) παρέχουν μία ανασκόπηση της έρευνας που έχει διενεργηθεί για την αντιμετώπιση και την επίλυση του προβλήματος.

Το πρόβλημα σχετίζεται με ένα σύστημα του οποίου η κατάσταση επιδεινώνεται εξαιτίας της διαρκούς λειτουργίας του. Το σύστημα επιθεωρείται σε κάποιες χρονικές στιγμές και η κατάστασή του παρατηρείται σε κάθε χρονική στιγμή επιθεώρησης κατά την οποία μία ενέργεια επιλέγεται από ένα σύνολο εναλλακτικών ενεργειών. Μία ενέργεια η οποία μπορεί να ελέγξει την κατάσταση του συστήματος είναι εκείνη η οποία θέτει σε λειτουργία μία προληπτική συντήρησή του. Μία άλλη ενέργεια είναι εκείνη η οποία θέτει σε λειτουργία μία επισκευή του. Μία ακόμη ενέργεια είναι εκείνη η οποία αντικαθιστά το σύστημα από ένα καινούργιο σύστημα. Η λειτουργία, η προληπτική συντήρηση, η επισκευή και η αντικατάσταση του συστήματος επιφέρουν αντίστοιχα κόστη. Μία πολιτική είναι ένας κανόνας με τον οποίον επιλέγονται οι ενέργειες σε κάθε χρονική στιγμή επιθεώρησης. Στις περισσότερες περιπτώσεις το πρόβλημα είναι η εύρεση της πολιτικής η οποία, για οποιαδήποτε αρχική κατάσταση του συστήματος, ελαχιστοποιεί το μακροπρόθεσμο αναμενόμενο μέσο κόστος ανά μονάδα χρόνου.

Πολλές είναι εκείνες οι εργασίες (βλέπε π.χ. Berg (1976), Federgruen and So (1989), Ozekici and Gunlunk (1992), So (1992), Vanneste (1992), Douer and Yechiali (1994), Chen and Feldman (1997), Love et al. (2000)) στις οποίες έχει αποδειχθεί ότι η βέλτιστη πολιτική θέτει σε λειτουργία μία συντήρηση ή αντικαθιστά ένα σύστημα αν και μόνο αν ο βαθμός επιδείνωσης του συστήματος είναι μεγαλύτερος ή ίσος με μία κρίσιμη τιμή. Μία τέτοια πολιτική είναι γνωστή ως μονότονη πολιτική.

Οι Van der Duyn Schouten and Vanneste (1995) εισήγαγαν ένα Μαρκοβιανό μοντέλο αποφάσεων για τη βέλτιστη προληπτική συντήρηση ενός συστήματος παραγωγής. Το σύστημα αποτελείται από ένα μηχανισμό, μία μονάδα παραγωγής και έναν ενδιάμεσο αποθηκευτικό χώρο. Ο μηχανισμός τροφοδοτεί τη μονάδα παραγωγής με ένα ακατέργαστο υλικό. Ο αποθηκευτικός χώρος του ακατέργαστου υλικού έχει τοποθετηθεί ανάμεσα στο μηχανισμό και στη μονάδα παραγωγής ώστε να διευκολυνθεί η παραγωγική διαδικασία. Οι Van der Duyn Schouten and Vanneste υπέθεσαν ότι, σε κάθε χρονική στιγμή επιθεώρησης του μηχανισμού είτε δεν γίνεται καμία επέμβαση στη λειτουργία του είτε τίθεται σε λειτουργία μία προληπτική συντήρηση ή μία επισκευή του. Υπέθεσαν επίσης ότι, αν σε μία χρονική στιγμή επιθεώρησης δεν γίνεται καμία επέμβαση στη λειτουργία του μηχανισμού, το σύστημα παραγωγής μεταβαίνει από μία κατάσταση σε μία άλλη κατάσταση με μία πιθανότητα η οποία εξαρτάται μόνο από την ηλικία του μηχανισμού. Θεώρησαν ότι, αν κατά τη διάρκεια μιας προληπτικής συντήρησης ή μιας επισκευής του μηχανισμού ο αποθηκευτικός χώρος αδειάσει, η έλλειψη του ακατέργαστου υλικού επιφέρει ένα κόστος για κάθε μονάδα χρόνου κατά την οποία ο αποθηκευτικός χώρος είναι κενός. Θεώρησαν επίσης ότι, οι απαιτούμενοι χρόνοι για μία προληπτική συντήρηση και μία επισκευή του μηχανισμού είναι τυχαίες μεταβλητές οι οποίες ακολουθούν τη Γεωμετρική κατανομή.

Μία κατάλληλη Μαρκοβιανή διαδικασία αποφάσεων σε διακριτό χρόνο με πεπερασμένο χώρο καταστάσεων χρησιμοποιήθηκε για να περιγράψει την κατάσταση του συστήματος παραγωγής σε κάθε χρονική στιγμή επιθεώρησης. Η κατάσταση του συστήματος παραγωγής αποτελείται από την ηλικία του μηχανισμού και από το περιεχόμενο του αποθηκευτικού χώρου. Ορίστηκαν κατάλληλες συνθήκες οι οποίες σχετίζονται με το κόστος της έλλειψης του ακατέργαστου υλικού στον αποθηκευτικό χώρο κατά τη διάρκεια μιας προληπτικής συντήρησης ή μιας επισκευής του μηχανισμού, με τους αναμενόμενους χρόνους της προληπτικής συντήρησης και τους αναμενόμενους χρόνους της επισκευής του μηχανισμού. Για κάθε σταθερό περιεχόμενο του αποθηκευτικού χώρου αποδείχθηκε ότι η βέλτιστη πολιτική είναι μονότονη διότι θέτει σε λειτουργία μία προληπτική συντήρηση αν και μόνο αν η ηλικία του μηχανισμού είναι μεγαλύτερη ή ίση με μία κρίσιμη τιμή.

Στο παρόν κεφάλαιο γενικεύουμε το μοντέλο των Van der Duyn Schouten and Vanneste (1995). Υποθέτουμε ότι, αν σε μία χρονική στιγμή επιθεώρησης δεν γίνεται καμία επέμβαση στη λειτουργία του μηχανισμού, το σύστημα παραγωγής μεταβαίνει την επόμενη χρονική στιγμή επιθεώρησης σε μία άλλη κατάσταση με μία πιθανότητα η οποία εξαρτάται από το βαθμό επιδείνωσης του μηχανισμού και από την ηλικία του μηχανισμού. Θεωρούμε ότι η επιδείνωση του μηχανισμού δεν είναι στάσιμη διότι οι πιθανότητες μετάβασης εξαρτώνται από την ηλικία του μηχανισμού. Η έννοια της «μη-στάσιμης» επιδείνωσης επινοήθηκε από τους Benyamini and Yechiali (1999). Θεωρούμε ότι η λειτουργία του μηχανισμού, μία προληπτική συντήρησή του ή μία επισκευή του επιφέρουν αντίστοιχα κόστη για κάθε μονάδα χρόνου κατά την οποία ο μηχανισμός λειτουργεί, συντηρείται προληπτικά ή επισκευάζεται. Υποθέτουμε ακόμη ότι υπάρχει ένα κόστος αποθήκευσης του ακατέργαστου υλικού στον αποθηκευτικό χώρο και ένα κόστος όταν ο αποθηκευτικός χώρος είναι κενός.

Μία κατάλληλη Μαρκοβιανή διαδικασία αποφάσεων σε διακριτό χρόνο με άπειρο χώρο καταστάσεων χρησιμοποιείται για να περιγράψει την κατάσταση του συστήματος παραγωγής σε κάθε χρονική στιγμή επιθεώρησης. Η κατάσταση του συστήματος παραγωγής αποτελείται από το βαθμό επιδείνωσης του μηχανισμού, από την ηλικία του μηχανισμού και από το περιεχόμενο του αποθηκευτικού χώρου. Ορίζονται κατάλληλες συνθήκες οι οποίες αφορούν τα κόστη της λειτουργίας, τα κόστη της προληπτικής συντήρησης, τα κόστη της επισκευής, τις πιθανότητες μετάβασης, τους αναμενόμενους χρόνους της προληπτικής συντήρησης και τους αναμενόμενους χρόνους της επισκευής του μηχανισμού. Για κάθε σταθερό περιεχόμενο του αποθηκευτικού χώρου και κάθε σταθερή ηλικία του μηχανισμού, αποδεικνύεται ότι η βέλτιστη πολιτική είναι μονότονη διότι θέτει σε λειτουργία μία προληπτική συντήρηση αν και μόνο αν ο βαθμός επιδείνωσης του μηχανισμού είναι μεγαλύτερος ή ίσος με μία κρίσιμη τιμή.

Δύο παρόμοια μοντέλα μελετήθηκαν από τους Meller and Kim (1996) και Salameh and Ghattas (2001). Οι Meller and Kim θεώρησαν ένα σύστημα παραγωγής το οποίο αποτελείται από δύο μηχανές και έναν ενδιάμεσο αποθηκευτικό χώρο. Υπέθεσαν ότι ο χρόνος μέχρι η πρώτη μηχανή να υποστεί μία βλάβη είναι μία τυχαία μεταβλητή η οποία ακολουθεί την Εκθετική κατανομή. Θεώρησαν ότι μία προληπτική συντήρηση της πρώτης μηχανής είναι δυνατή όταν το περιεχόμενο του αποθηκευτικού χώρου ισούται με μία κρίσιμη τιμή. Όρισαν κατάλληλα κόστη και επινόησαν ένα μοντέλο για το οποίο το μέσο κόστος υπολογίστηκε ως μία συνάρτηση της κρίσιμης τιμής. Οι Salameh and Ghattas θεώρησαν μία μονάδα παραγωγής με έναν αποθηκευτικό χώρο. Υπέθεσαν ότι η μονάδα παράγει ένα υλικό και στο τέλος μιας προκαθορισμένης χρονικής περιόδου διακόπτει τη λειτουργία της για να συντηρηθεί προληπτικά. Θεώρησαν ότι μία προληπτική συντήρηση της μονάδας παραγωγής διαρκεί τυχαίο χρόνο. Καθόρισαν την τιμή του περιεχομένου του αποθηκευτικού χώρου η οποία ελαχιστοποιεί το άθροισμα του κόστους της αποθήκευσης του υλικού ανά μονάδα χρόνου και του κόστους της έλλειψης του υλικού ανά μονάδα χρόνου.

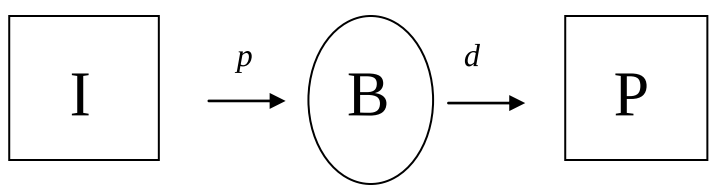
Στο επόμενο εδάφιο περιγράφουμε το μοντέλο και παρουσιάζουμε τις συνθήκες που αφορούν τα κόστη, τις πιθανότητες μετάβασης, τους αναμενόμενους χρόνους της προληπτικής συντήρησης και τους αναμενόμενους χρόνους της επισκευής του μηχανισμού. Στο Εδάφιο 4.3 αποδεικνύουμε αναλυτικά την ύπαρξη και τη μονότονη μορφή της βέλτιστης πολιτικής. Στο υπόλοιπο του κεφαλαίου μελετούμε την ειδική περίπτωση της στάσιμης επιδείνωσης του μηχανισμού. Στην περίπτωση αυτή οι πιθανότητες μετάβασης εξαρτώνται μόνο από το βαθμό επιδείνωσης του μηχανισμού και δεν εξαρτώνται από την ηλικία του μηχανισμού. Στο Εδάφιο 4.4 επισημαίνεται ότι, όταν η επιδείνωση του μηχανισμού είναι στάσιμη, για κάθε σταθερό περιεχόμενο του αποθηκευτικού χώρου, η βέλτιστη πολιτική είναι μονότονη διότι θέτει σε λειτουργία μία προληπτική συντήρηση αν και μόνο αν ο βαθμός επιδείνωσης του μηχανισμού είναι μεγαλύτερος ή ίσος με μία κρίσιμη τιμή. Στο Εδάφιο 4.5, για την περίπτωση της στάσιμης επιδείνωσης του μηχανισμού, σχεδιάζεται ένας κατάλληλος αλγόριθμος βελτίωσης των πολιτικών ο οποίος εφαρμόζεται στο σύνολο των μονότονων πολιτικών και παράγει μία ακολουθία αυστηρά βελτιωμένων πολιτικών που έχουν τη μονότονη μορφή. Ένα αριθμητικό παράδειγμα του αλγορίθμου επίσης παρουσιάζεται. Στο Εδάφιο 4.6 δύο γενικεύσεις του μοντέλου του Εδαφίου 4.4 μελετώνται και δύο αριθμητικά παραδείγματα παρουσιάζονται. Στο Εδάφιο 4.7 τροποποιούμε το μοντέλο του Εδαφίου 4.4 και θεωρούμε ότι οι απαιτούμενοι χρόνοι για μία προληπτική συντήρηση και μία επισκευή του μηχανισμού είναι συνεχείς τυχαίες μεταβλητές. Στο Εδάφιο 4.8 εφαρμόζουμε τον αλγόριθμο του Εδαφίου 4.5 για το μοντέλο του Εδαφίου 4.7 και παρουσιάζουμε δύο αριθμητικά παραδείγματα στα οποία οι απαιτούμενοι χρόνοι για μία προληπτική συντήρηση και μία επισκευή του μηχανισμού είναι τυχαίες μεταβλητές που ακολουθούν την κατανομή Γάμα και την κατανομή Weibull, αντίστοιχα.

Τα αποτελέσματα των Εδαφίων 4.2-4.6 έχουν δημοσιευτεί στην εργασία των Kyriakidis and Dimitrakos (2004). Τα αποτελέσματα των Εδαφίων 4.7 και 4.8 περιλαμβάνονται στην εργασία των Dimitrakos and Kyriakidis (2005b) η οποία έχει υποβληθεί προς δημοσίευση. Στο Παράρτημα της διατριβής παρουσιάζονται αποσπάσματα του προγράμματος Matlab για τον αλγόριθμο του Εδαφίου 4.5.

4.2 Περιγραφή του μοντέλου

Θεωρούμε ένα σύστημα παραγωγής το οποίο αποτελείται από ένα μηχανισμό τροφοδοσίας (I), έναν αποθηκευτικό χώρο (B) και μία μονάδα παραγωγής (P). Ο μηχανισμός τροφοδοτεί τη μονάδα παραγωγής με ένα ακατέργαστο υλικό. Ο αποθηκευτικός χώρος του ακατέργαστου υλικού έχει τοποθετηθεί ανάμεσα στο μηχανισμό και τη μονάδα παραγωγής με σκοπό να αποφεύγονται συχνές διακοπές της παραγωγικής διαδικασίας εξαιτίας ενδεχόμενων σοβαρών βλαβών του μηχανισμού τροφοδοσίας.

Η μονάδα παραγωγής δέχεται το ακατέργαστο υλικό με σταθερό ρυθμό ίσο με d μονάδες του υλικού ανά μονάδα χρόνου. Υποθέτουμε ότι η μέγιστη χωρητικότητα του αποθηκευτικού χώρου είναι ίση με K μονάδες του ακατέργαστου υλικού. Όσο ο αποθηκευτικός χώρος δεν είναι πλήρης, ο μηχανισμός τροφοδοτεί τη μονάδα παραγωγής με το ακατέργαστο υλικό με σταθερό ρυθμό ίσο με p μονάδες του υλικού ανά μονάδα χρόνου. Υποθέτουμε ότι $p > d$. Όταν ο αποθηκευτικός χώρος είναι πλήρης, ο μηχανισμός ελαττώνει την ταχύτητα τροφοδοσίας του από p μονάδες του υλικού ανά μονάδα χρόνου σε d μονάδες του υλικού ανά μονάδα χρόνου.



Σχήμα 4.1. *Το σύστημα παραγωγής*

Το σύστημα παραγωγής που περιγράψαμε απεικονίζεται στο παραπάνω σχήμα. Όπως αναφέρεται στην εργασία των Van der Duyn Schouten and Vanneste (1995) ως ένα παράδειγμα αυτού του συστήματος μπορεί να θεωρηθεί μία πλατφόρμα η οποία έχει τοποθετηθεί στη θάλασσα με σκοπό την άντληση αργού πετρελαίου. Το πετρέλαιο μεταφέρεται μέσω σωλήνων από την πλατφόρμα πρώτα σε δεξαμενές και μετέπειτα προς τα διυλιστήρια που βρίσκονται στην ακτή. Στην περίπτωση αυτή το αργό πετρέλαιο, η πλατφόρμα, τα διυλιστήρια και οι δεξαμενές είναι το ακατέργαστο υλικό, ο μηχανισμός, η μονάδα παραγωγής και ο αποθηκευτικός χώρος, αντίστοιχα.

Όπως αναφέρεται στην εργασία των Meller and Kim (1996) μία άλλη πιθανή εφαρμογή του συστήματος σχετίζεται με ένα εργοστάσιο το οποίο κατασκευάζει καθίσματα για τα αυτοκίνητα. Θεωρούμε ότι το εργοστάσιο διαθέτει ένα μηχανισμό ο οποίος παράγει καλύμματα καθισμάτων. Τα καλύμματα προωθούνται σε έναν αποθηκευτικό χώρο με μεγάλη χωρητικότητα. Ως μονάδα παραγωγής μπορεί να θεωρηθεί το τμήμα του εργοστασίου το οποίο δέχεται τα καλύμματα από το μηχανισμό μέσω του αποθηκευτικού χώρου και είναι υπεύθυνο για την τοποθέτησή τους στα καθίσματα.

Υποθέτουμε ότι ο μηχανισμός επιθεωρείται τις χρονικές στιγμές $\tau = 0, 1, \ldots$ οι οποίες θεωρούμε ότι ισαπέχουν μεταξύ τους και ταξινομείται σε μία από τις $m+2$ καταστάσεις $0, 1, \ldots, m+1,$ οι οποίες αναπαριστούν το βαθμό επιδείνωσής του. Ο βαθμός επιδείνωσης 0 σημαίνει ότι η λειτουργία του μηχανισμού είναι αρίστη. Ο βαθμός επιδείνωσης $m+1$ σημαίνει ότι ο μηχανισμός δεν λειτουργεί. Στους ενδιάμεσους βαθμούς επιδείνωσης $1, \ldots, m$ θεωρούμε ότι ο μηχανισμός λειτουργεί. Ο βαθμός επιδείνωσης 1 σημαίνει ότι η λειτουργία του μηχανισμού είναι σχεδόν αρίστη, ο βαθμός επιδείνωσης 2 σημαίνει ότι η λειτουργία του μηχανισμού είναι πάρα πολύ καλή, ..., ο βαθμός επιδείνωσης m σημαίνει ότι ο μηχανισμός έχει υποστεί σοβαρή βλάβη αλλά ωστόσο λειτουργεί.

Αν τη χρονική στιγμή επιθεώρησης τ ο βαθμός επιδείνωσης του μηχανισμού είναι ίσος με $i < m+1$ και το περιεχόμενο του αποθηκευτικού χώρου είναι ίσο με $x < K$ τότε την επόμενη χρονική στιγμή επιθεώρησης $\tau+1$ το περιεχόμενο του αποθηκευτικού χώρου θα είναι ίσο με $\min(x+p-d, K).$ Αυτή η αύξηση του περιεχομένου του αποθηκευτικού χώρου θα συμβεί ακόμη και όταν ο βαθμός επιδείνωσης του μηχανισμού τη χρονική στιγμή επιθεώρησης $\tau+1$ είναι ίσος με $m+1.$

Σε κάθε χρονική στιγμή επιθεώρησης η κατάσταση του μηχανισμού μπορεί να αναπαρασταθεί από το ζεύγος των μεταβλητών $(i, t),$ όπου $i \in \{0, 1, \ldots, m+1\}$ είναι ο βαθμός επιδείνωσης του μηχανισμού και $t \in \{0, 1, \ldots\}$ είναι η ηλικία του μηχανισμού. Θεωρούμε ότι, αν ο βαθμός επιδείνωσης του μηχανισμού είναι ίσος με i τη χρονική στιγμή $t,$ τη χρονική στιγμή $t+1$ γίνεται ίσος με j με πιθανότητα $p_{ij}(t).$ Υποθέτουμε ότι η πιθανότητα να γίνει τελικά ο βαθμός επιδείνωσης του μηχανισμού ίσος με $m+1$ από οποιοδήποτε αρχικό βαθμό επιδείνωσης είναι μη-μηδενική.

Αν σε κάποια χρονική στιγμή επιθεώρησης ο μηχανισμός βρίσκεται στην κατάσταση $(m+1, t),$ τότε υποχρεωτικά θέτουμε σε λειτουργία μία επισκευή του. Αν σε κάποια χρονική στιγμή επιθεώρησης ο μηχανισμός βρίσκεται σε μία κατάσταση $(i, t),$ $i \leq m,$ τότε δεν επεμβαίνουμε στη λειτουργία του ή θέτουμε σε λειτουργία μία προληπτική συντήρησή του. Θεωρούμε ότι μία προληπτική συντήρηση και μία επισκευή του μηχανισμού δεν είναι δυνατόν να διακοπούν και επαναφέρουν το μηχανισμό στην κατάσταση $(0, 0).$ Υποθέτουμε ότι, οι απαιτούμενοι χρόνοι για μία προληπτική συντήρηση και μία επισκευή του μηχανισμού ακολουθούν τη Γεωμετρική κατανομή με πιθανότητα επιτυχίας a και $b,$ αντίστοιχα. Συνεπώς η πιθανότητα μία προληπτική συντήρηση και μία επισκευή του μηχανισμού να διαρκέσουν $t \geq 1$ χρονικές μονάδες είναι ίση με $(1-a)^{t-1}a$ και $(1-b)^{t-1}b,$ αντίστοιχα.

Κατά τη διάρκεια μιας προληπτικής συντήρησης ή μιας επισκευής του μηχανισμού η τροφοδοσία του αποθηκευτικού χώρου με ακατέργαστο υλικό διακόπτεται. Αν κατά τη διάρκεια της προληπτικής συντήρησης ή της επισκευής ο αποθηκευτικός χώρος περιέχει κάποια ποσότητα του ακατέργαστου υλικού, η μονάδα παραγωγής λειτουργεί κανονικά καθώς δέχεται το υλικό του αποθηκευτικού χώρου με σταθερό ρυθμό ίσο με d μονάδες του υλικού ανά μονάδα χρόνου. Αν κατά τη διάρκεια της προληπτικής συντήρησης ή της επισκευής ο αποθηκευτικός χώρος αδειάσει, τότε η μονάδα παραγωγής διακόπτει τη λειτουργία της. Θεωρούμε ότι, αν κατά τη διάρκεια μιας προληπτικής συντήρησης ή μιας επισκευής του μηχανισμού ο αποθηκευτικός χώρος αδειάσει, η έλλειψη του ακατέργαστου υλικού στον αποθηκευτικό χώρο επιφέρει ένα κόστος για κάθε μονάδα χρόνου κατά την οποία ο αποθηκευτικός χώρος είναι κενός. Η μονάδα του κόστους επιλέγεται έτσι ώστε, για κάθε μονάδα χρόνου κατά την οποία ο μηχανισμός συντηρείται προληπτικά ή επισκευάζεται και ο αποθηκευτικός χώρος είναι κενός, το κόστος είναι ίσο με d μονάδες κόστους.

Αν σε κάποια χρονική στιγμή επιθεώρησης ο μηχανισμός βρίσκεται στην κατάσταση $(i, t), 0 \leq i \leq m, t \geq 0,$ και δεν επεμβαίνουμε στη λειτουργία του θεωρούμε ότι υπάρχει ένα κόστος λειτουργίας μέχρι την επόμενη χρονική στιγμή επιθεώρησης το οποίο είναι ίσο με $c_i(t),$ αν ο αποθηκευτικός χώρος δεν είναι πλήρης, και ίσο με $\widetilde{c}_i(t),$ αν ο αποθηκευτικός χώρος είναι πλήρης, αντίστοιχα. Για κάθε μονάδα χρόνου κατά την οποία ο μηχανισμός συντηρείται προληπτικά ή επισκευάζεται θεωρούμε ότι υπάρχει ένα κόστος ίσο με $c_p(t)$ ή με $c_f(t)$, αντίστοιχα, όπου t είναι η ηλικία του μηχανισμού. Υποθέτουμε επίσης ότι υπάρχει ένα κόστος ίσο με $h > 0$ για κάθε μονάδα χρόνου κατά την οποία μία μονάδα του ακατέργαστου υλικού είναι αποθηκευμένη στον αποθηκευτικό χώρο.

Έστω PM η κατάσταση στην οποία βρίσκεται ο μηχανισμός όταν θέτουμε σε λειτουργία μία προληπτική συντήρησή του. Ο χώρος καταστάσεων του συστήματος είναι το ακόλουθο άπειρο σύνολο:

$$S = \{0, ..., m+1, PM\} \times \{0, 1, ...\} \times \{0, ..., K\},$$

όπου, $(i, t, x) \in S$ αναπαριστά την κατάσταση του συστήματος στην οποία i είναι ο βαθμός επιδείνωσης του μηχανισμού, t είναι η ηλικία του μηχανισμού και x είναι το περιεχόμενο του αποθηκευτικού χώρου.

Μία πολιτική είναι ένας κανόνας με τον οποίον επιλέγονται οι ενέργειες κατά τις χρονικές στιγμές επιθεώρησης του μηχανισμού $\tau = 0, 1, ...$. Υπάρχουν τρεις ενέργειες: ενέργεια 0 (δεν επεμβαίνουμε στη λειτουργία του μηχανισμού), ενέργεια 1 (θέτουμε σε λειτουργία μία προληπτική συντήρηση του μηχανισμού), ενέργεια 2 (θέτουμε σε λειτουργία μία επισκευή του μηχανισμού). Στις καταστάσεις $(m+1, t, x),\ t \geq 0,\ 0 \leq x \leq K,$ η ενέργεια 2 είναι υποχρεωτική. Στις καταστάσεις $(PM, t, x),\ t \geq 0,\ 0 \leq x \leq K,$ η μόνη δυνατή ενέργεια είναι η ενέργεια 0. Στις καταστάσεις $(i, t, x),\ 0 \leq i \leq m,\ t \geq 0,\ 0 \leq x \leq K,$ οι ενέργειες 0 και 1 είναι δυνατές.

Θεωρούμε μία Μαρκοβιανή διαδικασία αποφάσεων σε διακριτό χρόνο με χώρο καταστάσεων S και αναζητούμε τη στάσιμη πολιτική η οποία, για οποιαδήποτε αρχική κατάσταση του συστήματος, ελαχιστοποιεί το μακροπρόθεσμο αναμενόμενο μέσο κόστος ανά μονάδα χρόνου.

Υποθέτουμε ότι ισχύουν οι ακόλουθες συνθήκες για τα κόστη, τις πιθανότητες μετάβασης, τους αναμενόμενους χρόνους της προληπτικής συντήρησης και τους αναμενόμενους χρόνους της επισκευής του μηχανισμού.

Συνθήκη 1: Για κάθε $t = 0, 1, ...,$ $c_0(t) \le c_1(t) \le \cdots \le c_m(t)$ και $\tilde{c}_0(t) \le \tilde{c}_1(t) \le \cdots \le \tilde{c}_m(t).$ Δηλαδή, για κάθε σταθερή ηλικία του μηχανισμού, το κόστος της λειτουργίας του αυξάνεται όσο αυξάνεται ο βαθμός επιδείνωσής του.

Συνθήκη 2: Για κάθε $t = 0, 1, ...,$ $\tilde{c}_i(t) \le c_i(t), 0 \le i \le m.$ Δηλαδή, η μείωση της ταχύτητας με την οποία ο μηχανισμός τροφοδοτεί τη μονάδα παραγωγής με ακατέργαστο υλικό από p μονάδες του υλικού ανά μονάδα χρόνου σε d μονάδες του υλικού ανά μονάδα χρόνου, όταν ο αποθηκευτικός χώρος είναι πλήρης, επιφέρει αντίστοιχη μείωση στο κόστος της λειτουργίας του μηχανισμού.

Συνθήκη 3: $0 < b < a \le 1.$ Δηλαδή, ο αναμενόμενος χρόνος για μία προληπτική συντήρηση είναι μικρότερος από τον αναμενόμενο χρόνο για μία επισκευή του μηχανισμού.

Συνθήκη 4: Για κάθε $t = 0, 1, ...,$ $c_p(t) \le c_f(t).$ Δηλαδή, ο ρυθμός κόστους της προληπτικής συντήρησης δεν υπερβαίνει το ρυθμό κόστους της επισκευής για οποιαδήποτε σταθερή ηλικία του μηχανισμού.

Συνθήκη 5: Θεωρούμε ότι $c_i(t), \tilde{c}_i(t), 0 \le i \le m, c_p(t), c_f(t)$ είναι αύξουσες συναρτήσεις ως προς t, δηλαδή η αύξηση της ηλικίας του μηχανισμού έχει ως επακόλουθο την αύξηση στο κόστος της λειτουργίας του, στο κόστος της προληπτικής συντήρησής του και στο κόστος της επισκευής του.

Συνθήκη 6: Θεωρούμε ότι $\sup_{t \ge 0} c_m(t) < \infty$ και $\sup_{t \ge 0} c_f(t) < \infty.$ Η συνθήκη αυτή εγγυάται ότι τα αναμενόμενα κόστη ενός βήματος του μοντέλου είναι φραγμένα.

Συνθήκη 7: Για κάθε $k = 0, 1, ..., m+1,$ η συνάρτηση

$$D_k(i,t) = \sum_{j=k}^{m+1} p_{ij}(t)$$

είναι αύξουσα ως προς $i, 0 \le i \le m$ και ως προς $t, t \ge 0.$ Η συνθήκη αυτή υποδηλώνει ότι $I_i(t) \le_{st} I_{i+1}(t), 0 \le i \le m,$ όπου $I_i(t)$ είναι μία τυχαία μεταβλητή η οποία αναπαριστά τον επόμενο βαθμό επιδείνωσης αν ο παρών βαθμός επιδείνωσης του μηχανισμού και η παρούσα ηλικία του είναι i και $t,$ αντίστοιχα. Το σύμβολο “$\le_{st}$” σημαίνει «στοχαστικά μικρότερο ή ίσο» (βλέπε σελ. 153 του βιβλίου του Ross (1983)).

Μπορεί να αποδειχθεί ότι η Συνθήκη 7 είναι ισοδύναμη με την ακόλουθη συνθήκη (βλέπε σελ. 122-123 του βιβλίου του Derman (1970)).

Συνθήκη 8: Για κάθε συνάρτηση $h(j,t),\ 0 \le j \le m+1,\ t \ge 0,$ η οποία είναι αύξουσα ως προς j και ως προς $t,$ η συνάρτηση $\sum_{j=0}^{m+1} p_{ij}(t)h(j,t)$ είναι επίσης αύξουσα ως προς i και ως προς $t.$

4.3 Η μορφή της βέλτιστης πολιτικής

Χάριν απλότητας και χωρίς βλάβη της γενικότητας στο παρόν εδάφιο και για το υπόλοιπο του κεφαλαίου υποθέτουμε ότι $p-d=1.$ Για κάθε αρχική κατάσταση (i,t,x) της διαδικασίας, το ελάχιστο αναμενόμενο αποπληθωρισμένο κόστος $V_\alpha(i,t,x,N),$ όταν απομένουν N βήματα μέχρι το τερματισμό της διαδικασίας μπορεί να υπολογιστεί αναδρομικά για κάθε $N=1, 2, ...,$ από τις παρακάτω εξισώσεις (βλέπε Θεώρημα 2.1 στη σελ. 18 της παρούσας διατριβής):

$$V_\alpha(i,t,x,N) = \min\left\{c_i(t) + hx + \alpha\sum_{j=0}^{m+1} p_{ij}(t)V_\alpha(j,t+1,x+1,N-1), V_\alpha(PM,t,x,N)\right\},$$

$$0 \le i \le m, \quad t \ge 0, \quad 0 \le x \le K-1,$$

$$V_\alpha(i,t,K,N) = \min\left\{\widetilde{c}_i(t) + hK + \alpha\sum_{j=0}^{m+1} p_{ij}(t)V_\alpha(j,t+1,K,N-1), V_\alpha(PM,t,K,N)\right\},$$

$$0 \le i \le m, \quad t \ge 0,$$

$$V_\alpha(m+1,t,x,N) = c_f(t) + hx + (d-x)^+ + \alpha b V_\alpha(0,0,(x-d)^+,N-1)$$
$$+ \alpha(1-b)V_\alpha(m+1,t+1,(x-d)^+,N-1),$$

$$t \ge 0, \quad 0 \le x \le K,$$

$$V_\alpha(PM,t,x,N) = c_p(t) + hx + (d-x)^+ + \alpha a V_\alpha(0,0,(x-d)^+,N-1)$$
$$+ \alpha(1-a)V_\alpha(PM,t+1,(x-d)^+,N-1),$$

$$t \ge 0, \quad 0 \le x \le K,$$

όπου, $V_\alpha(i,t,x,0) = 0,\ (i,t,x) \in S,$ και $\alpha \in (0,1)$ είναι ο αποπληθωριστικός παράγοντας.

Επισημαίνουμε ότι η έκφραση $(d-x)^+ = \max(d-x, 0)$ αναπαριστά τη ζήτηση σε ακατέργαστο υλικό η οποία χάνεται κατά τη διάρκεια μιας χρονικής περιόδου κατά την οποία ο μηχανισμός συντηρείται προληπτικά ή επισκευάζεται και ο αποθηκευτικός χώρος περιέχει x μονάδες του ακατέργαστου υλικού στην αρχή της περιόδου.

Η έκφραση $(x-d)^+ = \max(x-d, 0)$ αναπαριστά το περιεχόμενο του αποθηκευτικού χώρου στο τέλος μιας χρονικής περιόδου κατά την οποία ο μηχανισμός συντηρείται προληπτικά ή επισκευάζεται και ο αποθηκευτικός χώρος περιέχει x μονάδες του ακατέργαστου υλικού στην αρχή της περιόδου.

Ο πρώτος όρος ανάμεσα στα άγκιστρα στις παραπάνω εξισώσεις αντιστοιχεί στην ενέργεια 0 (δεν επεμβαίνουμε στη λειτουργία του μηχανισμού) και ο δεύτερος όρος αντιστοιχεί στην ενέργεια 1 (θέτουμε σε λειτουργία μία προληπτική συντήρηση του μηχανισμού).

Από τις παραπάνω εξισώσεις παίρνουμε το ακόλουθο αποτέλεσμα.

Λήμμα 4.1. Για κάθε $N = 0, 1, \ldots$ ισχύει ότι:

(i) $V_\alpha(i,t,x,N) \le V_\alpha(i+1,t,x,N),\ 0 \le i \le m,\ t \ge 0,\ 0 \le x \le K,$

(ii) $V_\alpha(PM,t,x,N) \le V_\alpha(m+1,t,x,N),\ t \ge 0,\ 0 \le x \le K,$

(iii) $V_\alpha(i,t,x,N) \le V_\alpha(i,t+1,x,N),\ 0 \le i \le m+1,\ t \ge 0,\ 0 \le x \le K,$

$V_\alpha(PM,t,x,N) \le V_\alpha(PM,t+1,x,N),\ t \ge 0,\ 0 \le x \le K.$

Απόδειξη. Η απόδειξη γίνεται με επαγωγή ως προς N. Το λήμμα ισχύει για $N = 0,$ διότι $V_\alpha(i,t,x,0) = 0$ για κάθε $(i,t,x) \in S.$ Υποθέτουμε ότι ισχύει για $N-1\,(\ge 0).$ Θα πρέπει να αποδείξουμε ότι ισχύει για $N.$ Αρχικά θα αποδείξουμε το (ii) στη συνέχεια το (i) και μετέπειτα το (iii).

(ii): Έστω $D = V_\alpha(m+1,t+1,(x-d)^+,N-1) - V_\alpha(0,0,(x-d)^+,N-1).$

Διαδοχικά έχουμε ότι:

$$\begin{aligned}
V_\alpha(PM,t,x,N) &= c_p(t) + hx + (d-x)^+ + \alpha a V_\alpha(0,0,(x-d)^+,N-1) \\
&\quad + \alpha(1-a)V_\alpha(PM,t+1,(x-d)^+,N-1) \\
&\le c_f(t) + hx + (d-x)^+ + \alpha a V_\alpha(0,0,(x-d)^+,N-1) \\
&\quad + \alpha(1-a)V_\alpha(m+1,t+1,(x-d)^+,N-1) \\
&= c_f(t) + hx + (d-x)^+ + \alpha V_\alpha(m+1,t+1,(x-d)^+,N-1) - \alpha a D
\end{aligned}$$

$$\leq c_f(t) + hx + (d-x)^+ + \alpha V_\alpha(m+1, t+1, (x-d)^+, N-1) - \alpha bD$$

$$= V_\alpha(m+1, t, x, N).$$

Η πρώτη ανισότητα προκύπτει από την επαγωγική υπόθεση για το (ii) και από τη Συνθήκη 4. Η δεύτερη ανισότητα προκύπτει από τη Συνθήκη 3 και την ανισότητα $D \geq 0$ η οποία είναι συνέπεια της επαγωγικής υπόθεσης για το (i) και το (iii).

(i): Θα πρέπει να δείξουμε ότι:

$$V_\alpha(i, t, x, N) \leq V_\alpha(i+1, t, x, N), \ \ 0 \leq i \leq m-1, \ \ t \geq 0, \ \ 0 \leq x \leq K, \tag{4.1}$$

και

$$V_\alpha(m, t, x, N) \leq V_\alpha(m+1, t, x, N), \ \ t \geq 0, \ \ 0 \leq x \leq K. \tag{4.2}$$

Η ανισότητα (4.2) μπορεί να επαληθευτεί εύκολα χρησιμοποιώντας το (ii), για $0 \leq x \leq K-1$:

$$V_\alpha(m, t, x, N) = \min\left\{ c_m(t) + hx + \alpha \sum_{j=0}^{m+1} p_{mj}(t) V_\alpha(j, t+1, x+1, N-1), V_\alpha(PM, t, x, N) \right\}$$

$$\leq V_\alpha(PM, t, x, N)$$

$$\leq V_\alpha(m+1, t, x, N).$$

Για $x = K,$ η αντίστοιχη ανισότητα προκύπτει με παρόμοιο τρόπο. Για $0 \leq x \leq K-1,$ $t \geq 0$ και $0 \leq i \leq m-1,$ έχουμε ότι:

$$V_\alpha(i,t,x,N) = \min\left\{c_i(t) + hx + \alpha\sum_{j=0}^{m+1} p_{ij}(t)V_\alpha(j,t+1,x+1,N-1), V_\alpha(PM,t,x,N)\right\}$$

$$\leq \min\left\{c_{i+1}(t) + hx + \alpha\sum_{j=0}^{m+1} p_{i+1,j}(t)V_\alpha(j,t+1,x+1,N-1), V_\alpha(PM,t,x,N)\right\}$$

$$= V_\alpha(i+1,t,x,N).$$

Η παραπάνω ανισότητα προκύπτει από τη Συνθήκη 1 και από την ανισότητα:

$$\sum_{j=0}^{m+1} p_{ij}(t)V_\alpha(j,t+1,x+1,N-1) \leq \sum_{j=0}^{m+1} p_{i+1,j}(t)V_\alpha(j,t+1,x+1,N-1),$$

η οποία είναι συνέπεια της επαγωγικής υπόθεσης των (i) και (iii) και της Συνθήκης 8. Έτσι, η ανισότητα (4.1) έχει αποδειχθεί για $0 \leq x \leq K-1.$ Με παρόμοιο τρόπο η ανισότητα (4.1) μπορεί να αποδειχθεί για $x = K.$

(iii): Από τη Συνθήκη 5 και την επαγωγική υπόθεση για το (iii) προκύπτει ότι οι εκφράσεις $V_\alpha(m+1,t,x,N)$ και $V_\alpha(PM,t,x,N)$ είναι αύξουσες ως προς $t \geq 0.$ Από την επαγωγική υπόθεση για το (i) και το (iii) και τη Συνθήκη 8 προκύπτει ότι η έκφραση:

$$\sum_{j=0}^{m+1} p_{ij}(t)V_\alpha(j,t+1,x+1,N-1)$$

είναι αύξουσα ως προς $t \geq 0.$ Από την μονοτονία της έκφρασης $V_\alpha(PM,t,x,N)$ ως προς t και τη Συνθήκη 5, προκύπτει ότι η έκφραση $V_\alpha(i,t,x,N), 0 \leq i \leq m, 0 \leq x \leq K-1,$ είναι αύξουσα ως προς $t.$ Με παρόμοιο τρόπο μπορεί να αποδειχθεί ότι η έκφραση $V_\alpha(i,t,K,N), 0 \leq i \leq m,$ είναι επίσης αύξουσα ως προς $t.$ ■

Για κάθε αρχική κατάσταση $(i,t,x)\in S$ της διαδικασίας, έστω $V_\alpha(i,t,x)$ το ελάχιστο συνολικό αναμενόμενο αποπληθωρισμένο κόστος σε άπειρο χρονικό ορίζοντα. Από το Θεώρημα 2.3 στη σελ. 20 της παρούσας διατριβής προκύπτει ότι: $\lim_{N\to\infty} V_\alpha(i,t,x,N)=V_\alpha(i,t,x).$ Από το (i) και το (iii) του Λήμματος 4.1 παίρνουμε το παρακάτω αποτέλεσμα.

Λήμμα 4.2. $V_\alpha(i,t,x)\le V_\alpha(i+1,t,x),\ \ 0\le i\le m,\ \ t\ge 0,\ \ 0\le x\le K,$

$V_\alpha(i,t,x)\le V_\alpha(i,t+1,x),\ \ 0\le i\le m+1,\ \ t\ge 0,\ \ 0\le x\le K.$

Έστω π^* η στάσιμη πολιτική η οποία ελαχιστοποιεί το συνολικό αναμενόμενο αποπληθωρισμένο κόστος για οποιαδήποτε αρχική κατάσταση της διαδικασίας. Η ύπαρξη της πολιτικής π^* είναι επακόλουθο της Συνθήκης 6. Το παρακάτω λήμμα εγγυάται την ύπαρξη μιας βέλτιστης στάσιμης πολιτικής η οποία για οποιαδήποτε αρχική κατάσταση της διαδικασίας ελαχιστοποιεί το μακροπρόθεσμο αναμενόμενο μέσο κόστος ανά μονάδα χρόνου (βλέπε Θεώρημα 2.4 και Θεώρημα 2.5 στις σελ. 21 και 22 της παρούσας διατριβής).

Λήμμα 4.3. Υπάρχει ένας θετικός πραγματικός αριθμός B τέτοιος ώστε

$$|V_\alpha(i,t,x)-V_\alpha(0,0,x^*)|\le B$$

για κάθε κατάσταση $(i,t,x)\in S$ και κάθε $\alpha\in(0,1),$ όπου $V_\alpha(0,0,x^*)=\max_{0\le x\le K} V_\alpha(0,0,x).$

Απόδειξη. Υποθέτουμε ότι η αρχική κατάσταση του συστήματος είναι η κατάσταση $(i,t,x)\in S.$ Έστω $X_r, a_r, C(X_r,a_r), r=0,1,...$ η κατάσταση του συστήματος τη χρονική στιγμή $r,$ η ενέργεια η οποία επιλέγεται τη χρονική στιγμή r και το υφιστάμενο κόστος τη χρονική στιγμή $r,$ αντίστοιχα. Έστω π η μη-στάσιμη πολιτική η οποία θέτει σε λειτουργία μία προληπτική συντήρηση ή μία επισκευή του μηχανισμού όταν το σύστημα βρίσκεται αρχικά στην κατάσταση

(i,t,x) και μετά την ολοκλήρωση της προληπτικής συντήρησης ή της επισκευής επιλέγει τις ίδιες ενέργειες με εκείνες της βέλτιστης πολιτικής π^*. Έστω επίσης $V_\alpha(i,t,x,\pi)$ το συνολικό αναμενόμενο αποπληθωρισμένο κόστος σε άπειρο χρονικό ορίζοντα, υπό τον έλεγχο της πολιτικής π, αν (i,t,x) είναι η αρχική κατάσταση του συστήματος. Έστω $T \in \{1, 2, ...\}$ ο απαιτούμενος χρόνος μέχρι να ολοκληρωθεί μία προληπτική συντήρηση ή μία επισκευή του μηχανισμού. Από τους ορισμούς των π και T έχουμε διαδοχικά ότι:

$$V_\alpha(i,t,x) \le V_\alpha(i,t,x,\pi) = E_\pi\left[\sum_{r=0}^{\infty}\alpha^r C(X_r,a_r) \mid X_0=(i,t,x)\right]$$

$$= E_\pi\left[\sum_{r=0}^{T-1}\alpha^r C(X_r,a_r) \mid X_0=(i,t,x)\right] + E_\pi\left[\sum_{r=T}^{\infty}\alpha^r C(X_r,a_r) \mid X_0=(i,t,x)\right]$$

$$= E_\pi\left[\sum_{r=0}^{T-1}\alpha^r C(X_r,a_r) \mid X_0=(i,t,x)\right] + \sum_{i=1}^{\infty} P(T=i)\alpha^i V_\alpha(0,0,(x-id)^+)$$

$$\le E_\pi\left[\sum_{r=0}^{T-1}\alpha^r C(X_r,a_r) \mid X_0=(i,t,x)\right] + V_\alpha(0,0,x^*)$$

$$\le \max\{d+hK+\sup_{t\ge 0} c_f(t), \sup_{t\ge 0} c_m(t)+hK\} E_\pi\left[\sum_{r=0}^{T-1}\alpha^r \mid X_0=(i,t,x)\right] + V_\alpha(0,0,x^*)$$

$$\le \max\{d+hK+\sup_{t\ge 0} c_f(t), \sup_{t\ge 0} c_m(t)+hK\} E(T) + V_\alpha(0,0,x^*)$$

$$\le B_1 + V_\alpha(0,0,x^*),$$

όπου, $B_1 = b^{-1}\max\{d+hK+\sup_{t\ge 0} c_f(t), \sup_{t\ge 0} c_m(t)+hK\}$, E αναπαριστά την αναμενόμενη τιμή και E_π αναπαριστά την υπό-συνθήκη αναμενόμενη τιμή, δοθέντος ότι η πολιτική π έχει υιοθετηθεί για τον έλεγχο της διαδικασίας. Η προτελευταία παραπάνω ανισότητα ισχύει διότι $\alpha \in (0,1)$. Από το (i) και το (iii) του Λήμματος 4.1, έχουμε ότι:

$$V_\alpha(i,t,x) - V_\alpha(0,0,x^*) \ge V_\alpha(0,0,\tilde{x}) - V_\alpha(0,0,x^*), \quad (i,t,x) \in S,$$

όπου, $V_\alpha(0,0,\widetilde{x}) = \min_{0\le x\le K} V_\alpha(0,0,x)$. Επομένως, ισχύει ότι:

$$|V_\alpha(i,t,x) - V_\alpha(0,0,x^*)| \le \max\{B_1, V_\alpha(0,0,x^*) - V_\alpha(0,0,\widetilde{x})\} = B, \quad (i,t,x)\in S. \quad \blacksquare$$

Το Λήμμα 4.3 υποδηλώνει ότι υπάρχουν αριθμοί $v(s), s\in S,$ και μία σταθερά g έτσι ώστε (βλέπε Θεώρημα 2.5 στη σελ. 22 της παρούσας διατριβής):

$$v(i,t,x) = \min\left\{c_i(t) + hx - g + \sum_{j=0}^{m+1} p_{ij}(t)v(j,t+1,x+1),\ v(PM,t,x)\right\},$$

$$0\le i\le m,\ \ t\ge 0,\ \ 0\le x\le K-1,$$

$$v(i,t,K) = \min\left\{\widetilde{c}_i(t) + hK - g + \sum_{j=0}^{m+1} p_{ij}(t)v(j,t+1,K),\ v(PM,t,K)\right\},$$

$$0\le i\le m,\ \ t\ge 0,$$

$$v(m+1,t,x) = c_f(t) + hx + (d-x)^+ - g + bv(0,0,(x-d)^+) + (1-b)v(m+1,t+1,(x-d)^+),$$

$$t\ge 0,\ \ 0\le x\le K,$$

$$v(PM,t,x) = c_p(t) + hx + (d-x)^+ - g + av(0,0,(x-d)^+) + (1-a)v(PM,t+1,(x-d)^+),$$

$$t\ge 0,\ \ 0\le x\le K.$$

Ο πρώτος όρος ανάμεσα στα άγκιστρα στις παραπάνω εξισώσεις αντιστοιχεί στην ενέργεια 0 (δεν επεμβαίνουμε στη λειτουργία του μηχανισμού) και ο δεύτερος όρος αντιστοιχεί στην ενέργεια 1 (θέτουμε σε λειτουργία μία προληπτική συντήρηση του μηχανισμού). Αν για μία κατάσταση (i, t, x) του συστήματος ο πρώτος όρος είναι μικρότερος από το δεύτερο όρο, τότε η βέλτιστη πολιτική επιλέγει στην κατάσταση αυτή την ενέργεια 0. Αν, αντιθέτως, ο πρώτος όρος είναι μεγαλύτερος από το δεύτερο όρο, τότε η βέλτιστη πολιτική επιλέγει την ενέργεια 1. Αν ο πρώτος και ο δεύτερος όρος είναι ίσοι, τότε και οι δύο ενέργειες θεωρούμε ότι είναι βέλτιστες. Σύμφωνα με το Θεώρημα 2.5, οι αριθμοί $v(s), s \in S,$ είναι τέτοιοι ώστε:

$$v(s) = \lim_{n\to\infty}[V_{\alpha_n}(s) - V_{\alpha_n}(0,0,x^*)],$$

για μία ακολουθία αριθμών $\{\alpha_n\},\ 0 < \alpha_n < 1,$ για την οποία ισχύει ότι: $\lim_{n\to\infty} \alpha_n = 1.$

Από το Λήμμα 4.2 και τις παραπάνω οριακές εκφράσεις προκύπτει το ακόλουθο αποτέλεσμα.

Πόρισμα 4.1. $v(i,t,x) \le v(i+1,t,x),\ 0 \le i \le m,\ t \ge 0,\ 0 \le x \le K,$

$v(i,t,x) \le v(i,t+1,x),\ 0 \le i \le m+1,\ t \ge 0,\ 0 \le x \le K.$

Η παρακάτω πρόταση παρέχει ένα χαρακτηρισμό της μορφής της βέλτιστης πολιτικής.

Πρόταση 4.1. Για κάθε σταθερό περιεχόμενο $x, 0 \le x \le K,$ του αποθηκευτικού χώρου και κάθε σταθερή ηλικία $t \ge 0$ του μηχανισμού, υπάρχει μία κρίσιμη τιμή $i^* (= i^*(t,x))$ του βαθμού επιδείνωσης του μηχανισμού τέτοια ώστε, η βέλτιστη πολιτική θέτει σε λειτουργία μία προληπτική συντήρηση αν και μόνο αν ο βαθμός επιδείνωσης i του μηχανισμού είναι μεγαλύτερος ή ίσος με την κρίσιμη τιμή i^*.

Απόδειξη. Υποθέτουμε ότι η βέλτιστη πολιτική επιλέγει την ενέργεια 1 στην κατάσταση $(i,t,x),$ $t \ge 0,\ 0 \le x \le K-1.$ Τότε ισχύει ότι:

$$v(PM,t,x) \le c_i(t) + hx - g + \sum_{j=0}^{m+1} p_{ij}(t) v(j,t+1,x+1). \tag{4.3}$$

Για να δείξουμε ότι η βέλτιστη πολιτική επιλέγει την ενέργεια 1 στην κατάσταση $(i+1,t,x)$ αρκεί να δείξουμε ότι:

$$v(PM,t,x) \le c_{i+1}(t) + hx - g + \sum_{j=0}^{m+1} p_{i+1,j}(t) v(j,t+1,x+1). \tag{4.4}$$

Από το Πόρισμα 4.1 και τις Συνθήκες 1 και 8 προκύπτει ότι το δεξιό μέλος της ανισότητας (4.4) είναι μεγαλύτερο ή ίσο με το δεξιό μέλος της ανισότητας (4.3). Επομένως η ανισότητα (4.3) συνεπάγεται την ανισότητα (4.4). Με παρόμοιο τρόπο αποδεικνύεται το ίδιο αποτέλεσμα για την περίπτωση κατά την οποία $x = K$. ■

Η Πρόταση 4.1 υποδηλώνει ότι η βέλτιστη πολιτική είναι μονότονη για σταθερό περιεχόμενο του αποθηκευτικού χώρου και σταθερή ηλικία του μηχανισμού.

4.4 Στάσιμη επιδείνωση του μηχανισμού

Θεωρούμε το μοντέλο που περιγράψαμε στο Εδάφιο 4.2 με τις παρακάτω υποθέσεις:

$$p_{ij}(t) = p_{ij}, \; 0 \le i, j \le m+1, \; t \ge 0,$$

$$c_i(t) = c_i, \; 0 \le i \le m, \; t \ge 0,$$

$$\tilde{c}_i(t) = \tilde{c}_i, \; 0 \le i \le m, \; t \ge 0,$$

$$c_p(t) = c_p, \; t \ge 0,$$

$c_f(t) = c_f, \; t \geq 0.$

Οι υποθέσεις αυτές υποδηλώνουν ότι σε κάθε χρονική στιγμή επιθεώρησης η κατάσταση του μηχανισμού προσδιορίζεται μόνο από το βαθμό επιδείνωσης $i \in \{0, 1, ..., m+1\}$ και δεν εξαρτάται από την ηλικία του. Η επιδείνωση του μηχανισμού θεωρούμε ότι είναι στάσιμη διότι οι πιθανότητες μετάβασης δεν εξαρτώνται από την ηλικία του. Ο χώρος καταστάσεων του συστήματος στην περίπτωση αυτή είναι το ακόλουθο πεπερασμένο σύνολο:

$$S = \{0, 1, ..., m+1, PM\} \times \{0, 1, ..., K\},$$

όπου, $(i, x) \in S$ αναπαριστά την κατάσταση του συστήματος στην οποία i είναι ο βαθμός επιδείνωσης του μηχανισμού και x είναι το περιεχόμενο του αποθηκευτικού χώρου.

Η διαδικασία επιδείνωσης του συστήματος είναι μία διαδικασία η οποία, υπό τον έλεγχο οποιασδήποτε πολιτικής, αναγεννάται οποτεδήποτε το σύστημα βρεθεί στην κατάσταση $(0, 0)$. Από γνωστά αποτελέσματα της θεωρίας των Μαρκοβιανών διαδικασιών αποφάσεων υπάρχει μία βέλτιστη στάσιμη πολιτική η οποία ελαχιστοποιεί το μακροπρόθεσμο αναμενόμενο μέσο κόστος ανά μονάδα χρόνου (βλέπε π.χ. Κεφάλαιο 3 του βιβλίου του Tijms (1994)).

Η Πρόταση 4.1 του προηγούμενου εδαφίου παίρνει τώρα την ακόλουθη μορφή.

Πρόταση 4.2. Για κάθε σταθερό περιεχόμενο $x, 0 \leq x \leq K,$ του αποθηκευτικού χώρου υπάρχει μία κρίσιμη τιμή $i^* (= i^*(x))$ τέτοια ώστε, η βέλτιστη πολιτική θέτει σε λειτουργία μία προληπτική συντήρηση αν και μόνο αν ο βαθμός επιδείνωσης i του μηχανισμού είναι μεγαλύτερος ή ίσος με την κρίσιμη τιμή i^*.

Σημείωση 4.1. Επισημαίνουμε ότι στην περίπτωση κατά την οποία ισχύει: $c_p = c_f = h = 0,$ $c_i = \tilde{c}_i = 0, \quad 0 \leq i \leq m, \quad p_{m,m+1} = 1, \quad p_{i,i+1} = r_i, \quad p_{i,m+1} = 1 - r_i, \quad 0 \leq i \leq m-1,$ το μοντέλο που περιγράψαμε στο παρόν εδάφιο συμπίπτει με αυτό που επινοήθηκε από τους Van der Duyn Schouten and Vanneste (1995). Η ακολουθία $\{r_i\}$, $0 \le i \le m$, είναι τέτοια ώστε $0 < r_i < 1$ και είναι αύξουσα ως προς i. Στην εργασία των Van der Duyn Schouten and Vanneste το μακροπρόθεσμο αναμενόμενο μέσο κόστος ανά μονάδα χρόνου υπό τον έλεγχο οποιασδήποτε πολιτικής αναπαριστά τη μέση ζήτηση σε ακατέργαστο υλικό η οποία χάνεται ανά μονάδα χρόνου. Η Πρόταση 4.2 γενικεύει ένα παρόμοιο αποτέλεσμα των Van der Duyn Schouten and Vanneste.

Σημείωση 4.2. Πολλά αριθμητικά παραδείγματα παρέχουν ισχυρή ένδειξη ότι, για κάθε σταθερό περιεχόμενο x του αποθηκευτικού χώρου, η κρίσιμη τιμή $i^*(x)$ η οποία, χαρακτηρίζει τη μονότονη μορφή της βέλτιστης πολιτικής, είναι φθίνουσα ως προς x. Μία αναλυτική απόδειξη αυτής της εικασίας φαίνεται δύσκολη.

Η Πρόταση 4.2 υποδηλώνει ότι στην περίπτωση της στάσιμης επιδείνωσης, όπως και στην περίπτωση της μη-στάσιμης επιδείνωσης του μηχανισμού, η βέλτιστη πολιτική είναι μονότονη για σταθερό περιεχόμενο του αποθηκευτικού χώρου.

4.5 Ένας αλγόριθμος όταν οι χρόνοι προληπτικής συντήρησης και επισκευής ακολουθούν τη Γεωμετρική κατανομή

Στην περίπτωση του μοντέλου που περιγράψαμε στο προηγούμενο εδάφιο, η βέλτιστη πολιτική μπορεί να υπολογιστεί με εφαρμογή του αλγορίθμου βελτίωσης των πολιτικών, του αλγορίθμου των διαδοχικών προσεγγίσεων και της μεθόδου του γραμμικού προγραμματισμού.

Ένας κατάλληλος αλγόριθμος βελτίωσης των πολιτικών μπορεί επίσης να σχεδιαστεί ο οποίος εφαρμόζεται στο σύνολο των μονότονων πολιτικών και παράγει μία ακολουθία αυστηρά βελτιωμένων πολιτικών που έχουν τη μονότονη μορφή. Ο αλγόριθμος παράγει πολιτικές οι οποίες για κάθε σταθερό περιεχόμενο x του αποθηκευτικού χώρου θέτουν σε λειτουργία μία προληπτική συντήρηση αν και μόνο αν ο βαθμός επιδείνωσης του μηχανισμού είναι μεγαλύτερος ή ίσος με μία κρίσιμη τιμή $i(x)$. Πολλά αριθμητικά παραδείγματα παρέχουν ισχυρή ένδειξη ότι η τελική πολιτική που παράγει ο αλγόριθμος είναι η βέλτιστη πολιτική διότι συμπίπτει με την τελική πολιτική που δημιουργεί ο αλγόριθμος βελτίωσης των πολιτικών και ο αλγόριθμος των διαδοχικών προσεγγίσεων.

Ο σχεδιασμός του αλγορίθμου βασίζεται στην τεχνική εμφύτευσης του Tijms (βλέπε σελ. 234 του βιβλίου του Tijms (1994)). Παρόμοιοι αλγόριθμοι έχουν σχεδιαστεί σε μοντέλα βέλτιστου ελέγχου ουρών αναμονής, βέλτιστου ελέγχου αποθεμάτων, βέλτιστης συντήρησης μηχανημάτων και βέλτιστου ελέγχου βιολογικών πληθυσμών (βλέπε π.χ. σελ. 234-248 του βιβλίου του Tijms (1994), Nobel and Tijms (1999, 2000), Kyriakidis (1993, 1999b, 2004)). Στη συνέχεια θα περιγράψουμε το σχεδιασμό του αλγορίθμου.

Έστω μία μονότονη πολιτική R η οποία χαρακτηρίζεται από τις κρίσιμες τιμές $i(x), 0 \le x \le K.$ Έστω επίσης $w(s), s \in S,$ οι σχετικές τιμές της πολιτικής R (βλέπε σελ. 25 της παρούσας διατριβής). Θεωρούμε ότι:

$$w(0, 0) = 0. \tag{4.5}$$

Αν η μονότονη πολιτική R υιοθετηθεί για τον έλεγχο του συστήματος, τότε το σύστημα είναι δυνατόν να μεταβεί σε κάποια κατάσταση του συνόλου των καταστάσεων:

$$E = \bigcup_{x=0}^{K} \{(i, x), 0 \le i \le i(x)\},$$

από οποιαδήποτε αρχική κατάσταση $s \in S.$ Σύμφωνα με τη σχέση (3.6.1) στη σελ. 235 του βιβλίου του Tijms (1994), οι σχετικές τιμές $w(s), s \in S,$ ικανοποιούν το ακόλουθο σύστημα των γραμμικών εξισώσεων:

$$w(s) = C_s^E - g(R)T_s^E + \sum_{\ell \in E} p_{s\ell}^E w(\ell), \quad s \in S, \tag{4.6}$$

όπου, $g(R)$ είναι το μακροπρόθεσμο αναμενόμενο μέσο κόστος ανά μονάδα χρόνου υπό τον έλεγχο της πολιτικής $R,$ $p_{s\ell}^E$ είναι η πιθανότητα η πρώτη κατάσταση κατά την είσοδο του συστήματος στο σύνολο E να είναι η κατάσταση $\ell,$ δοθέντος ότι η πολιτική R έχει υιοθετηθεί για τον έλεγχο του συστήματος και η αρχική κατάσταση είναι η κατάσταση $s \in S,$ και, T_s^E, C_s^E είναι ο αναμενόμενος χρόνος και το αναμενόμενο κόστος, αντίστοιχα. Υποθέτουμε ότι, αν η αρχική κατάσταση του συστήματος είναι μία κατάσταση $s \in E,$ τότε η πρώτη κατάσταση κατά την είσοδο του συστήματος στο σύνολο E είναι η κατάσταση στην οποία μεταβαίνει το σύστημα στην επόμενη επιστροφή του στο σύνολο E.

Οι ποσότητες T_s^E, C_s^E και $p_{s\ell}^E,$ $s \in S,$ $\ell \in E,$ μπορούν να υπολογιστούν εύκολα αν λάβουμε υπόψη τη δομή του μοντέλου υπό τον έλεγχο της μονότονης πολιτικής R και χρησιμοποιήσουμε διάφορες μορφές του Θεωρήματος Ολικής Πιθανότητας. Για παράδειγμα,

$$T_{(i,x)}^E = 1 + a^{-1} \sum_{j=i(x+1)+1}^{m} p_{ij} + b^{-1} p_{i,m+1}, \ 0 \le i < i(x), \ 0 \le x \le K-1,$$

$$C_{(i,x)}^E = a^{-1}\left\{c_p + (1-a)^{[x/d]}\left(d - ax + ad\left[\frac{x}{d}\right]\right)\right\} + \frac{h}{2}(1-a)^{[x/d]}\left(1+\left[\frac{x}{d}\right]\right)\left(2x - d\left[\frac{x}{d}\right]\right)$$
$$+ \frac{ha}{2}\sum_{t=1}^{[x/d]}(1-a)^{t-1}(2tx - dt^2 + dt), \quad i(x) \le i \le m, \quad 0 \le x \le K, \qquad (4.7)$$

$$p_{(i,K)(j,K)}^E = p_{ij}, \quad 0 \le i < i(K), \quad 0 \le j \le i(K),$$

$$p_{(i,x)(0,0)}^E = (1-a)^{[(x+1)/d]} \sum_{j=i(x+1)+1}^{m} p_{ij} + p_{i,m+1}(1-b)^{[(x+1)/d]}, \quad 0 \le i < i(x), \quad 0 \le x \le K-1, \qquad (4.8)$$

όπου $[x/d]$ και $[(x+1)/d]$ αναπαριστούν τα ακέραια μέρη των x/d και $(x+1)/d,$ αντίστοιχα.

Θα εξηγήσουμε αναλυτικά τον τρόπο με τον οποίον εξάγονται οι σχέσεις (4.7) και (4.8). Για τη σχέση (4.7) η αρχική κατάσταση του συστήματος είναι μία κατάσταση $(i,x) \in S$ τέτοια ώστε $i(x) \le i \le m$ και $0 \le x \le K,$ στην οποία, υπό τον έλεγχο της πολιτικής $R,$ επιλέγουμε να θέσουμε σε λειτουργία μία προληπτική συντήρηση του μηχανισμού.

Υπενθυμίζουμε ότι, κατά τη διάρκεια μιας προληπτικής συντήρησης, στη λειτουργία του συστήματος υπεισέρχονται τα εξής κόστη.

Ένα κόστος προληπτικής συντήρησης ίσο με c_p για κάθε μονάδα χρόνου κατά την οποία ο μηχανισμός συντηρείται προληπτικά.

Ένα κόστος έλλειψης του υλικού στον αποθηκευτικό χώρο ίσο με d μονάδες για κάθε μονάδα χρόνου κατά την οποία ο μηχανισμός συντηρείται προληπτικά και ο αποθηκευτικός χώρος είναι κενός.

Ένα κόστος αποθήκευσης του υλικού στον αποθηκευτικό χώρο ίσο με h για κάθε μονάδα χρόνου κατά την οποία μία μονάδα του υλικού είναι αποθηκευμένη στον αποθηκευτικό χώρο.

Έστω $C^{M}_{(i,x)}$ το αναμενόμενο κόστος της προληπτικής συντήρησης, $C^{SH}_{(i,x)}$ το αναμενόμενο κόστος της έλλειψης του υλικού στον αποθηκευτικό χώρο και $C^{H}_{(i,x)}$ το αναμενόμενο κόστος της αποθήκευσης του υλικού στον αποθηκευτικό χώρο, αν στην κατάσταση $(i,x)\in S$ επιλέγουμε να θέσουμε σε λειτουργία μία προληπτική συντήρηση του μηχανισμού. Ισχύει ότι:

$$C^{E}_{(i,x)} = C^{M}_{(i,x)} + C^{SH}_{(i,x)} + C^{H}_{(i,x)}, \quad i(x)\le i\le m, \quad 0\le x\le K. \tag{4.9}$$

Για να υπολογίσουμε τους τρεις όρους του αθροίσματος στο δεξιό μέλος της εξίσωσης (4.9) δεσμευόμαστε ως προς τη χρονική διάρκεια μιας προληπτικής συντήρησης. Το αναμενόμενο κόστος της προληπτικής συντήρησης δίνεται από τη σχέση:

$$C^{M}_{(i,x)} = c_p a^{-1}, \quad i(x)\le i\le m, \quad 0\le x\le K. \tag{4.10}$$

Αν η διάρκεια μιας προληπτικής συντήρησης είναι ίση με $t < x/d$ χρονικές μονάδες, το αναμενόμενο κόστος της έλλειψης του υλικού στον αποθηκευτικό χώρο είναι μηδενικό διότι μετά από την ολοκλήρωση της προληπτικής συντήρησης ο αποθηκευτικός χώρος περιέχει $x-td>0$ μονάδες του ακατέργαστου υλικού. Αν, αντιθέτως, η διάρκεια μιας προληπτικής συντήρησης είναι ίση με $t\ge x/d$ χρονικές μονάδες, το αναμενόμενο κόστος της έλλειψης του υλικού στον αποθηκευτικό χώρο δίνεται από το ακόλουθο άπειρο άθροισμα:

$$C^{SH}_{(i,x)} = \sum_{t=[x/d]+1}^{\infty}(td-x)(1-a)^{t-1}a, \quad i(x)\le i\le m, \quad 0\le x\le K,$$

διότι, για κάθε μονάδα χρόνου κατά την οποία ο μηχανισμός συντηρείται προληπτικά και ο αποθηκευτικός χώρος είναι κενός, το κόστος της έλλειψης του υλικού είναι ίσο με d. Ο υπολογισμός του παραπάνω αθροίσματος οδηγεί στη σχέση:

$$C^{SH}_{(i,x)} = a^{-1}(1-a)^{[x/d]}\left(d - ax + ad\left[\frac{x}{d}\right]\right), \quad i(x) \le i \le m, \quad 0 \le x \le K, \tag{4.11}$$

όπου, $[x/d]$ αναπαριστά το ακέραιο μέρος του x/d. Έστω $C^{H1}_{(i,x)}$ και $C^{H2}_{(i,x)}$ τα αναμενόμενα κόστη της αποθήκευσης του υλικού στον αποθηκευτικό χώρο αν στην κατάσταση $(i,x) \in S$ θέσουμε σε λειτουργία μία προληπτική συντήρηση του μηχανισμού η οποία διαρκεί $t < x/d$ και $t \ge x/d$ χρονικές μονάδες, αντίστοιχα. Ισχύει ότι:

$$C^{H}_{(i,x)} = C^{H1}_{(i,x)} + C^{H2}_{(i,x)}, \quad i(x) \le i \le m, \quad 0 \le x \le K. \tag{4.12}$$

Αν η διάρκεια μιας προληπτικής συντήρησης είναι ίση με $t < x/d$ χρονικές μονάδες, το αναμενόμενο κόστος της αποθήκευσης του υλικού είναι:

$$C^{H1}_{(i,x)} = \sum_{t=1}^{[x/d]}\left[\sum_{j=0}^{t-1} h(x-jd)\right](1-a)^{t-1}a, \quad i(x) \le i \le m, \quad 0 \le x \le K. \tag{4.13}$$

Αν η διάρκεια μιας προληπτικής συντήρησης είναι ίση με $t \ge x/d$ χρονικές μονάδες, το αναμενόμενο κόστος της αποθήκευσης του υλικού είναι:

$$C^{H2}_{(i,x)} = \sum_{t=[x/d]+1}^{\infty}\left[\sum_{j=0}^{[x/d]} h(x-jd)\right](1-a)^{t-1}a, \quad i(x) \le i \le m, \quad 0 \le x \le K. \tag{4.14}$$

Ο γενικός όρος $h(x-jd)$ του αθροίσματος ανάμεσα στις αγκύλες στις εξισώσεις (4.13) και (4.14) αναπαριστά το κόστος της αποθήκευσης του υλικού αν η προληπτική συντήρηση ολοκληρωθεί μετά από $j \ge 0$ χρονικές μονάδες. Ο γενικός όρος μπορεί να δικαιολογηθεί ως εξής. Για κάθε μονάδα χρόνου κατά την οποία ο μηχανισμός συντηρείται προληπτικά, ο αποθηκευτικός χώρος ελαττώνει το περιεχόμενο του κατά d μονάδες. Επιπλέον υπάρχει ένα κόστος ίσο με h για κάθε μονάδα χρόνου κατά την οποία μία μονάδα του υλικού είναι αποθηκευμένη στον αποθηκευτικό χώρο. Το άνω όριο $[x/d]$ του αθροίσματος ανάμεσα στις αγκύλες στην εξίσωση (4.14) δικαιολογείται από το γεγονός ότι ο αποθηκευτικός χώρος παύει να περιέχει κάποια ποσότητα του ακατέργαστου υλικού και κατά συνέπεια το κόστος της αποθήκευσης του υλικού μηδενίζεται αν η διάρκεια μιας προληπτικής συντήρησης είναι ίση με $t \geq x/d$ χρονικές μονάδες. Από τον υπολογισμό του αθροίσματος ανάμεσα στις αγκύλες, το διπλό άθροισμα της εξίσωσης (4.13) μπορεί να γραφεί με την ακόλουθη μορφή:

$$C_{(i,x)}^{H1} = \frac{ha}{2}\sum_{t=1}^{[x/d]}(1-a)^{t-1}(2tx - dt^2 + dt), \quad i(x) \leq i \leq m, \quad 0 \leq x \leq K. \tag{4.15}$$

Από τον υπολογισμό του διπλού αθροίσματος της εξίσωσης (4.14), έχουμε ότι:

$$C_{(i,x)}^{H2} = \frac{h}{2}(1-a)^{[x/d]}\left(1+\left[\frac{x}{d}\right]\right)\left(2x - d\left[\frac{x}{d}\right]\right), \quad i(x) \leq i \leq m, \quad 0 \leq x \leq K. \tag{4.16}$$

Από τις εξισώσεις (4.15) και (4.16) η εξίσωση (4.12) μπορεί να γραφεί ως εξής:

$$C_{(i,x)}^{H} = \frac{ha}{2}\sum_{t=1}^{[x/d]}(1-a)^{t-1}(2tx - dt^2 + dt) + \frac{h}{2}(1-a)^{[x/d]}\left(1+\left[\frac{x}{d}\right]\right)\left(2x - d\left[\frac{x}{d}\right]\right), \tag{4.17}$$

$$i(x) \leq i \leq m, \quad 0 \leq x \leq K.$$

Από τις εξισώσεις (4.10), (4.11), (4.17) και τη σχέση (4.9) προκύπτει η έκφραση (4.7).

Για τη σχέση (4.8) η αρχική κατάσταση του συστήματος είναι μία κατάσταση $(i,x) \in E$ τέτοια ώστε $0 \leq i < i(x)$ και $0 \leq x \leq K-1$, στην οποία, υπό τον έλεγχο της πολιτικής R, δεν επεμβαίνουμε στη λειτουργία του μηχανισμού.

Η κατάσταση $(0,0)$ είναι η πρώτη κατάσταση στην επόμενη επιστροφή του συστήματος στο σύνολο E, αν το σύστημα μεταβεί με πιθανότητα p_{ij} σε μία κατάσταση $(j, x+1)$ τέτοια ώστε

$i(x+1) < j \le m$ και $0 \le x \le K-1,$ στην οποία, υπό τον έλεγχο της πολιτικής $R,$ θέτουμε σε λειτουργία μία προληπτική συντήρηση ή αν το σύστημα μεταβεί με πιθανότητα $p_{i,m+1}$ σε μία κατάσταση $(m+1, x+1)$ τέτοια ώστε $0 \le x \le K-1,$ στην οποία, υπό τον έλεγχο της πολιτικής $R,$ θέτουμε σε λειτουργία μία επισκευή του μηχανισμού. Μετά από την ολοκλήρωση μιας προληπτικής συντήρησης ή μιας επισκευής, η λειτουργία του μηχανισμού είναι αρίστη και ο βαθμός επιδείνωσής του ισούται με 0. Αν η διάρκεια μιας προληπτικής συντήρησης ή μιας επισκευής είναι ίση με $t \ge (x+1)/d$ χρονικές μονάδες, ο αποθηκευτικός χώρος είναι κενός. Το σύστημα μεταβαίνει στην κατάσταση $(0, 0)$ με πιθανότητα:

$$p^E_{(i,x)(0,0)} = \left(\sum_{j=i(x+1)+1}^{m} p_{ij} \right) \left(\sum_{t=[(x+1)/d]+1}^{\infty} (1-a)^{t-1} a \right) + p_{i,m+1} \left(\sum_{t=[(x+1)/d]+1}^{\infty} (1-b)^{t-1} b \right), \tag{4.18}$$

$$0 \le i < i(x), \quad 0 \le x \le K-1,$$

όπου, $[(x+1)/d]$ αναπαριστά το ακέραιο μέρος του $(x+1)/d.$ Ισχύει ότι:

$$\sum_{t=[(x+1)/d]+1}^{\infty} (1-a)^{t-1} a = (1-a)^{[(x+1)/d]} \quad \text{και} \quad \sum_{t=[(x+1)/d]+1}^{\infty} (1-b)^{t-1} b = (1-b)^{[(x+1)/d]} \tag{4.19}$$

Η έκφραση (4.8) προκύπτει από την εξίσωση (4.18) και τις σχέσεις (4.19).

Οι σχετικές τιμές $w(s),\ s \in E,$ και το μακροπρόθεσμο αναμενόμενο μέσο κόστος $g(R)$ ανά μονάδα χρόνου, υπό τον έλεγχο της πολιτικής $R,$ μπορούν να υπολογιστούν επιλύνοντας το ακόλουθο σύστημα των γραμμικών εξισώσεων:

$$w(s) = C^E_s - g(R) T^E_s + \sum_{\ell \in E} p^E_{s\ell} w(\ell), \quad s \in E, \tag{4.20}$$

και την εξίσωση (4.5).

Θεωρούμε την ποσότητα $Q_R(s; a)$ η οποία ορίζεται ως εξής:

$$Q_R(s;a) = C(s,a) - g(R)T(s,a) + \sum_{\ell \in S} p_{s\ell}(a)w(\ell), \quad s \in \{0, ..., m\} \times \{0, ..., K\}, \quad a \in \{0, 1\}, \quad (4.21)$$

όπου, $p_{s\ell}(a)$ είναι η πιθανότητα η επόμενη κατάσταση της διαδικασίας να είναι η κατάσταση $\ell,$ αν η παρούσα κατάσταση είναι η κατάσταση s και επιλέγεται η ενέργεια $a,$ και $T(s,a),$ $C(s,a)$ είναι ο αναμενόμενος χρόνος ενός βήματος και το αναμενόμενο κόστος ενός βήματος, αντίστοιχα, για τη μετάβαση της διαδικασίας στην επόμενη κατάσταση, αν η παρούσα κατάσταση είναι η κατάσταση s και επιλέγεται η ενέργεια $a.$

Υποθέτουμε ότι για κάποιο περιεχόμενο $x,$ $0 \le x \le K,$ του αποθηκευτικού χώρου, υπάρχει ένας ακέραιος αριθμός $\tilde{i}(x)$ τέτοιος ώστε $0 \le \tilde{i}(x) < i(x)$ και $Q_R((i,x);1) < w(i,x),$ $\tilde{i}(x) \le i < i(x).$ Τότε, σύμφωνα με το Θεώρημα 2.7 στη σελ. 24 της παρούσας διατριβής, το μακροπρόθεσμο αναμενόμενο μέσο κόστος υπό τον έλεγχο της μονότονης πολιτικής η οποία χαρακτηρίζεται από τις κρίσιμες τιμές $i(0), ..., i(x-1), \tilde{i}(x), i(x+1), ..., i(K)$ είναι μικρότερο από $g(R).$

Παρόμοια, αν για κάποιο περιεχόμενο $x,$ $0 \le x \le K,$ του αποθηκευτικού χώρου, υπάρχει ένας ακέραιος αριθμός $\tilde{i}(x)$ τέτοιος ώστε $i(x) < \tilde{i}(x) \le m+1$ και $Q_R((i,x);0) < w(i,x),$ $i(x) \le i < \tilde{i}(x),$ τότε, σύμφωνα πάλι με το Θεώρημα 2.7, το μακροπρόθεσμο αναμενόμενο μέσο κόστος υπό τον έλεγχο της μονότονης πολιτικής η οποία χαρακτηρίζεται από τις κρίσιμες τιμές $i(0), ..., i(x-1), \tilde{i}(x), i(x+1), ..., i(K)$ είναι μικρότερο από $g(R).$

Οι παραπάνω παρατηρήσεις μας οδηγούν στο σχεδιασμό του ακόλουθου αλγορίθμου ο οποίος παράγει μία ακολουθία αυστηρά βελτιωμένων μονότονων πολιτικών.

Αλγόριθμος

Βήμα 1. *(Έναρξη).* Επιλέγουμε μία αρχική μονότονη πολιτική R η οποία χαρακτηρίζεται από τις κρίσιμες τιμές $i(x),$ $0 \le x \le K.$

Βήμα 2. *(Εύρεση σχετικών τιμών και μέσου κόστους).* Για τη μονότονη πολιτική $R,$ υπολογίζουμε το μακροπρόθεσμο αναμενόμενο μέσο κόστος $g(R)$ ανά μονάδα χρόνου και τις σχετικές τιμές $w(s),$ $s \in E,$ επιλύνοντας το σύστημα των γραμμικών εξισώσεων (4.5) και (4.20).

Βήμα 3. *(Βελτίωση των πολιτικών).* Για κάθε σταθερό περιεχόμενο x, $0 \le x \le K$, του αποθηκευτικού χώρου:

(α) Βρίσκουμε, αν υπάρχει, το μικρότερο ακέραιο αριθμό $\tilde{i}(x)$ τέτοιο ώστε $0 \le \tilde{i}(x) < i(x)$ και $Q_R((i,x);1) < w(i,x),\ \ \tilde{i}(x) \le i < i(x),$

διαφορετικά,

(β) βρίσκουμε, αν υπάρχει, το μεγαλύτερο ακέραιο αριθμό $\tilde{i}(x)$ τέτοιο ώστε $i(x) < \tilde{i}(x) \le m+1$ και $Q_R((i,x);0) < w(i,x),\ \ i(x) \le i < \tilde{i}(x).$

Οι ποσότητες $Q_R((i,x);1)$ και $Q_R((i,x);0)$ δίνονται από την εξίσωση (4.21). Οι σχετικές τιμές $w(s),\ \ s \in S-E,$ υπολογίζονται από την εξίσωση (4.6).

Αντικαθιστούμε την κρίσιμη τιμή $i(x)$ με την κρίσιμη τιμή $\tilde{i}(x)$ για εκείνες τις τιμές του x, $0 \le x \le K,$ για τις οποίες είναι δυνατόν να βρούμε μία κρίσιμη τιμή $\tilde{i}(x)$ και επιστρέφουμε στο Βήμα 2.

Βήμα 4. *(Κριτήριο Σύγκλισης).* Αν δεν είναι δυνατόν να βρεθεί καμία κρίσιμη τιμή $\tilde{i}(x),\ 0 \le x \le K,$ ο αλγόριθμος σταματά. Η τελική πολιτική που παράγει ο αλγόριθμος είναι η πολιτική R με μακροπρόθεσμο αναμενόμενο μέσο κόστος $g(R)$ ανά μονάδα χρόνου.

Ο αλγόριθμος παράγει μία ακολουθία αυστηρά βελτιωμένων μονότονων πολιτικών και σταματά μετά από ένα πεπερασμένο αριθμό επαναλήψεων διότι το σύνολο των μονότονων πολιτικών είναι πεπερασμένο. Πολλά αριθμητικά παραδείγματα παρέχουν ισχυρή ένδειξη ότι η τελική πολιτική που δημιουργεί ο αλγόριθμος είναι η βέλτιστη πολιτική διότι συμπίπτει με την τελική πολιτική που δίνει ο αλγόριθμος βελτίωσης των πολιτικών και ο αλγόριθμος των διαδοχικών προσεγγίσεων. Μία αναλυτική απόδειξη αυτής της εικασίας φαίνεται δύσκολη. Ο απαιτούμενος υπολογιστικός χρόνος για την εκτέλεση του αλγορίθμου είναι κατά πολύ μικρότερος από τον απαιτούμενο υπολογιστικό χρόνο για την εκτέλεση του αλγορίθμου βελτίωσης των πολιτικών. Αυτό οφείλεται στο γεγονός ότι ο αριθμός των αγνώστων στο σύστημα των γραμμικών εξισώσεων του Βήματος 2 του αλγορίθμου είναι ίσος με τον αριθμό των στοιχείων του συνόλου $E,$ ενώ ο αριθμός των αγνώστων στο σύστημα των γραμμικών εξισώσεων του αντίστοιχου βήματος της εύρεσης των σχετικών τιμών και του μέσου κόστους στον αλγόριθμο βελτίωσης των πολιτικών είναι ίσος με τον αριθμό των στοιχείων του χώρου καταστάσεων $S.$ Από πολλά αριθμητικά παραδείγματα που έχουμε εξετάσει συμπεραίνουμε επίσης ότι, ο αριθμός των επαναλήψεων του αλγορίθμου δεν επηρεάζεται ιδιαίτερα από την αρχική μονότονη πολιτική που επιλέγουμε κατά την έναρξη του αλγορίθμου στο Βήμα 1.

Ένα αριθμητικό παράδειγμα παρουσιάζεται παρακάτω. Θεωρούμε την περίπτωση κατά την οποία $m = 50,$ $K = 10,$ $a = 0.9,$ $b = 0.2,$ $p = 9,$ $d = 8,$ $c_p = 0.4,$ $c_f = 0.8,$ $h = 0.5,$ $c_i = 0.1(i+1),$ $\tilde{c}_i = 0.05(i+1),$ $0 \le i \le m.$ Υποθέτουμε ότι οι μη-μηδενικές πιθανότητες μετάβασης p_{ij}, $0 \le i, j \le m+1,$ είναι τέτοιες ώστε:

$$p_{ij} = \frac{1}{m+2-i}, \quad i \le j \le m+1.$$

Αυτό σημαίνει ότι αν ο παρών βαθμός επιδείνωσης του μηχανισμού είναι ίσος με $i,$ τότε ο επόμενος βαθμός επιδείνωσης κατανέμεται ομοιόμορφα στο σύνολο $\{i, i+1, ..., m+1\}$. Οι πιθανότητες μετάβασης ικανοποιούν τη Συνθήκη 7 διότι, για κάθε $k = 0, ..., m+1,$ η ποσότητα:

$$\sum_{j=k}^{m+1} p_{ij} = \frac{m+2-k}{m+2-i}$$

είναι αύξουσα ως προς $i,$ $0 \le i \le 50.$ Ως αρχική μονότονη πολιτική στο Βήμα 1 του αλγορίθμου επιλέγουμε την πολιτική η οποία χαρακτηρίζεται από τις κρίσιμες τιμές $i(x) = 50,$ $0 \le x \le 10.$ Στον Πίνακα 4.1 παρουσιάζονται οι διαδοχικές μονότονες πολιτικές που παράγει ο αλγόριθμος και τα αντίστοιχα μακροπρόθεσμα αναμενόμενα μέσα κόστη τους ανά μονάδα χρόνου.

Πίνακας 4.1. Οι διαδοχικές μονότονες πολιτικές

Οι κρίσιμες τιμές $i(x),\ \ 0 \le x \le 10,$ των πολιτικών	Μέσο κόστος
$i(0) = \cdots = i(10) = 50$	6.416
$i(0) = 13,\ \ i(1) = \cdots = i(10) = 0$	4.392
$i(0) = 37,\ \ i(1) = 34,\ \ i(2) = 30,\ \ i(3) = 27,\ \ i(4) = 23,$ $i(5) = 18,\ \ i(6) = 14,\ \ i(7) = 9,\ \ i(8) = i(9) = i(10) = 0$	3.872
$i(0) = 33,\ \ i(1) = 29,\ \ i(2) = 26,\ \ i(3) = 22,\ \ i(4) = 17,\ \ i(5) = 13,$ $i(6) = 9,\ \ i(7) = 4,\ \ i(8) = i(9) = i(10) = 0$	3.855

Ο απαιτούμενος υπολογιστικός χρόνος για την εκτέλεση του αλγορίθμου ήταν 8.1 δευτερόλεπτα ενώ ο απαιτούμενος υπολογιστικός χρόνος για την εκτέλεση του αλγορίθμου βελτίωσης των πολιτικών ήταν 26.4 δευτερόλεπτα. Στο παρόν αριθμητικό παράδειγμα παρατηρούμε ότι, όταν το περιεχόμενο του αποθηκευτικού χώρου είναι αρκετά υψηλό (ανάμεσα στις τιμές 8 και 10), η βέλτιστη πολιτική θέτει σε λειτουργία μία προληπτική συντήρηση ακόμα και όταν η λειτουργία του μηχανισμού είναι αρίστη, δηλαδή ο βαθμός επιδείνωσής του ισούται με 0. Αυτό φαίνεται διαισθητικά λογικό διότι οι πιθανότητες μετάβασης σε καταστάσεις με υψηλά κόστη λειτουργίας δεν είναι αμελητέες.

4.6 Δύο γενικεύσεις του μοντέλου

Στο παρόν εδάφιο θεωρούμε δύο γενικεύσεις του μοντέλου που περιγράψαμε στο Εδάφιο 4.4.

(i) Υποθέτουμε ότι, αν σε κάποια χρονική στιγμή επιθεώρησης ο βαθμός επιδείνωσης του μηχανισμού είναι ίσος με $i \le m$ και θέτουμε σε λειτουργία μία προληπτική συντήρηση, τότε ο απαιτούμενος χρόνος για την προληπτική συντήρηση είναι μία τυχαία μεταβλητή η οποία ακολουθεί τη Γεωμετρική κατανομή με πιθανότητα επιτυχίας a_i. Θεωρούμε ότι $a_0 \geq a_1 \geq \cdots \geq a_m \geq b.$ Αυτό σημαίνει ότι η αύξηση του βαθμού επιδείνωσης του μηχανισμού έχει ως επακόλουθο την αύξηση του αναμενόμενου χρόνου μιας προληπτικής συντήρησης. Πολλά αριθμητικά παραδείγματα παρέχουν ισχυρή ένδειξη ότι, για κάθε σταθερό περιεχόμενο x του αποθηκευτικού χώρου, η βέλτιστη πολιτική είναι μονότονη διότι θέτει σε λειτουργία μία προληπτική συντήρηση αν και μόνο αν ο βαθμός επιδείνωσης του μηχανισμού είναι μεγαλύτερος ή ίσος με μία κρίσιμη τιμή $i^*(x)$. Μία αναλυτική απόδειξη φαίνεται δύσκολη. Ένα αριθμητικό παράδειγμα παρουσιάζεται παρακάτω. Θεωρούμε την περίπτωση κατά την οποία $m=10,\ K=5,\ a_i=10(10+i)^{-1},\ 0\leq i\leq 10,\ b=0.1,\ p=3,\ d=2,\ c_p=0.4,\ c_f=0.8,\ h=0.2,$ $c_i=0.1(i+1),\quad \tilde{c}_i=0.05(i+1),\quad 0\leq i\leq 10.$ Οι μη-μηδενικές πιθανότητες μετάβασης p_{ij}, $0\leq i,j\leq 11,$ είναι τέτοιες ώστε:

$$p_{ij}=\frac{1}{12-i},\quad i\leq j\leq 11.$$

Η βέλτιστη πολιτική χαρακτηρίζεται από τις κρίσιμες τιμές $i^*(0)=6,\ \ i^*(1)=5,\ \ i^*(2)=2,$ $i^*(3)=i^*(4)=i^*(5)=0.$ Το μακροπρόθεσμο αναμενόμενο μέσο κόστος ανά μονάδα χρόνου βρέθηκε ίσο με 1.51.

(ii) Υποθέτουμε ότι, αν σε κάποια χρονική στιγμή επιθεώρησης ο βαθμός επιδείνωσης του μηχανισμού είναι ίσος με $i\leq m,$ τότε μία προληπτική συντήρηση είναι δυνατόν να αποκαταστήσει μερικώς το μηχανισμό έτσι ώστε να βρεθεί σε μία καλύτερη κατάσταση με βαθμό επιδείνωσης ίσο με j, όπου $j\in\{0,1,\ldots,i-1\}$. Η έννοια της «μερικής αποκατάστασης» χρησιμοποιήθηκε στην εργασία των Douer and Yechiali (1994) και στην εργασία των Benyamini and Yechiali (1999). Υποθέτουμε ότι ο απαιτούμενος χρόνος για μία μερική αποκατάσταση του μηχανισμού είναι μία τυχαία μεταβλητή η οποία ακολουθεί τη Γεωμετρική κατανομή με πιθανότητα επιτυχίας a_{i-j}. Αν ο βαθμός επιδείνωσης του μηχανισμού είναι ίσος με $m+1,$ μία επισκευή είναι υποχρεωτική και επαναφέρει το μηχανισμό σε αρίστη κατάσταση στην οποία ο βαθμός επιδείνωσής του ισούται με 0. Θεωρούμε ότι ο αναμενόμενος χρόνος για μία επισκευή του μηχανισμού είναι ίσος με b^{-1}. Υποθέτουμε επίσης ότι $a_0 \geq a_1 \geq \cdots \geq a_m \geq b$. Πολλά αριθμητικά παραδείγματα παρέχουν ισχυρή ένδειξη ότι, για κάθε σταθερό περιεχόμενο x του αποθηκευτικού χώρου, η βέλτιστη πολιτική είναι μονότονη διότι θέτει σε λειτουργία μία προληπτική συντήρηση αν και μόνο αν ο βαθμός επιδείνωσης του μηχανισμού είναι μεγαλύτερος ή ίσος με μία κρίσιμη τιμή $i^*(x)$. Μία αναλυτική απόδειξη φαίνεται δύσκολη. Ένα αριθμητικό παράδειγμα παρουσιάζεται παρακάτω. Θεωρούμε την περίπτωση κατά την οποία $m = 3$, $K = 3$, $a_i = (3+i)^{-1}$, $0 \leq i \leq 3$, $b = 0.15$, $p = 2$, $d = 1$, $c_p = 0.3$, $c_f = 0.6$, $h = 0.4$, $c_i = 0.1(i+1)$, $\tilde{c}_i = 0.05(i+1)$, $0 \leq i \leq 3$. Ο πίνακας P με τις πιθανότητες μετάβασης p_{ij}, $0 \leq i, j \leq 4$, δίνεται παρακάτω:

$$P = \begin{pmatrix} 0.1 & 0.7 & 0.1 & 0.05 & 0.05 \\ 0 & 0.8 & 0.1 & 0.05 & 0.05 \\ 0 & 0 & 0.5 & 0.25 & 0.25 \\ 0 & 0 & 0 & 0.5 & 0.5 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Η βέλτιστη πολιτική χαρακτηρίζεται από τις κρίσιμες τιμές $i^*(0) = 2$, $i^*(1) = 2$, $i^*(2) = i^*(3) = 0$. Το μακροπρόθεσμο αναμενόμενο μέσο κόστος ανά μονάδα χρόνου βρέθηκε ίσο με 0.72.

4.7 Οι χρόνοι προληπτικής συντήρησης και επισκευής είναι συνεχείς τυχαίες μεταβλητές

Στο παρόν εδάφιο τροποποιούμε το μοντέλο που περιγράψαμε στο Εδάφιο 4.4. Υποθέτουμε ότι οι απαιτούμενοι χρόνοι για μία προληπτική συντήρηση και μία επισκευή του μηχανισμού είναι συνεχείς τυχαίες μεταβλητές με συναρτήσεις πυκνότητας πιθανότητας $f_1(x)$ και $f_2(x)$, αντίστοιχα.

Έστω m_{PM} και m_{RE} οι αναμενόμενοι χρόνοι για μία προληπτική συντήρηση και μία επισκευή του μηχανισμού, αντίστοιχα. Οι Συνθήκες 1, 2, 4, 6 και 7 ισχύουν ενώ η Συνθήκη 3 διαμορφώνεται ως εξής:

Συνθήκη 3:* $m_{PM} < m_{RE}$. Δηλαδή, ο αναμενόμενος χρόνος για μία προληπτική συντήρηση είναι μικρότερος από τον αναμενόμενο χρόνο για μία επισκευή του μηχανισμού.

Σε κάθε χρονική στιγμή επιθεώρησης του μηχανισμού επιλέγεται μία ενέργεια a και η κατάσταση του συστήματος παραγωγής μπορεί να αναπαρασταθεί από το ζεύγος των μεταβλητών (i,x), όπου i είναι ο βαθμός επιδείνωσης του μηχανισμού και x είναι το περιεχόμενο του αποθηκευτικού χώρου. Ο χώρος καταστάσεων S του συστήματος είναι το ακόλουθο σύνολο:

$$S = \{(i,x),\, 0 \le i \le m+1,\, 0 \le x \le K\}.$$

Ο χώρος καταστάσεων S είναι ένας διδιάστατος χώρος καταστάσεων ο οποίος αποτελείται από μία διακριτή μεταβλητή, το βαθμό επιδείνωσης i του μηχανισμού και μία συνεχής μεταβλητή, το περιεχόμενο x του αποθηκευτικού χώρου. Για να μετατρέψουμε τη συνεχή μεταβλητή x σε μία διακριτή μεταβλητή εισάγουμε την παράμετρο ξ η οποία διαιρεί το διάστημα $[0,K]$ σε K/ξ τμήματα έτσι ώστε η μεταβλητή $j\xi$, $j = 0, 1, ..., K/\xi$, αναπαριστά το περιεχόμενο του αποθηκευτικού χώρου σε κάθε χρονική στιγμή επιθεώρησης του μηχανισμού. Ο χώρος καταστάσεων του συστήματος παραγωγής μπορεί τώρα να ξαναγραφεί κατά προσέγγιση με την ακόλουθη μορφή:

$$S = \{(i, j\xi),\, 0 \le i \le m+1,\, 0 \le j \le K/\xi\}.$$

Η προσέγγιση γίνεται καλύτερη όταν η παράμετρος ξ παίρνει πολύ μικρές τιμές (π.χ. $\xi = 0.001$).

Αν σε κάποια χρονική στιγμή επιθεώρησης το σύστημα βρίσκεται σε μία κατάσταση $(m+1,\, j\xi)$, $0 \le j\xi \le K$, επιλέγουμε υποχρεωτικά την ενέργεια $a = 2$ σύμφωνα με την οποία θέτουμε σε λειτουργία μία επισκευή του μηχανισμού. Η επισκευή επαναφέρει το σύστημα στην κατάσταση $(0,0)$ ή σε μία από τις καταστάσεις $(0, j'\xi),$ όπου $1 \le j' \le j.$

Αν σε κάποια χρονική στιγμή επιθεώρησης το σύστημα βρίσκεται σε μία κατάσταση $(i, j\xi),$ $0 \le i \le m,$ $0 \le j\xi \le K,$ επιλέγουμε την ενέργεια $a=1$ σύμφωνα με την οποία θέτουμε σε λειτουργία μία προληπτική συντήρηση ή την ενέργεια $a=0$ σύμφωνα με την οποία δεν επεμβαίνουμε στη λειτουργία του μηχανισμού. Αν επιλέξουμε την ενέργεια $a=1,$ μία προληπτική συντήρηση επαναφέρει το σύστημα στην κατάσταση $(0,0)$ ή σε μία από τις καταστάσεις $(0, j'\xi),$ όπου $1 \le j' \le j.$ Αν επιλέξουμε την ενέργεια $a=0$ και ισχύει ότι $0 \le j\xi \le K-1,$ το σύστημα μεταβαίνει σε μία από τις καταστάσεις $(r, j\xi+1),$ όπου $0 \le r \le m+1.$ Αν επιλέξουμε την ενέργεια $a=0$ και ισχύει ότι $j\xi > K-1,$ το σύστημα μεταβαίνει σε μία από τις καταστάσεις $(r, K),$ όπου $0 \le r \le m+1.$

Μία πολιτική είναι ένας κανόνας με τον οποίον επιλέγονται οι ενέργειες κατά τις χρονικές στιγμές επιθεώρησης του μηχανισμού. Θεωρούμε μία ημι-Μαρκοβιανή διαδικασία αποφάσεων με χώρο καταστάσεων S και αναζητούμε τη στάσιμη πολιτική η οποία, για οποιαδήποτε αρχική κατάσταση του συστήματος, ελαχιστοποιεί το μακροπρόθεσμο αναμενόμενο μέσο κόστος ανά μονάδα χρόνου. Η σχετική θεωρία βρίσκεται στο Εδάφιο 2.5 της παρούσας διατριβής και στο Κεφάλαιο 3 του βιβλίου του Tijms (1994). Οι χρονικές στιγμές κατά τις οποίες επιλέγεται μία ενέργεια a είναι όλες οι χρονικές στιγμές $\tau = 0, 1, \ldots$ στις οποίες το σύστημα εισέρχεται σε μία κατάσταση του χώρου καταστάσεων S.

Επισημαίνουμε ότι, υπό τον έλεγχο οποιασδήποτε πολιτικής, το σύστημα είναι δυνατόν να μεταβεί στην κατάσταση $(0,0)$ από οποιαδήποτε αρχική κατάσταση $s \in S.$ Το Πόρισμα 6.20 στη σελ. 149 του βιβλίου του Ross (1992) εξασφαλίζει την ύπαρξη μιας βέλτιστης στάσιμης πολιτικής διότι ο χώρος καταστάσεων S του συστήματος είναι πεπερασμένος.

Έστω $p_{su}(a)$ η πιθανότητα η επόμενη κατάσταση του συστήματος να είναι η κατάσταση $u \in S,$ αν η παρούσα κατάσταση είναι η κατάσταση $s \in S$ και επιλέγεται η ενέργεια $a,$ και έστω $T(s,a),$ $C(s,a)$ ο αναμενόμενος χρόνος και το αναμενόμενο κόστος, αντίστοιχα, για τη μετάβαση του συστήματος στην επόμενη κατάσταση, αν η παρούσα κατάσταση είναι η κατάσταση $s \in S$ και επιλέγεται η ενέργεια $a.$

Οι πιθανότητες μετάβασης του συστήματος για τις τρεις δυνατές ενέργειες δίνονται από τις ακόλουθες εξισώσεις (4.22)-(4.27).

$$p_{(m+1,j\xi)(0,0)}(2)=\int_{\frac{j\xi}{d}}^{\infty} f_2(t)dt,\ 0\le j\xi\le K, \tag{4.22}$$

$$p_{(m+1,j\xi)(0,j'\xi)}(2)=\int_{\frac{\xi(j-j')-\xi/2}{d}}^{\frac{\xi(j-j')+\xi/2}{d}} f_2(t)dt,\quad \xi\le j\xi\le K,\ 1\le j'\le j. \tag{4.23}$$

$$p_{(i,j\xi)(0,0)}(1)=\int_{\frac{j\xi}{d}}^{\infty} f_1(t)dt,\quad 0\le i\le m,\quad 0\le j\xi\le K, \tag{4.24}$$

$$p_{(i,j\xi)(0,j'\xi)}(1)=\int_{\frac{\xi(j-j')-\xi/2}{d}}^{\frac{\xi(j-j')+\xi/2}{d}} f_1(t)dt,\quad 0\le i\le m,\quad \xi\le j\xi\le K,\quad 1\le j'\le j, \tag{4.25}$$

$$p_{(i,j\xi)(r,j\xi+1)}(0)=p_{ir},\quad 0\le i\le m,\quad 0\le j\xi\le K-1,\quad 0\le r\le m+1, \tag{4.26}$$

$$p_{(i,j\xi)(r,K)}(0)=p_{ir},\ 0\le i\le m,\quad K-1+\xi\le j\xi\le K,\quad 0\le r\le m+1. \tag{4.27}$$

Οι αναμενόμενοι χρόνοι $T(s,a)$ για τις τρεις δυνατές ενέργειες δίνονται από τις ακόλουθες εξισώσεις (4.28)-(4.30).

$$T((m+1,j\xi),2)=\int_0^{\infty} tf_2(t)dt,\quad 0\le j\xi\le K, \tag{4.28}$$

$$T((i,j\xi),1)=\int_0^\infty tf_1(t)dt, \quad 0\le i\le m, \quad 0\le j\xi\le K, \tag{4.29}$$

$$T((i,j\xi),0)=1, \quad 0\le i\le m, \quad 0\le j\xi\le K. \tag{4.30}$$

Αν σε μία κατάσταση $s\in S,$ επιλέξουμε την ενέργεια $a=1$ ή την ενέργεια $a=2,$ οι εκφράσεις για τα αναμενόμενα κόστη $C(s,a)$ είναι περισσότερο πολύπλοκες. Αν θέσουμε σε λειτουργία μία προληπτική συντήρηση $(a=1)$ ή μία επισκευή $(a=2),$ τα αναμενόμενα κόστη $C(s,a),$ αποτελούνται από τα αναμενόμενα κόστη $C^M(s,a)$ της προληπτικής συντήρησης ή της επισκευής, τα αναμενόμενα κόστη $C^{SH}(s,a)$ εξαιτίας της έλλειψης του υλικού στον αποθηκευτικό χώρο και τα αναμενόμενα κόστη $C^H(s,a)$ της αποθήκευσης του υλικού στον αποθηκευτικό χώρο. Οι ποσότητες αυτές δίνονται από τις ακόλουθες εξισώσεις (4.31)-(4.33) στην περίπτωση κατά την οποία θέτουμε σε λειτουργία μία επισκευή, δηλαδή επιλέγουμε την ενέργεια $a=2.$ Με αντικατάσταση του $m+1$ με $i,\ 0\le i\le m,$ του c_f με c_p και του f_2 με f_1 στις εξισώσεις (4.31)-(4.33), προκύπτουν οι αντίστοιχες εκφράσεις για την περίπτωση κατά την οποία θέτουμε σε λειτουργία μία προληπτική συντήρηση, δηλαδή επιλέγουμε την ενέργεια $a=1.$ Για τους παρακάτω υπολογισμούς δεσμευόμαστε ως προς τη χρονική διάρκεια μιας προληπτικής συντήρησης $(a=1)$ ή μιας επισκευής $(a=2)$ του μηχανισμού. Κάθε μία από τις τρεις παρακάτω εξισώσεις είναι αποτέλεσμα συλλογισμών παρόμοιων με αυτούς που οδηγούν στην εξίσωση (4.7). Στις σελ. 82-85 της παρούσας διατριβής περιέχεται μία λεπτομερής περιγραφή του τρόπου με τον οποίον εξάγεται η σχέση (4.7).

$$C^M((m+1,j\xi),2)=c_f\int_0^\infty tf_2(t)dt, \quad 0\le j\xi\le K, \tag{4.31}$$

$$C^{SH}((m+1,j\xi),2)=\int_{\frac{j\xi}{d}}^\infty (td-j\xi)f_2(t)dt, \quad 0\le j\xi\le K, \tag{4.32}$$

$$C^H((m+1, j\xi),2) = \int_0^{\frac{j\xi}{d}} \left[\int_0^t h(j\xi - sd)ds \right] f_2(t)dt + \int_{\frac{j\xi}{d}}^{\infty} \left[\int_0^{\frac{j\xi}{d}} h(j\xi - sd)ds \right] f_2(t)dt, \quad 0 \le j\xi \le K. \quad (4.33)$$

Τα αναμενόμενα κόστη $C(s,a)$ για τις τρεις δυνατές ενέργειες δίνονται από τις ακόλουθες εξισώσεις (4.34)-(4.37).

$$C((m+1, j\xi),2) = C^M((m+1, j\xi),2) + C^{SH}((m+1, j\xi),2) + C^H((m+1, j\xi),2), \quad (4.34)$$

$$0 \le j\xi \le K,$$

$$C((i, j\xi),1) = C^M((i, j\xi),1) + C^{SH}((i, j\xi),1) + C^H((i, j\xi),1), \quad 0 \le i \le m, \quad 0 \le j\xi \le K, \quad (4.35)$$

$$C((i, j\xi),0) = c_i + hj\xi, \quad 0 \le i \le m, \quad 0 \le j\xi \le K - \xi, \quad (4.36)$$

$$C((i,K),0) = \tilde{c}_i + hK, \quad 0 \le i \le m. \quad (4.37)$$

Παρατηρούμε ότι στις εξισώσεις (4.36) και (4.37), οι οποίες αντιστοιχούν στην ενέργεια $a = 0$, τα αναμενόμενα κόστη $C(s,a)$ αποτελούνται από τα κόστη της λειτουργίας του μηχανισμού και τα κόστη της αποθήκευσης του υλικού στον αποθηκευτικό χώρο.

4.8 Ένας αλγόριθμος όταν οι χρόνοι προληπτικής συντήρησης και επισκευής είναι συνεχείς τυχαίες μεταβλητές

Στο μοντέλο του Εδαφίου 4.4 υποθέτουμε ότι οι απαιτούμενοι χρόνοι για μία προληπτική συντήρηση και μία επισκευή του μηχανισμού είναι τυχαίες μεταβλητές που ακολουθούν τη Γεωμετρική κατανομή. Σύμφωνα με την Πρόταση 4.2, αν ισχύουν οι Συνθήκες 1-4, 6 και 7, για κάθε σταθερό περιεχόμενο του αποθηκευτικού χώρου, η βέλτιστη πολιτική είναι μονότονη διότι θέτει σε λειτουργία μία προληπτική συντήρηση αν και μόνο αν ο βαθμός επιδείνωσης του μηχανισμού είναι μεγαλύτερος ή ίσος με μία κρίσιμη τιμή.

Στο μοντέλο του Εδαφίου 4.7 υποθέτουμε ότι οι απαιτούμενοι χρόνοι για μία προληπτική συντήρηση και μία επισκευή του μηχανισμού είναι συνεχείς τυχαίες μεταβλητές. Φαίνεται διαισθητικά λογικό ότι, αν ισχύουν οι Συνθήκες 1, 2, 4, 6, 7 και η Συνθήκη 3*, για κάθε σταθερό περιεχόμενο του αποθηκευτικού χώρου, η βέλτιστη πολιτική είναι επίσης μονότονη. Μία αναλυτική απόδειξη αυτής της εικασίας φαίνεται δύσκολη. Είναι ωστόσο δυνατόν να υπολογίσουμε τη βέλτιστη πολιτική με εφαρμογή του αλγορίθμου βελτίωσης των πολιτικών, του αλγορίθμου των διαδοχικών προσεγγίσεων και της μεθόδου του γραμμικού προγραμματισμού. Πολλά αριθμητικά παραδείγματα παρέχουν ισχυρή ένδειξη ότι, για κάθε σταθερό περιεχόμενο $j\xi,\ \ j=0,1,\dots,K/\xi,$ του αποθηκευτικού χώρου, η βέλτιστη πολιτική είναι πράγματι μονότονη διότι θέτει σε λειτουργία μία προληπτική συντήρηση αν και μόνο αν ο βαθμός επιδείνωσης του μηχανισμού είναι μεγαλύτερος ή ίσος με μία κρίσιμη τιμή $i^*(j\xi)$.

Ο αλγόριθμος που παρουσιάσαμε στο Εδάφιο 4.5 για το μοντέλο του Εδαφίου 4.4, μπορεί επίσης να σχεδιαστεί για το μοντέλο του Εδαφίου 4.7. Είναι ένας κατάλληλος αλγόριθμος βελτίωσης των πολιτικών ο οποίος εφαρμόζεται στο σύνολο των μονότονων πολιτικών και παράγει μία ακολουθία αυστηρά βελτιωμένων πολιτικών που έχουν τη μονότονη μορφή. Για κάθε σταθερό περιεχόμενο $j\xi,\ \ j=0,1,\dots,K/\xi,$ του αποθηκευτικού χώρου, ο αλγόριθμος παράγει πολιτικές οι οποίες θέτουν σε λειτουργία μία προληπτική συντήρηση αν και μόνο αν ο βαθμός επιδείνωσης του μηχανισμού είναι μεγαλύτερος ή ίσος με μία κρίσιμη τιμή $i(j\xi)$. Πολλά αριθμητικά παραδείγματα παρέχουν ισχυρή ένδειξη ότι η τελική πολιτική που δημιουργεί ο αλγόριθμος είναι η βέλτιστη πολιτική διότι συμπίπτει με την τελική πολιτική που δίνει ο αλγόριθμος βελτίωσης των πολιτικών και ο αλγόριθμος των διαδοχικών προσεγγίσεων. Ο σχεδιασμός του αλγορίθμου βασίζεται στην τεχνική εμφύτευσης του Tijms (βλέπε σελ. 234 του βιβλίου του Tijms (1994)) και παρουσιάζεται παρακάτω.

Έστω μία μονότονη πολιτική R η οποία χαρακτηρίζεται από τις κρίσιμες τιμές $i(j\xi)$, $j\in\{0,1,\dots,K/\xi\}$. Αν η μονότονη πολιτική R υιοθετηθεί για τον έλεγχο του συστήματος, τότε το σύστημα είναι δυνατόν να μεταβεί σε κάποια κατάσταση του συνόλου των καταστάσεων:

$$E = \bigcup_{j=0}^{K/\xi} \{(i, j\xi), 0 \le i \le i(j\xi)\},$$

από οποιαδήποτε αρχική κατάσταση $s \in S.$

Έστω $g(R)$ το μακροπρόθεσμο αναμενόμενο μέσο κόστος ανά μονάδα χρόνου υπό τον έλεγχο της πολιτικής R και T_s^E, C_s^E ο αναμενόμενος χρόνος και το αναμενόμενο κόστος, αντίστοιχα, μέχρι την πρώτη είσοδο του συστήματος στο σύνολο E αν η αρχική κατάσταση είναι η κατάσταση $s \in S.$ Έστω επίσης $p_{s\ell}^E, \; s \in S, \; \ell \in E,$ η πιθανότητα η πρώτη κατάσταση κατά την είσοδο του συστήματος στο σύνολο E να είναι η κατάσταση $\ell,$ δοθέντος ότι η πολιτική R έχει υιοθετηθεί για τον έλεγχο του συστήματος και η αρχική κατάσταση είναι η κατάσταση $s \in S.$ Υποθέτουμε ότι, αν η αρχική κατάσταση του συστήματος είναι μία κατάσταση $s \in E,$ τότε η πρώτη κατάσταση κατά την είσοδο του συστήματος στο σύνολο E είναι η κατάσταση στην οποία μεταβαίνει το σύστημα στην επόμενη επιστροφή του στο σύνολο $E.$

Έστω $w(s), s \in S,$ οι σχετικές τιμές της πολιτικής R (βλέπε σελ. 25 της παρούσας διατριβής). Σύμφωνα με τη σχέση (3.6.1) στη σελ. 235 του βιβλίου του Tijms (1994) οι σχετικές τιμές $w(s), s \in S,$ ικανοποιούν το ακόλουθο σύστημα των γραμμικών εξισώσεων:

$$w(s) = C_s^E - g(R)T_s^E + \sum_{\ell \in E} p_{s\ell}^E w(\ell), \quad s \in S, \tag{4.38}$$

μαζί με την εξίσωση

$$w(0, 0) = 0. \tag{4.39}$$

Οι ποσότητες T_s^E, C_s^E και $p_{s\ell}^E, \; s \in S, \; \ell \in E,$ μπορούν να υπολογιστούν εύκολα αν λάβουμε υπόψη τις μεταβάσεις του συστήματος υπό τον έλεγχο της μονότονης πολιτικής R και χρησιμοποιήσουμε διάφορες μορφές του Θεωρήματος Ολικής Πιθανότητας. Παρακάτω παρουσιάζουμε κάποιες από αυτές τις εκφράσεις.

$$T^E_{(i,j\xi)} = 1 + \sum_{r=i(j\xi+1)+1}^{m} p_{ir}\left[\int_0^\infty t f_1(t)dt\right] + p_{i,m+1}\int_0^\infty t f_2(t)dt, \quad 0 \le i < i(j\xi), \quad 0 \le j\xi \le K-1,$$

$$C^E_{(i,j\xi)} = c_i + hj\xi + \sum_{r=i(j\xi+1)+1}^{m} p_{ir}[C^M((i,j\xi+1),1) + C^{SH}((i,j\xi+1),1) + C^H((i,j\xi+1),1)]$$
$$+ p_{i,m+1}[C^M((m+1,j\xi+1),2) + C^{SH}((m+1,j\xi+1),2) + C^H((m+1,j\xi+1),2)],$$
$$0 \le i < i(j\xi), \quad 0 \le j\xi \le K-1,$$

$$C^E_{(i,j\xi)} = C^M((i,j\xi),1) + C^{SH}((i,j\xi),1) + C^H((i,j\xi),1), \quad i(j\xi) \le i \le m, \quad 0 \le j\xi \le K,$$

$$p^E_{(i,j\xi)(r,j\xi+1)} = p_{ir}, \quad 0 \le i < i(j\xi), \quad 0 \le r \le i(j\xi+1), \quad 0 \le j\xi \le K-1,$$

$$p^E_{(i,j\xi)(0,0)} = \int_{\frac{j\xi}{d}}^{\infty} f_1(t)dt, \quad i(j\xi) \le i \le m, \quad 0 \le j\xi \le K.$$

Οι σχετικές τιμές $w(s),\ s \in E,$ και το μακροπρόθεσμο αναμενόμενο μέσο κόστος $g(R)$ ανά μονάδα χρόνου, υπό τον έλεγχο της πολιτικής $R,$ μπορούν να υπολογιστούν επιλύνοντας το ακόλουθο σύστημα των γραμμικών εξισώσεων:

$$w(s) = C^E_s - g(R)T^E_s + \sum_{\ell \in E} p^E_{s\ell} w(\ell), \quad s \in E, \tag{4.40}$$

και την εξίσωση (4.39). Η ποσότητα $Q_R(s;a)$ ορίζεται ως εξής:

$$Q_R(s;a) = C(s,a) - g(R)T(s,a) + \sum_{\ell \in S} p_{s\ell}(a)w(\ell), \quad s \in \{(i,j\xi), 0 \le i \le m, 0 \le j \le K/\xi\}, \tag{4.41}$$
$$a \in \{0,1\},$$

όπου, $p_{s\ell}(a)$ είναι η πιθανότητα η επόμενη κατάσταση του συστήματος να είναι η κατάσταση ℓ, αν η παρούσα κατάσταση είναι η κατάσταση s και επιλέγεται η ενέργεια a, και $T(s,a)$, $C(s,a)$ είναι ο αναμενόμενος χρόνος και το αναμενόμενο κόστος, αντίστοιχα, για τη μετάβαση του συστήματος στην επόμενη κατάσταση, αν η παρούσα κατάσταση είναι η κατάσταση s και επιλέγεται η ενέργεια a. Οι πιθανότητες μετάβασης $p_{s\ell}(a)$ δίνονται από τις εξισώσεις (4.22)-(4.27), οι αναμενόμενοι χρόνοι $T(s,a)$ δίνονται από τις εξισώσεις (4.28)-(4.30) και τα αναμενόμενα κόστη $C(s,a)$ δίνονται από τις εξισώσεις (4.34)-(4.37).

Υποθέτουμε ότι για κάποιο περιεχόμενο $j\xi$, $0 \le j\xi \le K$, του αποθηκευτικού χώρου, υπάρχει ένας ακέραιος αριθμός $\tilde{i}(j\xi)$ τέτοιος ώστε $0 \le \tilde{i}(j\xi) < i(j\xi)$ και $Q_R((i, j\xi);1) < w(i, j\xi)$, $\tilde{i}(j\xi) \le i < i(j\xi)$. Τότε, σύμφωνα με το Θεώρημα 2.7 για τις ημι-Μαρκοβιανές διαδικασίες αποφάσεων (βλέπε (ε) στη σελ. 28 της παρούσας διατριβής), το μακροπρόθεσμο αναμενόμενο μέσο κόστος υπό τον έλεγχο της μονότονης πολιτικής η οποία χαρακτηρίζεται από τις κρίσιμες τιμές $i(0), ..., i(j\xi - \xi), \tilde{i}(j\xi), i(j\xi + \xi), ..., i(K)$ είναι μικρότερο από $g(R)$.

Παρόμοια, αν για κάποιο περιεχόμενο $j\xi$, $0 \le j\xi \le K$, του αποθηκευτικού χώρου, υπάρχει ένας ακέραιος αριθμός $\tilde{i}(j\xi)$ τέτοιος ώστε $i(j\xi) < \tilde{i}(j\xi) \le m+1$ και $Q_R((i, j\xi);0) < w(i, j\xi)$, $i(j\xi) \le i < \tilde{i}(j\xi)$, τότε, σύμφωνα πάλι με το Θεώρημα 2.7 για τις ημι-Μαρκοβιανές διαδικασίες αποφάσεων, το μακροπρόθεσμο αναμενόμενο μέσο κόστος υπό τον έλεγχο της μονότονης πολιτικής η οποία χαρακτηρίζεται από τις κρίσιμες τιμές $i(0), ..., i(j\xi - \xi), \tilde{i}(j\xi), i(j\xi + \xi), ..., i(K)$ είναι μικρότερο από $g(R)$.

Οι παραπάνω παρατηρήσεις μας οδηγούν στο σχεδιασμό του ακόλουθου αλγορίθμου ο οποίος παράγει μία ακολουθία αυστηρά βελτιωμένων μονότονων πολιτικών.

Αλγόριθμος

Βήμα 1. *(Εναρξη)*. Επιλέγουμε μία αρχική μονότονη πολιτική R η οποία χαρακτηρίζεται από τις κρίσιμες τιμές $i(j\xi)$, $j = 0, 1, ..., K/\xi$.

Βήμα 2. *(Εύρεση σχετικών τιμών και μέσου κόστους)*. Για τη μονότονη πολιτική R, υπολογίζουμε το μακροπρόθεσμο αναμενόμενο μέσο κόστος $g(R)$ ανά μονάδα χρόνου και τις σχετικές τιμές $w(s)$, $s \in E$, επιλύνοντας το σύστημα των γραμμικών εξισώσεων (4.39), (4.40).

Βήμα 3. *(Βελτίωση των πολιτικών).* Για κάθε σταθερό περιεχόμενο $j\xi, \quad j = 0, 1, ..., K/\xi,$ του αποθηκευτικού χώρου:

(α) Βρίσκουμε, αν υπάρχει, το μικρότερο ακέραιο αριθμό $\tilde{i}\,(j\xi)$ τέτοιο ώστε $0 \le \tilde{i}\,(j\xi) < i(j\xi)$ και $Q_R((i, j\xi);1) < w(i, j\xi), \ \tilde{i}\,(j\xi) \le i < i(j\xi),$

διαφορετικά,

(β) βρίσκουμε, αν υπάρχει, το μεγαλύτερο ακέραιο αριθμό $\tilde{i}\,(j\xi)$ τέτοιο ώστε $i(j\xi) < \tilde{i}\,(j\xi) \le m+1$ και $Q_R((i, j\xi);0) < w(i, j\xi), \ i(j\xi) \le i < \tilde{i}\,(j\xi).$

Οι ποσότητες $Q_R((i, j\xi);1)$ και $Q_R((i, j\xi);0)$ δίνονται από την εξίσωση (4.41). Οι σχετικές τιμές $w(s), \ s \in S - E,$ υπολογίζονται από την εξίσωση (4.38).

Αντικαθιστούμε την κρίσιμη τιμή $i(j\xi)$ με την κρίσιμη τιμή $\tilde{i}\,(j\xi)$ για εκείνες τις τιμές του $j\xi, \quad 0 \le j\xi \le K,$ για τις οποίες είναι δυνατόν να βρούμε μία κρίσιμη τιμή $\tilde{i}\,(j\xi)$ και επιστρέφουμε στο Βήμα 2.

Βήμα 4. *(Κριτήριο Σύγκλισης).* Αν δεν είναι δυνατόν να βρεθεί καμία κρίσιμη τιμή $\tilde{i}\,(j\xi), \ 0 \le j\xi \le K,$ ο αλγόριθμος σταματά. Η τελική πολιτική που παράγει ο αλγόριθμος είναι η πολιτική R με μακροπρόθεσμο αναμενόμενο μέσο κόστος $g(R)$ ανά μονάδα χρόνου.

Ο αλγόριθμος παράγει μία ακολουθία αυστηρά βελτιωμένων μονότονων πολιτικών και σταματά μετά από ένα πεπερασμένο αριθμό επαναλήψεων διότι το σύνολο των μονότονων πολιτικών είναι πεπερασμένο. Πολλά αριθμητικά παραδείγματα παρέχουν ισχυρή ένδειξη ότι η τελική πολιτική που δημιουργεί ο αλγόριθμος είναι η βέλτιστη πολιτική διότι συμπίπτει με την τελική πολιτική που δίνει ο αλγόριθμος βελτίωσης των πολιτικών και ο αλγόριθμος των διαδοχικών προσεγγίσεων. Ο απαιτούμενος υπολογιστικός χρόνος για την εκτέλεση του αλγορίθμου είναι κατά πολύ μικρότερος από τον απαιτούμενο υπολογιστικό χρόνο για την εκτέλεση του αλγορίθμου βελτίωσης των πολιτικών. Αυτό οφείλεται στο γεγονός ότι ο αριθμός των αγνώστων στο σύστημα των γραμμικών εξισώσεων του Βήματος 2 του αλγορίθμου είναι ίσος με τον αριθμό των στοιχείων του συνόλου $E,$ ενώ ο αριθμός των αγνώστων στο σύστημα των γραμμικών εξισώσεων του αντίστοιχου βήματος της εύρεσης των σχετικών τιμών και του μέσου κόστους στον αλγόριθμο βελτίωσης των πολιτικών είναι ίσος με τον αριθμό των στοιχείων του χώρου καταστάσεων $S.$ Από πολλά αριθμητικά παραδείγματα που έχουμε εξετάσει συμπεραίνουμε επίσης ότι, ο αριθμός των επαναλήψεων του αλγορίθμου δεν επηρεάζεται ιδιαίτερα από την αρχική μονότονη πολιτική που επιλέγουμε κατά την έναρξη του αλγορίθμου στο Βήμα 1.

Δύο αριθμητικά παραδείγματα παρουσιάζονται παρακάτω. Στο πρώτο παράδειγμα θεωρούμε την περίπτωση κατά την οποία $m = 10,\ \ K = 5,\ \ p = 4,\ \ d = 3,\ \ c_p = 0.2,\ \ c_f = 0.4,\ \ h = 0.3,$ $c_i = 0.1(i+1),\ \ \tilde{c}_i = 0.05(i+1),\ \ 0 \le i \le m.$ Υποθέτουμε ότι οι μη-μηδενικές πιθανότητες μετάβασης $p_{ir},\ 0 \le i, r \le m+1$ είναι τέτοιες ώστε:

$$p_{ir} = \frac{1}{m+2-i}, \qquad i \le r \le m+1.$$

Αυτό σημαίνει ότι αν ο παρών βαθμός επιδείνωσης του μηχανισμού είναι ίσος με $i,$ τότε ο επόμενος βαθμός επιδείνωσης κατανέμεται ομοιόμορφα στο σύνολο $\{i, i+1, ..., m+1\}$. Οι πιθανότητες μετάβασης ικανοποιούν τη Συνθήκη 7 διότι, για κάθε $k = 0, ..., m+1,$ η ποσότητα:

$$\sum_{r=k}^{m+1} p_{ir} = \frac{m+2-k}{m+2-i}$$

είναι αύξουσα ως προς $i,\ 0 \le i \le 10.$ Υποθέτουμε ότι ο απαιτούμενος χρόνος για μία προληπτική συντήρηση του μηχανισμού είναι μία τυχαία μεταβλητή η οποία ακολουθεί την κατανομή Γάμα με παραμέτρους a_1 και b_1. Υποθέτουμε επίσης ότι ο απαιτούμενος χρόνος για μία επισκευή του μηχανισμού είναι μία τυχαία μεταβλητή η οποία ακολουθεί την κατανομή Γάμα με παραμέτρους a_2 και b_2. Οι συναρτήσεις πυκνότητας πιθανότητας δίνονται από τις εξισώσεις:

$$f_1(t) = \frac{1}{\Gamma(a_1) b_1^{a_1}} t^{a_1 - 1} \exp(-t b_1^{-1}) \quad \text{και} \quad f_2(t) = \frac{1}{\Gamma(a_2) b_2^{a_2}} t^{a_2 - 1} \exp(-t b_2^{-1}),\ \ t \ge 0,$$

αντίστοιχα, όπου, $\Gamma(\cdot)$ είναι η συνάρτηση Γάμα. Θεωρούμε ότι $a_1 = 2,\ \ b_1 = 1$ και $a_2 = 2,$ $b_2 = 3.$ Η Συνθήκη 3* ισχύει διότι η αναμενόμενη τιμή της κατανομής Γάμα με παραμέτρους a

και b είναι ίση με ab. Θεωρούμε ότι $\xi = 0.05$ έτσι ώστε το διάστημα $[0, K]$ διαιρείται σε $K/\xi = 100$ τμήματα. Ως αρχική μονότονη πολιτική στο Βήμα 1 του αλγορίθμου επιλέγουμε την πολιτική η οποία χαρακτηρίζεται από τις κρίσιμες τιμές $i(j\xi) = 10, \quad j = 0, ..., 100.$ Στον Πίνακα 4.2 παρουσιάζονται οι διαδοχικές μονότονες πολιτικές που παράγει ο αλγόριθμος και τα αντίστοιχα μακροπρόθεσμα αναμενόμενα μέσα κόστη τους ανά μονάδα χρόνου.

Πίνακας 4.2. Οι διαδοχικές μονότονες πολιτικές

Οι κρίσιμες τιμές $i(j\xi), \quad j = 0, ..., 100,$ των πολιτικών	Μέσο κόστος
$i(0) = \cdots = i(5) = 10$	2.2381
$i(0) = \cdots = i(3\xi) = 7, \; i(4\xi) = \cdots = i(8\xi) = 6,$ $i(9\xi) = \cdots = i(11\xi) = 5, \; i(12\xi) = \cdots = i(15\xi) = 4,$ $i(16\xi) = \cdots = i(20\xi) = 3, \; i(21\xi) = \cdots = i(23\xi) = 2,$ $i(24\xi) = \cdots = i(26\xi) = 1, \; i(27\xi) = \cdots = i(5) = 0$	1.9033
$i(0) = \cdots = i(14\xi) = 9, \; i(15\xi) = \cdots = i(30\xi) = 8,$ $i(31\xi) = \cdots = i(43\xi) = 7, \; i(44\xi) = \cdots = i(54\xi) = 6,$ $i(55\xi) = \cdots = i(64\xi) = 5, \; i(65\xi) = \cdots = i(72\xi) = 4,$ $i(75\xi) = \cdots = i(80\xi) = 3, \; i(81\xi) = \cdots = i(86\xi) = 2,$ $i(87\xi) = \cdots = i(92\xi) = 1, \; i(93\xi) = \cdots = i(5) = 0$	1.7970
$i(0) = \cdots = i(6\xi) = 9, \; i(7\xi) = \cdots = i(25\xi) = 8,$ $i(26\xi) = \cdots = i(39\xi) = 7, \; i(40\xi) = \cdots = i(50\xi) = 6,$ $i(51\xi) = \cdots = i(59\xi) = 5, \; i(60\xi) = \cdots = i(67\xi) = 4,$ $i(68\xi) = \cdots = i(75\xi) = 3, \; i(76\xi) = \cdots = i(81\xi) = 2,$ $i(82\xi) = \cdots = i(86\xi) = 1, \; i(87\xi) = \cdots = i(5) = 0$	1.7966

Ο απαιτούμενος υπολογιστικός χρόνος για την εκτέλεση του αλγορίθμου ήταν 30.4 δευτερόλεπτα ενώ ο απαιτούμενος υπολογιστικός χρόνος για την εκτέλεση του αλγόριθμου βελτίωσης των πολιτικών ήταν 59.2 δευτερόλεπτα.

Στο δεύτερο παράδειγμα θεωρούμε την περίπτωση κατά την οποία $m = 20$, $K = 10$, $p = 3$, $d = 2$, $c_p = 0.4$, $c_f = 0.8$, $h = 0.2$. Υποθέτουμε ότι τα κόστη της λειτουργίας c_i, $\tilde{c}_i$, $0 \le i \le m$ και οι μη-μηδενικές πιθανότητες μετάβασης p_{ir}, $0 \le i, r \le m+1$, συμπίπτουν με το πρώτο παράδειγμα. Υποθέτουμε ότι ο απαιτούμενος χρόνος για μία προληπτική συντήρηση του μηχανισμού είναι μία τυχαία μεταβλητή η οποία ακολουθεί την κατανομή Weibull με παραμέτρους α_1 και λ_1. Υποθέτουμε επίσης ότι ο απαιτούμενος χρόνος για μία επισκευή του μηχανισμού είναι μία τυχαία μεταβλητή η οποία ακολουθεί την κατανομή Weibull με παραμέτρους α_2 και λ_2. Οι συναρτήσεις πυκνότητας πιθανότητας δίνονται από τις εξισώσεις:

$$f_1(t) = \alpha_1 \lambda_1 (\lambda_1 t)^{\alpha_1 - 1} \exp[-(\lambda_1 t)^{\alpha_1}] \quad \text{και} \quad f_2(t) = \alpha_2 \lambda_2 (\lambda_2 t)^{\alpha_2 - 1} \exp[-(\lambda_2 t)^{\alpha_2}], \; t \ge 0,$$

αντίστοιχα. Θεωρούμε ότι $\alpha_1 = 1$, $\lambda_1 = 3$ και $\alpha_2 = 0.5, \lambda_2 = 5$. Η Συνθήκη 3* ισχύει διότι η αναμενόμενη τιμή της κατανομής Weibull με παραμέτρους α και λ είναι ίση με $\lambda^{-1}\Gamma(1+\alpha^{-1})$, όπου, $\Gamma(\cdot)$ είναι η συνάρτηση Γάμα. Θεωρούμε πάλι ότι $\xi = 0.05$ έτσι ώστε το διάστημα $[0, K]$ διαιρείται σε $K/\xi = 200$ τμήματα. Ως αρχική μονότονη πολιτική στο Βήμα 1 του αλγορίθμου επιλέγουμε την πολιτική η οποία χαρακτηρίζεται από τις κρίσιμες τιμές $i(j\xi) = 20$, $j = 0, ..., 200$. Ο αλγόριθμος σταματά μετά από πέντε επαναλήψεις. Η τελική πολιτική που παράγει ο αλγόριθμος είναι η βέλτιστη πολιτική και χαρακτηρίζεται από τις ακόλουθες κρίσιμες τιμές $i^*(j\xi)$, $j = 0, ..., 200$:

$$\begin{aligned}
&i^*(0) = \cdots = i^*(10\xi) = 14, i^*(11\xi) = \cdots = i^*(35\xi) = 12, i^*(36\xi) = \cdots = i^*(42\xi) = 10,\\
&i^*(43\xi) = \cdots = i^*(52\xi) = 9, i^*(53\xi) = \cdots = i^*(62\xi) = 8, i^*(63\xi) = \cdots = i^*(74\xi) = 7,\\
&i^*(75\xi) = \cdots = i^*(82\xi) = 6, i^*(83\xi) = \cdots = i^*(90\xi) = 5, i^*(91\xi) = \cdots = i^*(96\xi) = 4,\\
&i^*(97\xi) = \cdots = i^*(105\xi) = 3, i^*(106\xi) = \cdots = i^*(110\xi) = 2, i^*(111\xi) = \cdots = i^*(135\xi) = 1,\\
&i^*(136\xi) = \cdots = i^*(10) = 0.
\end{aligned}$$

Το μακροπρόθεσμο αναμενόμενο μέσο κόστος ανά μονάδα χρόνου για τη βέλτιστη πολιτική βρέθηκε ίσο με 0.46. Ο απαιτούμενος υπολογιστικός χρόνος για την εκτέλεση του αλγορίθμου ήταν 36.9 δευτερόλεπτα ενώ ο απαιτούμενος υπολογιστικός χρόνος για την εκτέλεση του αλγόριθμου βελτίωσης των πολιτικών ήταν 72.7 δευτερόλεπτα.

Σημείωση 4.3. Από τα δύο παραπάνω αριθμητικά παραδείγματα αλλά και από πολλά άλλα παραδείγματα που έχουμε εξετάσει υπάρχει ισχυρή ένδειξη ότι, για κάθε σταθερό περιεχόμενο $j\xi,\;\; j = 0, 1, ..., K/\xi,$ του αποθηκευτικού χώρου, η κρίσιμη τιμή $i^*(j\xi)$ η οποία, χαρακτηρίζει τη μονότονη μορφή της βέλτιστης πολιτικής, είναι φθίνουσα ως προς j.

ΚΕΦΑΛΑΙΟ 5

Ένας αλγόριθμος για τον υπολογισμό της βέλτιστης πολιτικής επισκευής ή αντικατάστασης ενός συστήματος

5.1 Εισαγωγή

Ένας μεγάλος αριθμός στοχαστικών μοντέλων έχουν επινοηθεί για να περιγράψουν τη συμπεριφορά ενός συστήματος το οποίο επιδεινώνεται με την πάροδο του χρόνου και η λειτουργία του διακόπτεται εξαιτίας ενδεχόμενων βλαβών. Μία ενέργεια η οποία μπορεί να ελέγξει την κατάσταση του συστήματος είναι μία επισκευή του. Μία άλλη ενέργεια είναι η αντικατάσταση του συστήματος από ένα καινούργιο σύστημα. Στα περισσότερα από αυτά τα μοντέλα (βλέπε π.χ. Barlow and Proschan (1965), Pierskala and Voelker (1976), Nakagawa and Kowada (1983), Phelps (1981, 1983)) οι ερευνητές υπέθεσαν ότι υπάρχουν δύο μόνο είδη επισκευής του συστήματος, η τέλεια επισκευή και η ελάχιστη επισκευή. Η τέλεια επισκευή επαναφέρει το σύστημα σε αρίστη κατάσταση ενώ η ελάχιστη επισκευή αποκαθιστά μερικώς το σύστημα και το επαναφέρει στην κατάσταση στην οποία βρισκόταν λίγο πριν υποστεί μία βλάβη.

Οι Kijima et al. (1988) μελέτησαν ένα γενικότερο μοντέλο στο οποίο μία επισκευή είναι δυνατόν να επαναφέρει το σύστημα σε μία ενδιάμεση κατάσταση ανάμεσα στην αρίστη και σε αυτήν στην οποία βρισκόταν λίγο πριν υποστεί μία βλάβη. Στο μοντέλο των Kijima et al. η τέλεια και η ελάχιστη επισκευή είναι δύο ειδικές περιπτώσεις. Υπέθεσαν ότι τα κόστη της επισκευής και της αντικατάστασης είναι σταθερά και θεώρησαν ότι το σύστημα μπορεί να αντικατασταθεί περιοδικά μόνο σε προκαθορισμένες χρονικές στιγμές $kT, \; k = 0, 1, ...$ και να επισκευαστεί οποτεδήποτε έχει υποστεί μία βλάβη. Κύριος στόχος της εργασίας είναι η εύρεση της τιμής της περιόδου T η οποία ελαχιστοποιεί το μακροπρόθεσμο αναμενόμενο μέσο κόστος ανά μονάδα χρόνου. Για το σκοπό αυτό πρότειναν μία προσεγγιστική μέθοδο.

Οι Makis and Jardine (1993) γενίκευσαν το μοντέλο των Kijima et al. κατασκευάζοντας μία κατάλληλη ημι-Μαρκοβιανή διαδικασία αποφάσεων. Θεώρησαν ότι σε κάθε χρονική στιγμή κατά την οποία το σύστημα έχει υποστεί μία βλάβη, μία επισκευή ή μία αντικατάστασή του είναι δυνατές. Ένας διδιάστατος άπειρος χώρος καταστάσεων χρησιμοποιήθηκε στον οποίον η κατάσταση της διαδικασίας αποτελείται από τον αριθμό των βλαβών που έχει υποστεί το σύστημα και από την ηλικία του. Ο αριθμός των βλαβών είναι μία διακριτή μεταβλητή ενώ η ηλικία του συστήματος είναι μία συνεχής μεταβλητή. Υπέθεσαν ότι το κόστος της αντικατάστασης είναι σταθερό και ότι το κόστος της επισκευής εξαρτάται από τον αριθμό των βλαβών που έχει υποστεί το σύστημα και από την ηλικία του. Όρισαν κατάλληλες συνθήκες οι οποίες αφορούν το ρυθμό επιδείνωσης (failure rate) του συστήματος και τα κόστη της επισκευής και της αντικατάστασής του. Μία πολιτική είναι ένας κανόνας επιλογής ενεργειών σε κάθε χρονική στιγμή κατά την οποία το σύστημα έχει υποστεί μία βλάβη και μία επισκευή ή μία αντικατάστασή του είναι δυνατές. Οι Makis and Jardine απέδειξαν ότι η στάσιμη πολιτική η οποία ελαχιστοποιεί το μακροπρόθεσμο αναμενόμενο μέσο κόστος ανά μονάδα χρόνου είναι μονότονη διότι αντικαθιστά το σύστημα όταν έχει υποστεί την $n-$οστή βλάβη αν και μόνο αν η ηλικία του είναι μεγαλύτερη ή ίση με μία κρίσιμη τιμή η οποία εξαρτάται από τον αριθμό n.

Οι Love et al. (2000) προσέγγισαν την ημι-Μαρκοβιανή διαδικασία αποφάσεων των Makis and Jardine με μία ημι-Μαρκοβιανή διαδικασία αποφάσεων με πεπερασμένο και διακριτό χώρο καταστάσεων αποκόπτοντας τον άπειρο χώρο καταστάσεων και μετατρέποντας τη συνεχή μεταβλητή της ηλικίας του συστήματος σε μία διακριτή μεταβλητή. Υπολόγισαν τις πιθανότητες μετάβασης και τους αναμενόμενους χρόνους της ημι-Μαρκοβιανής διαδικασίας αποφάσεων με τον πεπερασμένο χώρο καταστάσεων και σχεδίασαν έναν αλγόριθμο ο οποίος παράγει μία ακολουθία αυστηρά βελτιωμένων μονότονων πολιτικών. Στην τελευταία παράγραφο του Εδαφίου 3 της εργασίας τους αναφέρουν ότι ο αλγόριθμος υπολογίζει τη βέλτιστη μονότονη πολιτική. Για τον ισχυρισμό αυτόν, όπως θα εξηγήσουμε στο Εδάφιο 5.3, δεν υπάρχει μία αυστηρή απόδειξη αν και πολλά αριθμητικά παραδείγματα παρέχουν ισχυρή ένδειξη ότι είναι αληθινός.

Κύριος στόχος του παρόντος κεφαλαίου είναι να βελτιώσουμε τον αλγόριθμο των Love et al. προτείνοντας έναν κατάλληλο αλγόριθμο βελτίωσης των πολιτικών ο οποίος βασίζεται στην τεχνική εμφύτευσης του Tijms (1994). Στο παρόν πρόβλημα ο υπολογισμός όλων των απαραίτητων ποσοτήτων για την εφαρμογή αυτής της τεχνικής είναι εφικτός. Η τεχνική εμφύτευσης του Tijms ελαττώνει σημαντικά τους υπολογισμούς για την εύρεση των σχετικών τιμών και του μέσου κόστους των μονότονων πολιτικών, οι οποίοι έχουν αυξηθεί σημαντικά εξαιτίας της μετατροπής της ηλικίας του συστήματος σε μία διακριτή μεταβλητή.

Στο επόμενο εδάφιο περιγράφουμε το πρόβλημα και παρουσιάζουμε τη σχετική ημι-Μαρκοβιανή διαδικασία αποφάσεων με τον πεπερασμένο χώρο καταστάσεων. Στο Εδάφιο 5.3

σχεδιάζουμε έναν κατάλληλο αλγόριθμο βελτίωσης των πολιτικών εφαρμόζοντας την τεχνική εμφύτευσης του Tijms. Στο Εδάφιο 5.4 παρουσιάζουμε δύο αριθμητικά παραδείγματα στα οποία ο χρόνος ζωής ενός νέου συστήματος ακολουθεί την κατανομή Γάμα και την κατανομή Weibull, αντίστοιχα.

Τα αποτελέσματα των Εδαφίων 5.2-5.4 περιλαμβάνονται στην εργασία των Dimitrakos and Kyriakidis (2005a) η οποία έχει υποβληθεί προς δημοσίευση.

5.2 Κατασκευή του μοντέλου

Θεωρούμε ένα σύστημα το οποίο επιδεινώνεται με την πάροδο του χρόνου και η λειτουργία του διακόπτεται εξαιτίας ενδεχόμενων βλαβών. Σε κάθε χρονική στιγμή κατά την οποία το σύστημα έχει υποστεί μία βλάβη η κατάστασή του μπορεί να αναπαρασταθεί από το ζεύγος των μεταβλητών (n, t_n), όπου n είναι η $n-$οστή βλάβη και t_n είναι η ηλικία του συστήματος τη χρονική στιγμή της $n-$οστής βλάβης. Ο χώρος καταστάσεων είναι το ακόλουθο σύνολο:

$$I = \{(n, t_n), n = 1, 2, ..., t_n \geq 0\}.$$

Έστω N ο μέγιστος αριθμός των βλαβών που μπορεί να υποστεί το σύστημα πριν από μία αντικατάστασή του και έστω B ένα άνω φράγμα για την ηλικία του. Ακολουθώντας τους Love et al. (2000) μετατρέπουμε τη συνεχή μεταβλητή t_n σε μία διακριτή μεταβλητή ως εξής. Ο άξονας της ηλικίας διαιρείται σε ένα σύνολο ισομηκών ηλικιακών τμημάτων (age slices) και θεωρούμε ότι η $n-$οστή βλάβη συμβαίνει στο ηλικιακό τμήμα i_n. Εισάγουμε την παράμετρο ξ η οποία ορίζεται ως ο αριθμός των ηλικιακών τμημάτων σε μία μονάδα του χρόνου για να συσχετίσουμε το ηλικιακό τμήμα i_n με την ηλικία t_n. Αν η $n-$οστή βλάβη συμβεί στο ηλικιακό τμήμα i_n, υποθέτουμε ότι η ηλικία t_n του συστήματος τη χρονική στιγμή της $n-$οστής βλάβης βρίσκεται ανάμεσα στους αριθμούς i_n / ξ και $(i_n + 1)/\xi$.

Από εδώ και στο εξής παραλείπουμε το δείκτη n στα ηλικιακά τμήματα i_n και τα αναπαριστούμε απλά με i. Ο χώρος καταστάσεων I προσεγγίζεται από το ακόλουθο πεπερασμένο και διακριτό σύνολο:

$$S = \{(n,i), 1 \le n \le N, 0 \le i \le M\},$$

όπου, $M = B\xi$. Παρατηρούμε ότι μπορούμε να αυξήσουμε την ακρίβεια της προσέγγισης αυξάνοντας τις τιμές των ξ, N και M.

Στις χρονικές στιγμές κατά τις οποίες το σύστημα έχει υποστεί μία βλάβη, μία ενέργεια a μπορεί να επιλεγεί. Το σύστημα μπορεί να επισκευαστεί $(a = 1)$ ή να αντικατασταθεί $(a = 0)$ από ένα καινούργιο σύστημα. Υποθέτουμε ότι μία επισκευή και μία αντικατάσταση του συστήματος πραγματοποιούνται σε αμελητέο χρόνο.

Αν σε κάποια χρονική στιγμή το σύστημα βρεθεί σε μία από τις καταστάσεις (N,i), $0 \le i \le M$, και (n,M), $1 \le n \le N$, η αντικατάστασή του είναι υποχρεωτική και επαναφέρει το σύστημα σε μία από τις καταστάσεις $(1, j)$, όπου $0 \le j \le M$. Αν σε κάποια χρονική στιγμή το σύστημα βρεθεί σε μία από τις καταστάσεις (n,i), $1 \le n \le N-1$, $0 \le i \le M-1$, έχει υποστεί την $n-$οστή βλάβη και αποφασιστεί να επισκευαστεί, τότε μεταβαίνει σε μία από τις καταστάσεις $(n+1, j)$, όπου $i \le j \le M$. Αν σε κάποια χρονική στιγμή το σύστημα βρεθεί σε μία από τις καταστάσεις (n,i), $1 \le n \le N-1$, $0 \le i \le M-1$, έχει υποστεί την $n-$οστή βλάβη και αποφασιστεί να αντικατασταθεί, τότε μεταβαίνει σε μία από τις καταστάσεις $(1, j)$, όπου $0 \le j \le M$.

Υποθέτουμε ότι μία αντικατάσταση του συστήματος από ένα καινούργιο σύστημα επιφέρει ένα σταθερό κόστος ίσο με C_0. Αν σε κάποια χρονική στιγμή το σύστημα βρεθεί σε μία κατάσταση (n,i) και έχει υποστεί την $n-$οστή βλάβη τότε μία επισκευή του επιφέρει ένα κόστος ίσο με $C_1(n,i)$. Υποθέτουμε επίσης ότι η συνάρτηση του κόστους επισκευής $C_1(n,i)$ είναι αύξουσα ως προς n και ως προς i.

Έστω X_n, $2 \le n \le N-1$, ο χρόνος που μεσολαβεί μεταξύ της $(n-1)-$οστής και της $n-$οστής βλάβης του συστήματος. Έστω επίσης $f(x)$ η συνάρτηση πυκνότητας πιθανότητας και $F(x)$ η συνάρτηση κατανομής, αντίστοιχα, του χρόνου μέχρι το σύστημα να υποστεί την πρώτη βλάβη. Σύμφωνα με τις εργασίες των Kijima et al. (1988), Makis and Jardine (1993) και Love et al. (2000), αν το σύστημα βρεθεί στην κατάσταση (n,i), $1 \le n \le N-1$, $0 \le i \le M-1$, και επιλεγεί η ενέργεια $a = 1$, η ουσιαστική ηλικία (virtual age) του συστήματος μετά από την επισκευή του, θεωρείται ότι είναι ίση με $\theta i/\xi,$ όπου $0 \le \theta \le 1,$ και η συνάρτηση πυκνότητας πιθανότητας του X_{n+1} είναι:

$$f_{\frac{\theta i}{\xi}}(x) = \frac{f\left(x + \frac{\theta i}{\xi}\right)}{\overline{F}\left(\frac{\theta i}{\xi}\right)}, \quad x > 0,$$

όπου, $\overline{F}(x) = 1 - F(x), \quad x > 0.$ Σημειώνουμε ότι η παράμετρος θ αναπαριστά το βαθμό επισκευής. Αν $\theta = 1,$ η ουσιαστική ηλικία του συστήματος μετά από την $n-$οστή επισκευή του είναι ίση με την ηλικία του. Αυτό σημαίνει ότι στην περίπτωση αυτή πραγματοποιείται μία ελάχιστη επισκευή του συστήματος. Αν $\theta < 1,$ η ουσιαστική ηλικία του συστήματος είναι αυστηρά μικρότερη από την ηλικία $i.$ Αυτό σημαίνει ότι στην περίπτωση αυτή μία επισκευή ανανεώνει το σύστημα στο βαθμό που καθορίζεται από την τιμή του $\theta.$

Θεωρούμε μία ημι-Μαρκοβιανή διαδικασία αποφάσεων με χώρο καταστάσεων S και αναζητούμε τη στάσιμη πολιτική η οποία, για οποιαδήποτε αρχική κατάσταση του συστήματος, ελαχιστοποιεί το μακροπρόθεσμο αναμενόμενο μέσο κόστος ανά μονάδα χρόνου.

Έστω $p_{su}(a)$ η πιθανότητα η επόμενη κατάσταση της διαδικασίας να είναι η κατάσταση $u \in S,$ αν η παρούσα κατάσταση είναι η κατάσταση $s \in S$ και επιλέγεται η ενέργεια $a \in \{0, 1\}.$ Έστω επίσης $T(s,a)$ ο αναμενόμενος χρόνος για τη μετάβαση της διαδικασίας στην επόμενη κατάσταση αν η παρούσα κατάσταση είναι η κατάσταση $s \in S$ και επιλέγεται η ενέργεια $a \in \{0, 1\}.$ Οι πιθανότητες μετάβασης δίνονται από τις ακόλουθες εξισώσεις (5.1)-(5.4).

$$p_{(n,i)(1,j)}(0) = \int_{\frac{j}{\xi}}^{\frac{j+1}{\xi}} f(x)dx, \quad 1 \le n \le N, \quad 0 \le i \le M, \quad 0 \le j \le M-1, \tag{5.1}$$

$$p_{(n,i)(1,M)}(0) = \int_{\frac{M}{\xi}}^{\infty} f(x)dx, \quad 1 \le n \le N, \quad 0 \le i \le M, \tag{5.2}$$

$$p_{(n,i)(n+1,j)}(1)=\int_{\frac{j-i}{\xi}}^{\frac{j-i+1}{\xi}} f_{\frac{\theta i}{\xi}}(x)dx,\quad 1\le n\le N-1,\quad 0\le i\le j\le M-1, \tag{5.3}$$

$$p_{(n,i)(n+1,M)}(1)=\int_{\frac{M-i}{\xi}}^{\infty} f_{\frac{\theta i}{\xi}}(x)dx,\quad 1\le n\le N-1,\quad 0\le i\le M-1. \tag{5.4}$$

Παρατηρούμε ότι στις εξισώσεις (5.2) και (5.4) έχουμε ενοποιήσει όλες τις περιπτώσεις κατά τις οποίες $j\ge M$ στις καταστάσεις $(1,M)$ και $(n+1,M)$, αντίστοιχα, έτσι ώστε οι βλάβες που συμβαίνουν σε ηλικιακά τμήματα μεγαλύτερα του M, τίθενται στο M.

Οι αναμενόμενοι χρόνοι δίνονται από τις ακόλουθες εξισώσεις (5.5) και (5.6) για τις δύο δυνατές ενέργειες της αντικατάστασης και της επισκευής, αντίστοιχα.

$$T((n,i),0)=\int_0^{\infty} xf(x)dx,\quad 1\le n\le N,\quad 0\le i\le M, \tag{5.5}$$

$$T((n,i),1)=\int_0^{\infty} xf_{\frac{\theta i}{\xi}}(x)dx,\quad 1\le n\le N-1,\quad 0\le i\le M-1. \tag{5.6}$$

Επισημαίνουμε ότι, υπό τον έλεγχο οποιασδήποτε στάσιμης πολιτικής, το σύστημα είναι δυνατόν να μεταβεί στην κατάσταση $(1,0)$ από οποιαδήποτε αρχική κατάσταση $s\in S$. Επομένως, σύμφωνα με το Πόρισμα 6.20 στη σελ. 149 του βιβλίου του Ross (1992), υπάρχει μία βέλτιστη στάσιμη πολιτική διότι ο χώρος καταστάσεων S του συστήματος είναι πεπερασμένος. Οι Makis and Jardine (1993) απέδειξαν ότι, αν η συνάρτηση $f(x)$ και ο ρυθμός επιδείνωσης του συστήματος ικανοποιούν κάποιες κατάλληλες συνθήκες, η βέλτιστη πολιτική είναι μονότονη διότι αντικαθιστά το σύστημα όταν έχει υποστεί την $n-$οστή βλάβη, $n=1,\ldots,N-1$, αν και μόνο αν το αντίστοιχο ηλικιακό τμήμα i είναι μεγαλύτερο ή ίσο με μία κρίσιμη τιμή s_n^*. Απέδειξαν επίσης ότι: $s_1^*\ge s_2^*\ge\cdots\ge s_{N-1}^*$.

5.3 Ο αλγόριθμος

Η βέλτιστη πολιτική μπορεί να υπολογιστεί με εφαρμογή του αλγορίθμου βελτίωσης των πολιτικών, του αλγορίθμου των διαδοχικών προσεγγίσεων και της μεθόδου του γραμμικού προγραμματισμού. Για μία λεπτομερή περιγραφή των αλγορίθμων με διάφορες εφαρμογές τους παραπέμπουμε στο Κεφάλαιο 3 του βιβλίου του Tijms (1994). Είναι ωστόσο δυνατόν να σχεδιάσουμε έναν κατάλληλο αλγόριθμο βελτίωσης των πολιτικών ο οποίος εφαρμόζεται στο σύνολο των μονότονων πολιτικών και παράγει μία ακολουθία αυστηρά βελτιωμένων πολιτικών που έχουν τη μονότονη μορφή. Ο αλγόριθμος παράγει πολιτικές τέτοιες ώστε για κάθε τιμή του $n,\ 1 \le n \le N-1,$ και δοθέντος ενός βαθμού επισκευής $\theta,\ 0 \le \theta \le 1,$ αντικαθιστούν το σύστημα αν και μόνο αν το αντίστοιχο ηλικιακό τμήμα του είναι μεγαλύτερο ή ίσο με μία κρίσιμη τιμή s_n. Πολλά αριθμητικά παραδείγματα παρέχουν ισχυρή ένδειξη ότι η τελική πολιτική που δημιουργεί ο αλγόριθμος είναι η βέλτιστη πολιτική διότι συμπίπτει με την τελική πολιτική που δίνει ο αλγόριθμος βελτίωσης των πολιτικών και ο αλγόριθμος των διαδοχικών προσεγγίσεων. Ο σχεδιασμός του αλγορίθμου βασίζεται στην τεχνική εμφύτευσης του Tijms (βλέπε σελ. 234 του βιβλίου του Tijms (1994)).

Στη συνέχεια θα περιγράψουμε το σχεδιασμό του αλγορίθμου. Θεωρούμε μία μονότονη πολιτική R η οποία, για κάθε τιμή του αριθμού των βλαβών $n,\ 1 \le n \le N-1,$ και δοθέντος ενός βαθμού επισκευής $\theta,$ χαρακτηρίζεται από τις κρίσιμες τιμές $s_n,\ 1 \le n \le N-1.$ Αν η μονότονη πολιτική R υιοθετηθεί για τον έλεγχο του συστήματος, τότε το σύστημα είναι δυνατόν να μεταβεί σε κάποια κατάσταση του συνόλου των καταστάσεων:

$$E = \{(n,i), 1 \le n \le N-1, 0 \le i \le s_n - 1\},$$

από οποιαδήποτε αρχική κατάσταση $(n,i) \in S.$ Έστω $g(R)$ το μακροπρόθεσμο αναμενόμενο μέσο κόστος ανά μονάδα χρόνου υπό τον έλεγχο της πολιτικής $R,$ και, T_s^E, C_s^E ο αναμενόμενος χρόνος και το αναμενόμενο κόστος, αντίστοιχα, μέχρι την πρώτη είσοδο του συστήματος στο σύνολο $E,$ αν η αρχική κατάσταση είναι η κατάσταση $s \in S$ και η πολιτική R έχει υιοθετηθεί για τον έλεγχο του συστήματος. Έστω επίσης $p_{s\ell}^E$ η πιθανότητα η πρώτη κατάσταση κατά την είσοδο του συστήματος στο σύνολο E να είναι η κατάσταση ℓ, δοθέντος ότι η πολιτική R έχει υιοθετηθεί για τον έλεγχο του συστήματος και η αρχική κατάσταση είναι η κατάσταση $s \in S$. Υποθέτουμε ότι, αν η αρχική κατάσταση του συστήματος είναι μία κατάσταση $s \in E$, τότε η πρώτη κατάσταση κατά την είσοδο του συστήματος στο σύνολο E είναι η κατάσταση στην οποία μεταβαίνει το σύστημα στην επόμενη επιστροφή του στο σύνολο E.

Σύμφωνα με τη σχέση (3.6.1) στη σελ. 235 του βιβλίου του Tijms (1994), οι σχετικές τιμές $w(s)$, $s \in S$, της πολιτικής R, ικανοποιούν το ακόλουθο σύστημα των γραμμικών εξισώσεων:

$$w(s) = C_s^E - g(R)T_s^E + \sum_{\ell \in E} p_{s\ell}^E w(\ell), \quad s \in S, \tag{5.7}$$

$$w(1, 0) = 0. \tag{5.8}$$

Οι ποσότητες T_s^E, C_s^E και $p_{s\ell}^E$, $s \in S$, $\ell \in E$, μπορούν να υπολογιστούν αν λάβουμε υπόψη τις μεταβάσεις του συστήματος υπό τον έλεγχο της μονότονης πολιτικής R και χρησιμοποιήσουμε διάφορες μορφές του Θεωρήματος Ολικής Πιθανότητας. Οι εκφράσεις για τις ποσότητες αυτές δίνονται παρακάτω.

$$T_{(n,i)}^E = \int_0^\infty x f_{\frac{\theta i}{\xi}}(x)dx + \int_{\frac{s_{n+1}-i}{\xi}}^\infty f_{\frac{\theta i}{\xi}}(x)dx \frac{\int_0^\infty xf(x)dx}{\int_0^{\frac{s_1}{\xi}} f(x)dx}, \quad (n,i) \in E,$$

$$T_{(n,i)}^E = \frac{\int_0^\infty xf(x)dx}{\int_0^{\frac{s_1}{\xi}} f(x)dx}, \quad (n,i) \in S - E,$$

$$C^E_{(n,i)} = C_1(n,i) + C_0 \frac{\int_{\frac{s_{n+1}-i}{\xi}}^{\infty} f_{\frac{\theta i}{\xi}}(x)dx}{\int_0^{\frac{s_1}{\xi}} f(x)dx}, \quad (n,i) \in E,$$

$$C^E_{(n,i)} = \frac{C_0}{\int_0^{\frac{s_1}{\xi}} f(x)dx}, \quad (n,i) \in S - E,$$

$$p^E_{(n,i)(n+1,j)} = \int_{\frac{j-i}{\xi}}^{\frac{j-i+1}{\xi}} f_{\frac{\theta i}{\xi}}(x)dx, \ 0 \le i < s_n, \quad i \le j < s_{n+1}, \quad 1 \le n \le N-2,$$

$$p^E_{(n,i)(1,j)} = \int_{\frac{s_{n+1}-i}{\xi}}^{\infty} f_{\frac{\theta i}{\xi}}(x)dx \int_{\frac{j}{\xi}}^{\frac{j+1}{\xi}} f(x)dx \left(\int_0^{\frac{s_1}{\xi}} f(x)dx \right)^{-1}, \quad 0 \le i < s_n, \quad 0 \le j < s_1, \quad 1 \le n \le N-2, \quad (5.9)$$

$$p^E_{(n,i)(1,j)} = \int_{\frac{j}{\xi}}^{\frac{j+1}{\xi}} f(x)dx \left(\int_0^{\frac{s_1}{\xi}} f(x)dx \right)^{-1}, \quad s_n \le i \le M, \quad 0 \le j < s_1, \quad 1 \le n \le N-2,$$

$$p^E_{(N-1,i)(1,j)} = \int_{\frac{j}{\xi}}^{\frac{j+1}{\xi}} f(x)dx \left(\int_0^{\frac{s_1}{\xi}} f(x)dx \right)^{-1}, \quad 0 \le i \le M, \quad 0 \le j < s_1.$$

Θα εξηγήσουμε αναλυτικά τον τρόπο με τον οποίο εξάγεται η σχέση (5.9). Οι υπόλοιπες παραπάνω σχέσεις προκύπτουν με παρόμοιο τρόπο. Υποθέτουμε ότι η αρχική κατάσταση του συστήματος είναι η κατάσταση (n,i), $0 \le i \le s_n - 1$, $1 \le n \le N-2$, και η μονότονη πολιτική R

η οποία χαρακτηρίζεται από τις κρίσιμες τιμές $s_1, ..., s_{N-1}$ έχει υιοθετηθεί για τον έλεγχο του συστήματος. Για να είναι η κατάσταση $(1, j),\ 0 \le j \le s_1 - 1,$ η πρώτη κατάσταση στην επόμενη επιστροφή του συστήματος στο σύνολο $E,$ πρέπει η ηλικία του συστήματος όταν υποστεί την $(n+1)-$οστή βλάβη να είναι μεγαλύτερη από $s_{n+1}/\xi.$ Η πιθανότητα του ενδεχομένου αυτού είναι:

$$P_1 = \int_{\frac{s_{n+1}-i}{\xi}}^{\infty} f_{\frac{\theta i}{\xi}}(x)dx. \tag{5.10}$$

Επομένως,

$$p^E_{(n,i)(1,j)} = P_1 P_2, \tag{5.11}$$

όπου, P_2 είναι η πιθανότητα η πρώτη κατάσταση κατά την είσοδο του συστήματος στο σύνολο $E,$ υπό τον έλεγχο της πολιτικής $R,$ να είναι η κατάσταση $(1, j),$ δοθέντος ότι η αρχική κατάσταση του συστήματος είναι μία κατάσταση $(n+1, i'),\ i' > s_{n+1}.$

Υποθέτουμε ότι η αρχική κατάσταση είναι η κατάσταση $(n+1, i'),\ i' > s_{n+1}.$ Το σύστημα αντικαθίσταται και η ηλικία του όταν υποστεί την πρώτη βλάβη μπορεί να είναι μεγαλύτερη ή μικρότερη από $s_1/\xi.$ Αν η ηλικία του είναι μεγαλύτερη από $s_1/\xi,$ το σύστημα θα αντικατασταθεί ξανά. Η κατάσταση $(1, j)$ είναι η πρώτη κατάσταση κατά την είσοδο του συστήματος στο σύνολο E μετά από μία αντικατάστασή του με πιθανότητα ίση με $\int_{j/\xi}^{(j+1)/\xi} f(x)dx.$ Η κατάσταση $(1, j)$ είναι η πρώτη κατάσταση κατά την είσοδο του συστήματος στο σύνολο E μετά από δύο αντικαταστάσεις του με πιθανότητα ίση με $P_3 \int_{j/\xi}^{(j+1)/\xi} f(x)dx,$ όπου $P_3 = \int_{s_1/\xi}^{\infty} f(x)dx.$ Η κατάσταση $(1, j)$ είναι η πρώτη κατάσταση κατά την είσοδο του συστήματος στο σύνολο E

μετά από τρεις αντικαταστάσεις του με πιθανότητα ίση με $P_3^2 \int_{j/\xi}^{(j+1)/\xi} f(x)dx$ και ούτω κάθ' εξής. Επομένως, προσθέτοντας όλες τις παραπάνω πιθανότητες, έχουμε ότι:

$$P_2 = \int_{\frac{j}{\xi}}^{\frac{j+1}{\xi}} f(x)dx + P_3 \int_{\frac{j}{\xi}}^{\frac{j+1}{\xi}} f(x)dx + P_3^2 \int_{\frac{j}{\xi}}^{\frac{j+1}{\xi}} f(x)dx + \cdots = \frac{1}{1-P_3} \int_{\frac{j}{\xi}}^{\frac{j+1}{\xi}} f(x)dx. \tag{5.12}$$

Από τις σχέσεις (5.10)-(5.12) προκύπτει η έκφραση (5.9).

Οι σχετικές τιμές $w(s),\ s \in E,$ και το μακροπρόθεσμο αναμενόμενο μέσο κόστος $g(R)$ ανά μονάδα χρόνου, υπό τον έλεγχο της πολιτικής $R,$ μπορούν να υπολογιστούν επιλύνοντας το ακόλουθο σύστημα των γραμμικών εξισώσεων:

$$w(s) = C_s^E - g(R)T_s^E + \sum_{\ell \in E} p_{s\ell}^E w(\ell), \quad s \in E, \tag{5.13}$$

και την εξίσωση (5.8).

Θεωρούμε την ποσότητα $Q_R(s;a)$ η οποία ορίζεται ως εξής:

$$Q_R(s;a) = C(s,a) - g(R)T(s,a) + \sum_{\ell \in S} p_{s\ell}(a) w(\ell),\ s \in \{1, ..., N-1\} \times \{0, ..., M-1\}, \tag{5.14}$$

$$a \in \{0, 1\},$$

όπου, $p_{s\ell}(a)$ είναι η πιθανότητα η επόμενη κατάσταση του συστήματος να είναι η κατάσταση $\ell,$ αν η παρούσα κατάσταση είναι η κατάσταση s και επιλέγεται η ενέργεια $a,$ και $T(s,a),$ $C(s,a)$ είναι ο αναμενόμενος χρόνος και το αναμενόμενο κόστος, αντίστοιχα, για τη μετάβαση του συστήματος στην επόμενη κατάσταση, αν η παρούσα κατάσταση είναι η κατάσταση s και επιλέγεται η ενέργεια $a.$ Οι πιθανότητες μετάβασης $p_{s\ell}(a)$ δίνονται από τις εξισώσεις (5.1)-(5.4) και οι αναμενόμενοι χρόνοι $T(s,a)$ δίνονται από τις εξισώσεις (5.5) και (5.6). Αν σε μία κατάσταση $(n,i)\in S,$ επιλεγεί η ενέργεια $a=0,$ τότε $C((n,i),0)=C_0,$ και αν επιλεγεί η ενέργεια $a=1,$ τότε $C((n,i),1)=C_1(n,i).$

Υποθέτουμε ότι για κάποια τιμή του $n,\ 1\le n\le N-1,$ υπάρχει μία κρίσιμη τιμή $\tilde{s}_n$ τέτοια ώστε $0\le \tilde{s}_n < s_n$ και $Q_R((n,i);0)<w(n,i),\ \tilde{s}_n\le i<s_n.$ Τότε, σύμφωνα με το Θεώρημα 2.7 για τις ημι-Μαρκοβιανές διαδικασίες αποφάσεων (βλέπε (ε) στη σελ. 28 της παρούσας διατριβής), το μακροπρόθεσμο αναμενόμενο μέσο κόστος υπό τον έλεγχο της μονότονης πολιτικής η οποία χαρακτηρίζεται από τις κρίσιμες τιμές $s_1,\ldots,s_{n-1},\tilde{s}_n,s_{n+1},\ldots,s_{N-1}$ είναι μικρότερο από $g(R).$ Παρόμοια, αν για κάποια τιμή του $n,\ 1\le n\le N-1,$ υπάρχει μία κρίσιμη τιμή $\tilde{s}_n$ τέτοια ώστε $s_n<\tilde{s}_n\le M$ και $Q_R((n,i);1)<w(n,i),\ s_n\le i<\tilde{s}_n,$ τότε, σύμφωνα πάλι με το Θεώρημα 2.7 για τις ημι-Μαρκοβιανές διαδικασίες αποφάσεων, το μακροπρόθεσμο αναμενόμενο μέσο κόστος υπό τον έλεγχο της μονότονης πολιτικής η οποία χαρακτηρίζεται από τις κρίσιμες τιμές $s_1,\ldots,s_{n-1},\tilde{s}_n,s_{n+1},\ldots,s_{N-1}$ είναι μικρότερο από $g(R).$

Οι παραπάνω παρατηρήσεις μας οδηγούν στο σχεδιασμό του ακόλουθου αλγορίθμου ο οποίος παράγει μία ακολουθία αυστηρά βελτιωμένων μονότονων πολιτικών.

Αλγόριθμος

Βήμα 1. *(Έναρξη).* Επιλέγουμε ένα άνω φράγμα N στον αριθμό των επισκευών του συστήματος, ένα άνω φράγμα M για τον αριθμό των ηλικιακών τμημάτων και ένα βαθμό επισκευής $\theta\in[0,1].$ Επιλέγουμε επίσης μία αρχική μονότονη πολιτική R η οποία χαρακτηρίζεται από τις κρίσιμες τιμές $s_n,\ 1\le n\le N-1.$

Βήμα 2. *(Εύρεση σχετικών τιμών και μέσου κόστους).* Για τη μονότονη πολιτική $R,$ υπολογίζουμε το μακροπρόθεσμο αναμενόμενο μέσο κόστος $g(R)$ ανά μονάδα χρόνου και τις σχετικές τιμές $w(s),\ s\in E,$ επιλύνοντας το σύστημα των γραμμικών εξισώσεων (5.13) και (5.8).

Βήμα 3. *(Βελτίωση των πολιτικών).* Για κάθε $n=1,\ldots,N$:

(α) Βρίσκουμε, αν υπάρχει, το μικρότερο αριθμό $\tilde{s}_n$ τέτοιο ώστε $0\le\tilde{s}_n<s_n$ και $Q_R((n,i);0)<w(n,i),\ \tilde{s}_n\le i<s_n,$

διαφορετικά,

(β) βρίσκουμε, αν υπάρχει, το μεγαλύτερο αριθμό $\tilde{s}_n$ τέτοιο ώστε $s_n < \tilde{s}_n \le M$ και $Q_R((n,i);1) < w(n,i),\ s_n \le i < \tilde{s}_n$.

Οι ποσότητες $Q_R((n,i);1)$ και $Q_R((n,i);0)$ δίνονται από την εξίσωση (5.14). Οι σχετικές τιμές $w(s),\ s \in S - E,$ υπολογίζονται από την εξίσωση (5.7).

Αντικαθιστούμε την κρίσιμη τιμή s_n με την κρίσιμη τιμή $\tilde{s}_n$ για εκείνες τις τιμές του $n,$ $1 \le n \le N-1,$ για τις οποίες είναι δυνατόν να βρούμε μία κρίσιμη τιμή $\tilde{s}_n$ και επιστρέφουμε στο Βήμα 2.

Βήμα 4. *(Κριτήριο Σύγκλισης).* Αν δεν είναι δυνατόν να βρεθεί καμία κρίσιμη τιμή $\tilde{s}_n,$ $1 \le n \le N-1,$ ο αλγόριθμος σταματά. Η τελική πολιτική που παράγει ο αλγόριθμος είναι η πολιτική R με μακροπρόθεσμο αναμενόμενο μέσο κόστος $g(R)$ ανά μονάδα χρόνου.

Ο αλγόριθμος παράγει μία ακολουθία αυστηρά βελτιωμένων μονότονων πολιτικών και σταματά μετά από ένα πεπερασμένο αριθμό επαναλήψεων διότι το σύνολο των μονότονων πολιτικών είναι πεπερασμένο. Διαφέρει από τον αλγόριθμο των Love et al. (βλέπε Εδάφιο 3 στην εργασία των Love et al. (2000)) στο Βήμα 2 της εύρεσης των σχετικών τιμών και του μέσου κόστους. Πολλά αριθμητικά παραδείγματα παρέχουν ισχυρή ένδειξη ότι η τελική πολιτική που δημιουργεί ο αλγόριθμος είναι η βέλτιστη πολιτική διότι συμπίπτει με την τελική πολιτική που δίνει ο αλγόριθμος βελτίωσης των πολιτικών και ο αλγόριθμος των διαδοχικών προσεγγίσεων. Ο απαιτούμενος υπολογιστικός χρόνος για την εκτέλεση του αλγορίθμου είναι κατά πολύ μικρότερος από τον απαιτούμενο υπολογιστικό χρόνο για την εκτέλεση του αλγορίθμου βελτίωσης των πολιτικών και τον απαιτούμενο υπολογιστικό χρόνο για την εκτέλεση του αλγορίθμου των Love et al. (2000). Αυτό οφείλεται στο γεγονός ότι ο αριθμός των αγνώστων στο σύστημα των γραμμικών εξισώσεων του Βήματος 2 του αλγορίθμου είναι ίσος με τον αριθμό των στοιχείων του συνόλου $E,$ ενώ ο αριθμός των αγνώστων στο σύστημα των γραμμικών εξισώσεων του αντίστοιχου βήματος της εύρεσης των σχετικών τιμών και του μέσου κόστους στον αλγόριθμο βελτίωσης των πολιτικών και στον αλγόριθμο των Love et al. είναι ίσος με τον αριθμό των στοιχείων του χώρου καταστάσεων $S.$ Από πολλά αριθμητικά παραδείγματα που έχουμε εξετάσει συμπεραίνουμε επίσης ότι, ο αριθμός των επαναλήψεων του αλγορίθμου δεν επηρεάζεται ιδιαίτερα από την αρχική μονότονη πολιτική που επιλέγουμε κατά την έναρξη του αλγορίθμου στο Βήμα 1.

Οι Love et al. (2000) αναφέρουν ότι η τελική πολιτική του αλγορίθμου είναι η βέλτιστη πολιτική. Μία αυστηρή απόδειξη αυτού του ισχυρισμού φαίνεται δύσκολη, αν και πολλά αριθμητικά παραδείγματα παρέχουν ισχυρή ένδειξη ότι είναι αληθινός. Επισημαίνουμε ότι, σε διάφορα Μαρκοβιανά μοντέλα αποφάσεων (βλέπε π.χ. Federgruen and So (1989), Kyriakidis (1999b)) υπάρχει η δυνατότητα να αποδειχθεί ότι η τελική πολιτική που δημιουργούν παρόμοιοι κατάλληλοι αλγόριθμοι βελτίωσης των πολιτικών είναι η βέλτιστη πολιτική. Ο χώρος καταστάσεων αυτών των μοντέλων ήταν μονοδιάστατος και ήταν δυνατόν να αποδειχθεί ότι, υπό τον έλεγχο μιας μονότονης πολιτικής, το μακροπρόθεσμο αναμενόμενο μέσο κόστος ανά μονάδα χρόνου είναι μία μονοκόρυφη συνάρτηση ως προς την κρίσιμη τιμή. Μία συνέπεια αυτού του αποτελέσματος είναι το γεγονός ότι η τελική πολιτική του αλγορίθμου είναι η βέλτιστη πολιτική.

5.4 Αριθμητικά παραδείγματα

Θα παρουσιάσουμε δύο αριθμητικά παραδείγματα. Υποθέτουμε ότι ο χρόνος ζωής ενός νέου συστήματος είναι μία τυχαία μεταβλητή η οποία ακολουθεί την κατανομή Γάμα και την κατανομή Weibull, αντίστοιχα. Για τους τύπους των πιθανοτήτων μετάβασης (Εξισώσεις (5.1)-(5.4)) και των αναμενόμενων χρόνων (Εξισώσεις (5.5) και (5.6)) όπως διαμορφώνονται στα δύο παρακάτω παραδείγματα παραπέμπουμε στο Παράρτημα Α των Love et al. (2000).

Στο πρώτο παράδειγμα υποθέτουμε ότι ο χρόνος ζωής ενός νέου συστήματος είναι μία τυχαία μεταβλητή η οποία ακολουθεί την κατανομή Γάμα με παραμέτρους $\alpha > 0$ και $\lambda > 0$. Η συνάρτηση πυκνότητας πιθανότητας του χρόνου μέχρι την πρώτη βλάβη του συστήματος έχει τύπο:

$$f(x) = \frac{\lambda^{\alpha} x^{\alpha-1}}{\Gamma(\alpha)} \exp(-\lambda x), \quad x > 0,$$

όπου, $\Gamma(\cdot)$ είναι η συνάρτηση Γάμα. Θεωρούμε την περίπτωση κατά την οποία $N = 9$, $M = 100$, $\theta = 0.3$, $\xi = 10$, $\alpha = 3$, $\lambda = 3$. Θεωρούμε επίσης ότι $C_0 = 4$ και $C_1(n,i) = \sqrt{n+1}$, $1 \le n \le 8$, $0 \le i \le 99$. Ως αρχική μονότονη πολιτική στο Βήμα 1 του αλγορίθμου επιλέγουμε την πολιτική η οποία χαρακτηρίζεται από τις κρίσιμες τιμές $s_n = 100,\ 1 \le n \le 8.$ Στον Πίνακα 5.1 παρουσιάζονται οι διαδοχικές μονότονες πολιτικές που παράγει ο αλγόριθμος και τα αντίστοιχα μακροπρόθεσμα αναμενόμενα μέσα κόστη τους ανά μονάδα χρόνου. Τονίζουμε ότι για κάθε τιμή του $n,\ 1 \le n \le 8,$ μία μονότονη πολιτική η οποία χαρακτηρίζεται από τις κρίσιμες τιμές s_n αντικαθιστά το σύστημα αν και μόνο αν η ηλικία του είναι μεγαλύτερη ή ίση με s_n / ξ.

Πίνακας 5.1. Οι διαδοχικές μονότονες πολιτικές

Οι κρίσιμες τιμές $s_n,\ 1 \le n \le N-1,$ των πολιτικών	Μέσο κόστος
$s_1 = \cdots = s_8 = 100$	3.1903
$s_1 = 50,\ s_2 = 6,\ s_3 = 3,\ s_4 = \cdots = s_8 = 0$	3.0291
$s_1 = 50,\ s_2 = 26,\ s_3 = 17,\ s_4 = 11,\ s_5 = 8,$ $s_6 = 5,\ s_7 = 3,\ s_8 = 1$	2.9026
$s_1 = 50,\ s_2 = 23,\ s_3 = 14,\ s_4 = 10,\ s_5 = 6,$ $s_6 = 4,\ s_7 = 1,\ s_8 = 0$	2.8996

Ο απαιτούμενος υπολογιστικός χρόνος για την εκτέλεση του αλγορίθμου ήταν 16 δευτερόλεπτα ενώ ο απαιτούμενος υπολογιστικός χρόνος για την εκτέλεση του αλγορίθμου βελτίωσης των πολιτικών και του αλγορίθμου των Love et al., ήταν 50 και 44 δευτερόλεπτα, αντίστοιχα.

Στο δεύτερο παράδειγμα υποθέτουμε ότι ο χρόνος ζωής ενός νέου συστήματος είναι μία τυχαία μεταβλητή η οποία ακολουθεί την κατανομή Weibull με παραμέτρους $\alpha > 0$ και $\lambda > 0.$ Η συνάρτηση πυκνότητας πιθανότητας του χρόνου μέχρι την πρώτη βλάβη του συστήματος έχει τύπο:

$$f(x) = \alpha\lambda(\lambda x)^{\alpha-1} \exp[-(\lambda x)^{\alpha}], \quad x > 0.$$

Θεωρούμε την περίπτωση κατά την οποία $N=9$, $M=50$, $\theta=0.8$, $\xi=10$, $\alpha=5$, $\lambda=2$. Θεωρούμε επίσης ότι $C_0=6$ και $C_1(n,i)=i/10$, $1\le n\le 8$, $0\le i\le 49$. Ως αρχική μονότονη πολιτική στο Βήμα 1 του αλγορίθμου επιλέγουμε την πολιτική η οποία χαρακτηρίζεται από τις κρίσιμες τιμές $s_n=50$, $1\le n\le 8$. Η βέλτιστη πολιτική που δημιουργεί ο αλγόριθμος χαρακτηρίζεται από τις κρίσιμες τιμές $s_n^*=24$, $1\le n\le 8$, έτσι ώστε αντικαθιστά το σύστημα αν η ηλικία του είναι μεγαλύτερη ή ίση με 2.4. Οι κρίσιμες τιμές s_n^* είναι ίδιες για κάθε τιμή του n, $1\le n\le 8$, διότι το κόστος της επισκευής του συστήματος είναι σταθερό ως προς n. Το μακροπρόθεσμο αναμενόμενο μέσο κόστος ανά μονάδα χρόνου βρέθηκε ίσο με 2.08. Ο απαιτούμενος υπολογιστικός χρόνος για την εκτέλεση του αλγορίθμου ήταν 17 δευτερόλεπτα ενώ ο απαιτούμενος υπολογιστικός χρόνος για την εκτέλεση του αλγορίθμου βελτίωσης των πολιτικών και του αλγορίθμου των Love et al., ήταν 48 και 42 δευτερόλεπτα, αντίστοιχα.

ΚΕΦΑΛΑΙΟ 6

Υπολογισμός της βέλτιστης πολιτικής για τον έλεγχο μιας σύνθετης διαδικασίας μετανάστευσης με την εισαγωγή ολοκληρωτικών καταστροφών

6.1 Εισαγωγή

Θεωρούμε έναν πληθυσμό ατόμων ο οποίος αναπτύσσεται στοχαστικά σύμφωνα με μία σύνθετη διαδικασία Poisson με ρυθμό ίσο με $\lambda > 0.$ Η συνάρτηση πιθανότητας του πληθυσμιακού μεγέθους των ομαδικών αυξήσεων είναι ίση με $\{g_j,\ j = 1, 2, ...\}$ και είναι τέτοια ώστε $g_1 > 0.$ Υποθέτουμε ότι τα άτομα είναι βλαβερά. Για παράδειγμα, μπορεί να είναι έντομα που καταστρέφουν μία σοδειά ή μεταδίδουν μία ασθένεια. Αποδίδουμε στα άτομα αυτά τον όρο παράσιτα. Η βλάβη που προξενούν τα παράσιτα επιφέρει ένα κόστος ίσο με $c_i,\ i \geq 0,$ για κάθε μονάδα χρόνου κατά την οποία το μέγεθος του πληθυσμού των παρασίτων είναι ίσο με $i.$ Υποθέτουμε ότι η ακολουθία $\{c_i\}$ είναι αύξουσα ως προς i και ότι $c_i \leq Ai^m$ για μία σταθερά $A > 0$ και έναν ακέραιο αριθμό $m.$ Εισάγουμε επίσης τις δύο ακόλουθες συνθήκες που αφορούν τη συνάρτηση πιθανότητας $\{g_j,\ j = 1, 2, ...\}.$

Συνθήκη 1: Η πιθανογεννήτρια συνάρτηση $G(z) = \sum_{j=1}^{\infty} g_j z^j$ της συνάρτησης πιθανότητας $\{g_j,\ j = 1, 2, ...\}$ έχει ακτίνα σύγκλισης αυστηρά μεγαλύτερη από τη μονάδα.

Συνθήκη 2: Όλες οι ροπές της συνάρτησης πιθανότητας $\{g_j,\ j = 1, 2, ...\}$ μέχρι και τάξεως m είναι πεπερασμένες.

Υποθέτουμε ότι υπάρχει ένας ελεγκτής ο οποίος παρατηρεί συνεχώς την εξέλιξη του πληθυσμού των παρασίτων και μπορεί να επιλέξει μία ενέργεια η οποία θέτει σε λειτουργία ένα μηχανισμό που καταστρέφει ολοκληρωτικά τον πληθυσμό. Ο ψεκασμός της σοδειάς ή των εντόμων με κάποιο εντομοκτόνο μπορεί να είναι μία κατάλληλη ενέργεια για τον έλεγχο της ανάπτυξης του πληθυσμού των παρασίτων στην παραπάνω περίπτωση. Υποθέτουμε ότι ο ρυθμός με τον οποίον ο πληθυσμός των παρασίτων καταστρέφεται ολοκληρωτικά είναι ίσος με $\mu > 0.$ Συνεπώς, όταν επιλέγεται η ενέργεια που θέτει σε λειτουργία το μηχανισμό της καταστροφής, ο χρόνος μέχρι την πραγματοποίηση της καταστροφής είναι μία τυχαία μεταβλητή η οποία ακολουθεί την Εκθετική κατανομή με μέση τιμή μ^{-1}. Υποθέτουμε επίσης ότι η επιλογή της ενέργειας η οποία θέτει σε λειτουργία το μηχανισμό της καταστροφής επιφέρει ένα κόστος ίσο με k ανά μονάδα χρόνου, όπου k είναι μία θετική σταθερά.

Μία στάσιμη πολιτική f ορίζεται από μία ακολουθία $\{f_i\},\ i \geq 0,$ όπου f_i είναι η ενέργεια που επιλέγεται όταν η διαδικασία βρίσκεται στην κατάσταση i. Υποθέτουμε ότι $f_i = 1,$ όταν επιλέγεται η ενέργεια η οποία εισάγει μία ολοκληρωτική καταστροφή του πληθυσμού των παρασίτων και $f_i = 0,$ όταν δεν επιλέγεται αυτή η ενέργεια. Αν η στάσιμη πολιτική $f \equiv \{f_i\},$ $i \geq 0,$ υιοθετηθεί για τον έλεγχο της διαδικασίας, η ανάπτυξη του πληθυσμού των παρασίτων μπορεί να περιγραφεί από ένα Μαρκοβιανό μοντέλο σε συνεχή χρόνο με χώρο καταστάσεων $S = \{0, 1, ...\}$ και με τις ακόλουθες μεταβάσεις σε ένα χρονικό διάστημα $(t, t + \delta t),$ καθώς $\delta t \to 0:$

$$\begin{aligned} &i \to i + j \quad \text{με πιθανότητα} \quad \lambda g_j \delta t + o(\delta t), \quad i \geq 0, \quad j \geq 1, \\ &i \to 0 \quad \text{με πιθανότητα} \quad \mu f_i \delta t + o(\delta t), \quad i \geq 1. \end{aligned}$$

Η συνάρτηση $o(\cdot)$ είναι τέτοια ώστε $o(h)/h \to 0,$ καθώς $h \to 0.$ Το πρόβλημα είναι η εύρεση της πολιτικής, η οποία, ανάμεσα σε όλες τις πολιτικές, ελαχιστοποιεί το μακροπρόθεσμο αναμενόμενο μέσο κόστος ανά μονάδα χρόνου. Οι χρονικές στιγμές που επιλέγεται μία ενέργεια είναι οι χρονικές στιγμές κατά τις οποίες συμβαίνει μία άφιξη μιας ομάδας παρασίτων. Φαίνεται διαισθητικά λογικό ότι η βέλτιστη πολιτική ανήκει στην κατηγορία των μονότονων πολιτικών $\{P_n, n = 1, 2, ...\},$ όπου P_n είναι η στάσιμη πολιτική, υπό τον έλεγχο της οποίας, επιλέγεται η ενέργεια που θέτει σε λειτουργία το μηχανισμό της καταστροφής αν και μόνο αν το μέγεθος του πληθυσμού των παρασίτων είναι μεγαλύτερο ή ίσο με n.

Το γεγονός ότι μία βέλτιστη πολιτική είναι μονότονη μπορεί να επιταχύνει σημαντικά τον υπολογισμό της διότι σε πολλά προβλήματα είναι δυνατόν να σχεδιαστούν αποδοτικοί αλγόριθμοι οι οποίοι εφαρμόζονται στο σύνολο των μονότονων πολιτικών και συγκλίνουν στη βέλτιστη πολιτική. Αναφέρουμε τις εργασίες των Abakuks (1979), Federgruen and So (1989), Benyamini and Yechiali (1999) και Love et al. (2000) στις οποίες κατάλληλοι αλγόριθμοι έχουν σχεδιαστεί για τον υπολογισμό μιας βέλτιστης μονότονης πολιτικής.

Μία συνήθης μέθοδος η οποία συχνά μας οδηγεί να αποδείξουμε ότι μία μονότονη πολιτική είναι βέλτιστη είναι η μέθοδος των διαδοχικών προσεγγίσεων (βλέπε π.χ. Ross (1992)). Αρχικά θεωρούμε το πρόβλημα της ελαχιστοποίησης του συνολικού αναμενόμενου αποπληθωρισμένου κόστους σε πεπερασμένο χρονικό ορίζοντα και αποδεικνύουμε ότι μία μονότονη πολιτική είναι βέλτιστη. Στη συνέχεια η μονότονη μορφή της βέλτιστης πολιτικής μετατίθεται στο αντίστοιχο πρόβλημα του άπειρου χρονικού ορίζοντα. Τελικά αποδεικνύουμε τη μονότονη μορφή της βέλτιστης πολιτικής για το πρόβλημα της ελαχιστοποίησης του μακροπρόθεσμου αναμενόμενου μέσου κόστους ανά μονάδα χρόνου.

Ο Economou (2003) θεώρησε μία κατάλληλη Μαρκοβιανή διαδικασία αποφάσεων σε συνεχή χρόνο και μελέτησε το παραπάνω πρόβλημα στην περίπτωση κατά την οποία η ακολουθία $\{c_i\}$, $i \geq 0$, είναι άνω φραγμένη. Χρησιμοποίησε την τεχνική της ομοιομορφοποίησης (βλέπε π.χ. σελ. 30-31 της παρούσας διατριβής και Serfozo (1979)) και μετέτρεψε την αρχική Μαρκοβιανή διαδικασία αποφάσεων σε συνεχή χρόνο σε μία ισοδύναμη Μαρκοβιανή διαδικασία αποφάσεων σε διακριτό χρόνο. Στη συνέχεια χρησιμοποίησε τη μέθοδο των διαδοχικών προσεγγίσεων και απέδειξε ότι η βέλτιστη πολιτική είναι μονότονη ή είναι τέτοια ώστε, ο μηχανισμός της καταστροφής του πληθυσμού δεν τίθεται ποτέ σε λειτουργία. Στο τελευταίο εδάφιο της εργασίας του, πρότεινε έναν αλγόριθμο ο οποίος υπολογίζει, κατά προσέγγιση, τις στάσιμες πιθανότητες της διαδικασίας, υπό τον έλεγχο της μονότονης πολιτικής P_n. Ένα μειονέκτημα της χρήσης της τεχνικής της ομοιομορφοποίησης είναι ότι, η απόδειξη του παραπάνω αποτελέσματος ισχύει μόνο στην περίπτωση κατά την οποία το σύνολο των επιτρεπόμενων πολιτικών που μπορούν να υιοθετηθούν για τον έλεγχο της διαδικασίας αποτελείται αποκλειστικά από τις στάσιμες πολιτικές.

Στο επόμενο εδάφιο αποδεικνύουμε ότι η βέλτιστη πολιτική είναι μονότονη χωρίς να χρησιμοποιήσουμε την τεχνική της ομοιομορφοποίησης. Το ίδιο αποτέλεσμα προκύπτει στην περίπτωση κατά την οποία η ακολουθία $\{c_i\}$, $i \geq 0$, δεν είναι φραγμένη. Η μονότονη πολιτική είναι επίσης βέλτιστη ακόμη και αν εισάγουμε ένα πλασματικό κόστος ίσο με r για κάθε μονάδα χρόνου κατά την οποία η διαδικασία βρίσκεται στην κατάσταση 0. Στο Εδάφιο 6.3

αποδεικνύουμε ότι, για κάθε $n=1,2,...$ η πολιτική P_n είναι βέλτιστη όταν η πλασματική παράμετρος r ανήκει σε ένα κατάλληλο διάστημα των πραγματικών αριθμών. Μία συνέπεια αυτού του αποτελέσματος είναι ότι το μακροπρόθεσμο αναμενόμενο μέσο κόστος ανά μονάδα χρόνου, υπό τον έλεγχο της πολιτικής P_n, είναι μία μονοκόρυφη συνάρτηση του $n \geq 1$. Αυτό μας επιτρέπει να σχεδιάσουμε δύο αποδοτικούς αλγόριθμους για τον υπολογισμό της βέλτιστης πολιτικής όπως η μέθοδος της διχοτόμησης και ένας κατάλληλος αλγόριθμος βελτίωσης των πολιτικών ο οποίος εφαρμόζεται στο σύνολο των μονότονων πολιτικών και συγκλίνει στη βέλτιστη πολιτική. Στο Εδάφιο 6.4 παρουσιάζουμε το σχεδιασμό των αλγορίθμων και δύο αριθμητικά παραδείγματα. Επιπλέον προτείνουμε μία μέθοδο για τον ακριβή υπολογισμό των στάσιμων πιθανοτήτων υπό τον έλεγχο της πολιτικής P_n. Στο Εδάφιο 6.5 θεωρούμε ένα γενικότερο μοντέλο στο οποίο είναι δυνατόν να επιλεγεί μία ενέργεια που θέτει σε λειτουργία ένα μηχανισμό που προξενεί μία διωνυμική καταστροφή αντί μιας ολοκληρωτικής καταστροφής του πληθυσμού των παρασίτων.

Τα αποτελέσματα των Εδαφίων 6.2-6.5 έχουν δημοσιευτεί στην εργασία των Kyriakidis and Dimitrakos (2005b).

6.2 Η μορφή της βέλτιστης πολιτικής

Στο παρόν εδάφιο και στο Εδάφιο 6.3 εισάγουμε ένα πλασματικό κόστος ίσο με r για κάθε μονάδα χρόνου κατά την οποία η διαδικασία βρίσκεται στην κατάσταση 0. Υποθέτουμε ότι η παράμετρος r μπορεί να είναι οποιοσδήποτε πραγματικός αριθμός. Το κύριο αποτέλεσμα του παρόντος κεφαλαίου είναι η Πρόταση 6.4. Είναι ένα πόρισμα του Θεωρήματος 6.2 το οποίο βασίζεται στις μεταβολές των τιμών της παραμέτρου r στην ευθεία των πραγματικών αριθμών.

Παρακάτω αποδεικνύουμε ότι η βέλτιστη πολιτική η οποία ελαχιστοποιεί το μακροπρόθεσμο αναμενόμενο μέσο κόστος ανά μονάδα χρόνου είναι μονότονη. Στην Περίπτωση 1 θεωρούμε ότι η ακολουθία $\{c_i\}$, $i \geq 0$, είναι άνω φραγμένη και στην Περίπτωση 2 θεωρούμε ότι η ακολουθία $\{c_i\}$ τείνει στο άπειρο, καθώς $i \to \infty$.

Περίπτωση 1. Υποθέτουμε ότι η ακολουθία $\{c_i\}$, $i \geq 0$, είναι άνω φραγμένη. Χρησιμοποιούμε το συνήθη τρόπο απόδειξης ο οποίος βασίζεται στη μέθοδο των διαδοχικών προσεγγίσεων (βλέπε π.χ. Kyriakidis (1999a), Van der Duyn Schouten and Vanneste (1995)). Αρχικά θεωρούμε το πρόβλημα της ελαχιστοποίησης του συνολικού αναμενόμενου αποπληθωρισμένου κόστους σε πεπερασμένο χρονικό ορίζοντα. Η σχετική θεωρία βρίσκεται στο Εδάφιο 2.5 της παρούσας διατριβής, στο Κεφάλαιο 7 του βιβλίου του Ross (1992) και στο Κεφάλαιο 5 του βιβλίου των Heyman and Sobel (2004).

Έστω $C_\alpha(i,a)$ το αναμενόμενο αποπληθωρισμένο κόστος μέχρι τη μετάβαση στην επόμενη κατάσταση, αν η κατάσταση της διαδικασίας είναι η κατάσταση $i \geq 1$, ο αποπληθωριστικός παράγοντας είναι ίσος με $\alpha > 0$ και επιλέγεται η ενέργεια $a \in \{0, 1\}$. Δεσμευόμαστε ως προς το χρόνο για τη μετάβαση της διαδικασίας στην επόμενη κατάσταση και έχουμε ότι:

$$C_\alpha(i,0) = \int_0^\infty \left[\int_0^t c_i e^{-\alpha s} ds \right] \lambda e^{-\lambda t} dt = \frac{c_i}{\alpha + \lambda}, \quad i \geq 1,$$

$$C_\alpha(i,1) = \int_0^\infty \left[\int_0^t (c_i + k) e^{-\alpha s} ds \right] (\lambda + \mu) e^{-(\lambda+\mu)t} dt = \frac{c_i + k}{\alpha + \lambda + \mu}, \quad i \geq 1.$$

Για κάθε κατάσταση $i \in \{1, 2, \dots\}$ και κάθε ενέργεια $a \in \{0, 1\}$, προκύπτει ότι:

$$C_\alpha(i,a) < \frac{\sup_{i \geq 1} c_i + k}{\lambda} \equiv B. \tag{6.1}$$

Το ελάχιστο αναμενόμενο αποπληθωρισμένο κόστος $V_\alpha(i,N)$, αν η αρχική κατάσταση είναι η κατάσταση $i \geq 1$, και απομένουν N βήματα μέχρι το τερματισμό της διαδικασίας, ικανοποιεί για κάθε $N = 1, 2, \dots$ τις ακόλουθες εξισώσεις (βλέπε (α) στη σελ. 27 της παρούσας διατριβής):

$$V_\alpha(i,N) = \min \left\{ \frac{c_i + \lambda \sum_{j=1}^{\infty} g_j V_\alpha(i+j, N-1)}{\alpha + \lambda}, \frac{c_i + k + \lambda \sum_{j=1}^{\infty} g_j V_\alpha(i+j, N-1) + \mu V_\alpha(0, N-1)}{\alpha + \lambda + \mu} \right\},$$

$$i \geq 1,$$

με την ακόλουθη αρχική συνθήκη:

$$V_\alpha(i,0)=0, \quad i\geq 1.$$

Με επαγωγή ως προς N μπορεί να αποδειχθεί ότι η συνάρτηση $\{V_\alpha(i,N)\}$ είναι αύξουσα ως προς $i\geq 1,$ για κάθε $N=0,1,\ldots.$ Έστω $V_\alpha(i)$ το ελάχιστο συνολικό αναμενόμενο αποπληθωρισμένο κόστος σε άπειρο χρονικό ορίζοντα, αν η αρχική κατάσταση της διαδικασίας είναι η κατάσταση $i\geq 1.$ Από το Θεώρημα 2.3 στη σελ. 20 της παρούσας διατριβής έχουμε ότι: $\lim\limits_{N\to\infty} V_\alpha(i,N)=V_\alpha(i),\ i\geq 1.$ Συνεπώς η συνάρτηση $\{V_\alpha(i)\},\ i\geq 1,$ είναι επίσης αύξουσα ως προς $i.$

Υποθέτουμε ότι η αρχική κατάσταση της διαδικασίας είναι η κατάσταση $i\geq 1.$ Χρησιμοποιούμε τα ίδια επιχειρήματα με αυτά του Εδαφίου 2 της εργασίας του Kyriakidis (1999a). Έστω $X_u,\ u=0,1,\ldots$ η κατάσταση της διαδικασίας τη $u-$οστή χρονική στιγμή επιλογής ενέργειας και T_u ο χρόνος μέχρι τη $u-$οστή χρονική στιγμή επιλογής ενέργειας. Υποθέτουμε ότι $T_0=0.$ Έστω π^* η βέλτιστη πολιτική που ελαχιστοποιεί το συνολικό αναμενόμενο αποπληθωρισμένο κόστος σε άπειρο χρονικό ορίζοντα. Έστω επίσης π η μη-στάσιμη πολιτική η οποία επιλέγει την ενέργεια που θέτει σε λειτουργία το μηχανισμό της καταστροφής μέχρι η διαδικασία να εισέλθει για πρώτη φορά στην κατάσταση 0 και στη συνέχεια επιλέγει τις ίδιες ενέργειες με εκείνες της βέλτιστης πολιτικής $\pi^*.$ Έστω ότι η διαδικασία εισέρχεται για πρώτη φορά στην κατάσταση 0 τη χρονική στιγμή επιλογής ενέργειας $N.$ Αναπαριστούμε με $V_\alpha(i,\pi^*)$ και με $V_\alpha(i,\pi)$ το συνολικό αναμενόμενο αποπληθωρισμένο κόστος σε άπειρο χρονικό ορίζοντα, υπό τον έλεγχο των πολιτικών π^* και $\pi,$ αντίστοιχα, αν η αρχική κατάσταση της διαδικασίας είναι η κατάσταση $i.$ Ισχύει ότι:

$$V_\alpha(i)=V_\alpha(i,\pi^*)\leq V_\alpha(i,\pi)=E_\pi\left[\sum_{u=0}^{\infty} e^{-\alpha T_u} C_\alpha(X_u,a_u)\,\Big|\, X_0=i\right],$$

όπου, E_π αναπαριστά την υπό-συνθήκη αναμενόμενη τιμή, αν η πολιτική π υιοθετηθεί για τον έλεγχο της διαδικασίας. Έχουμε ότι:

$$
\begin{aligned}
V_\alpha(i) &\le E_\pi\left[\sum_{u=0}^{N-1} e^{-\alpha T_u} C_\alpha(X_u, a_u) \,\Big|\, X_0 = i\right] + E_\pi\left[\sum_{u=N}^{\infty} e^{-\alpha T_u} C_\alpha(X_u, a_u) \,\Big|\, X_0 = i\right] \\
&= E_\pi\left[\sum_{u=0}^{N-1} e^{-\alpha T_u} C_\alpha(X_u, a_u) \,\Big|\, X_0 = i\right] + E_\pi(e^{-\alpha T_N}) E_\pi\left[\sum_{u=N}^{\infty} e^{-\alpha (T_u - T_N)} C_\alpha(X_u, a_u) \,\Big|\, X_0 = i\right] \\
&\le B E_\pi\left(\sum_{u=0}^{N-1} e^{-\alpha T_u}\right) + E_\pi(e^{-\alpha T_N}) V_\alpha(0) \\
&\le B E_\pi(N) + V_\alpha(0). \qquad (6.2)
\end{aligned}
$$

Η παραπάνω ισότητα οφείλεται στο γεγονός ότι η πολιτική π δεν εξαρτάται από το χρόνο T_N μετά από την είσοδο της διαδικασίας στην κατάσταση 0. Η προτελευταία παραπάνω ανισότητα είναι συνέπεια της σχέσης (6.1) και των ορισμών της πολιτικής π και της χρονικής στιγμής N. Σύμφωνα με την εξίσωση του Wald (βλέπε π.χ. σελ. 38 του βιβλίου του Ross (1992)), η υπό-συνθήκη αναμενόμενη τιμή $E_\pi(N)$ είναι ίση με τον αναμενόμενο χρόνο μέχρι την πραγματοποίηση μιας καταστροφής του πληθυσμού των παρασίτων διαιρούμενο με τον αναμενόμενο χρόνο κάθε μετάβασης της διαδικασίας. Επομένως, ισχύει ότι:

$$E_\pi(N) = \frac{\mu^{-1}}{(\lambda+\mu)^{-1}} = \frac{\lambda}{\mu} + 1. \qquad (6.3)$$

Από τις σχέσεις (6.2) και (6.3), έχουμε ότι:

$$V_\alpha(i) \le B\left(\frac{\lambda}{\mu} + 1\right) + V_\alpha(0), \quad i \ge 1. \qquad (6.4)$$

Για κάθε κατάσταση $i = 1, 2, \ldots$ έχουμε διαδοχικά ότι:

$$
\begin{aligned}
|V_\alpha(i) - V_\alpha(0)| &\le |V_\alpha(i) - V_\alpha(1)| + |V_\alpha(1) - V_\alpha(0)| \\
&= V_\alpha(i) - V_\alpha(1) + |V_\alpha(1) - V_\alpha(0)|
\end{aligned}
$$

$$= V_\alpha(i) - V_\alpha(0) + V_\alpha(0) - V_\alpha(1) + |V_\alpha(1) - V_\alpha(0)|$$

$$\leq B\left(\frac{\lambda}{\mu} + 1\right) + V_\alpha(0) - V_\alpha(1) + |V_\alpha(1) - V_\alpha(0)|.$$

Η πρώτη παραπάνω ισότητα προκύπτει από τη μονοτονία της ακολουθίας $\{V_\alpha(i)\}$, $i \geq 1$, και η δεύτερη παραπάνω ανισότητα είναι συνέπεια της σχέσης (6.4). Επομένως η ανισότητα (2.4) του Θεωρήματος 2.5 στη σελ. 22 της παρούσας διατριβής ισχύει διότι η ποσότητα $|V_\alpha(i) - V_\alpha(0)|$ είναι ομοιόμορφα φραγμένη ως προς $i \geq 0$ και ως προς $\alpha > 0$. Συνεπώς υπάρχει μία φραγμένη ακολουθία αριθμών $\{h_i\}$, $i = 0, 1, \ldots$ και μία σταθερά G έτσι ώστε:

$$h_i = \lim_{n\to\infty}[V_{\alpha_n}(i) - V_{\alpha_n}(0)], \quad i \geq 0, \tag{6.5}$$

για μία ακολουθία αριθμών $\{\alpha_n\}$ τέτοια ώστε $\lim_{n\to\infty} \alpha_n = 0$, και

$$h_i = \min\left\{\frac{c_i - G}{\lambda} + \sum_{j=1}^{\infty} g_j h_{i+j}, \frac{c_i + k - G}{\lambda + \mu} + \sum_{j=1}^{\infty} \frac{\lambda g_j}{\lambda + \mu} h_{i+j} + \frac{\mu}{\lambda + \mu} h_0\right\}, \quad i \geq 1. \tag{6.6}$$

Επιπλέον υπάρχει μία βέλτιστη στάσιμη πολιτική η οποία ελαχιστοποιεί το μακροπρόθεσμο αναμενόμενο μέσο κόστος ανά μονάδα χρόνου και σε κάθε κατάσταση $i \geq 1$ επιλέγει εκείνη την ενέργεια η οποία ελαχιστοποιεί το δεξιό μέλος της εξίσωσης (6.6). Σε μία κατάσταση $i \geq 1$, η βέλτιστη πολιτική επιλέγει την ενέργεια που θέτει σε λειτουργία το μηχανισμό της καταστροφής του πληθυσμού αν και μόνο αν ισχύει ότι:

$$\frac{\lambda}{\mu} k + \lambda h_0 + G \leq c_i + \lambda \sum_{j=1}^{\infty} g_j h_{i+j}. \tag{6.7}$$

Από τη σχέση (6.5) προκύπτει ότι η ακολουθία $\{h_i\}$, $i \geq 1$, είναι αύξουσα ως προς i, διότι η ακολουθία $\{V_\alpha(i)\}$, $i \geq 1$, είναι αύξουσα ως προς i. Από τις μονοτονίες των ακολουθιών $\{c_i\}$ και $\{h_i\}$, $i \geq 1$, καταλήγουμε στο ακόλουθο αποτέλεσμα.

Πρόταση 6.1. Αν η ακολουθία $\{c_i\}$, $i \geq 0$, είναι άνω φραγμένη, η βέλτιστη πολιτική είτε δεν επιλέγει ποτέ την ενέργεια που θέτει σε λειτουργία το μηχανισμό της καταστροφής είτε είναι η μονότονη πολιτική P_{x*}, όπου $x*$ είναι ο μικρότερος ακέραιος αριθμός $i \geq 1$ για τον όποιον ισχύει η ανισότητα (6.7).

Επισημαίνουμε ότι το παραπάνω αποτέλεσμα γενικεύει το Θεώρημα 2.3 του Economou (2003) το οποίο ισχύει μόνο στην περίπτωση κατά την οποία οι επιτρεπόμενες πολιτικές για τον έλεγχο της διαδικασίας είναι αποκλειστικά οι στάσιμες πολιτικές.

Περίπτωση 2. Υποθέτουμε ότι $c_i \to \infty$, καθώς $i \to \infty$. Σύμφωνα με το Θεώρημα 3.1 του Economou (2003), υπό τον έλεγχο της πολιτικής P_n, $n \geq 1$, οι στάσιμες πιθανότητες της διαδικασίας ακολουθούν ασυμπτωτικά (δηλαδή για μεγάλες τιμές του πληθυσμιακού μεγέθους των παρασίτων) τη Γεωμετρική κατανομή διότι η πιθανογεννήτρια συνάρτηση της $\{g_j, j = 1, 2, ...\}$ έχει ακτίνα σύγκλισης αυστηρά μεγαλύτερη από τη μονάδα (βλέπε Συνθήκη 1 στη σελ. 123 της παρούσας διατριβής). Το αποτέλεσμα αυτό και η ανισότητα $c_i \leq Ai^m$, όπου $A > 0$ είναι μία σταθερά και m είναι ένας ακέραιος αριθμός, εγγυώνται ότι, υπό τον έλεγχο οποιασδήποτε μονότονης πολιτικής, το μακροπρόθεσμο αναμενόμενο μέσο κόστος ανά μονάδα χρόνου είναι πεπερασμένο. Παρατηρούμε επίσης ότι, υπό τον έλεγχο της πολιτικής η οποία δεν επιλέγει ποτέ την ενέργεια που θέτει σε λειτουργία το μηχανισμό της καταστροφής, η διαδικασία δεν έχει μία στάσιμη κατανομή. Επομένως, στην περίπτωση αυτή το μακροπρόθεσμο αναμενόμενο μέσο κόστος ανά μονάδα χρόνου είναι άπειρο για κάθε αρχική κατάσταση της διαδικασίας διότι η ακολουθία $\{c_i\}$ τείνει στο άπειρο (βλέπε Λήμμα 2.2.2 του Bather (1976)). Επιπλέον, όπως στην Περίπτωση 1, είναι δυνατόν να αποδειχθεί ότι το ελάχιστο συνολικό αναμενόμενο αποπληθωρισμένο κόστος $V_\alpha(i)$, $i \geq 1$, σε άπειρο χρονικό ορίζοντα, είναι μία αύξουσα συνάρτηση ως προς i.

Η Sennott (1989) κατασκεύασε πέντε υποθέσεις που εγγυώνται την ύπαρξη μιας βέλτιστης στάσιμης πολιτικής η οποία ελαχιστοποιεί το μακροπρόθεσμο αναμενόμενο μέσο κόστος ανά μονάδα χρόνου σε μία ημι-Μαρκοβιανή διαδικασία αποφάσεων. Οι ρυθμοί μετάβασης και τα κόστη του παρόντος μοντέλου ικανοποιούν αυτές τις υποθέσεις. Για να αποδείξουμε ότι οι υποθέσεις της Sennott ισχύουν στο παρόν μοντέλο αποδεικνύουμε ισοδύναμα τις Υποθέσεις 1, 2 και τη Συνθήκη SEN που παρουσιάζονται στις σελ. 28-30 της παρούσας διατριβής.

Η Υπόθεση 1 ισχύει για $\delta = 1$ και $\varepsilon = \exp[-(\lambda + \mu)]$. Η Υπόθεση 2 ισχύει για $B = (\lambda + \mu)^{-1}$. Η Συνθήκη SEN ισχύει αν επιλέξουμε την πολιτική $f = P_1$. Η Ιδιότητα 1 της Συνθήκης SEN ισχύει, διότι, υπό τον έλεγχο της πολιτικής P_1, οι στάσιμες πιθανότητες ακολουθούν ασυμπτωτικά τη Γεωμετρική κατανομή. Η Ιδιότητα 2 της Συνθήκης SEN ισχύει αν επιλέξουμε οποιοδήποτε $\varepsilon > 0$, το σύνολο $G = \{0, 1, ..., j\}$ έτσι ώστε $c_j \geq g + \varepsilon$, και τη στάσιμη πολιτική $\tilde{f}_i = P_j$, $i \in G$.

Οι υποθέσεις της Sennott ικανοποιούνται για το παρόν μοντέλο. Σύμφωνα με το Θεώρημα 2 της Sennott (1989), υπάρχει μία σταθερά G και μία αύξουσα ακολουθία αριθμών $\{h_i\}$, $i \geq 1$, έτσι ώστε, στην κατάσταση $i \geq 1$, η βέλτιστη πολιτική επιλέγει εκείνη την ενέργεια η οποία ελαχιστοποιεί το δεξιό μέλος της εξίσωσης (6.6). Τα παραπάνω μας οδηγούν στο ακόλουθο αποτέλεσμα.

Πρόταση 6.2. Αν $c_i \to \infty$, καθώς $i \to \infty$, η βέλτιστη πολιτική είναι μονότονη.

6.3 Η μορφή της συνάρτησης του μέσου κόστους υπό τον έλεγχο μιας μονότονης πολιτικής

Έστω $T_{i0}^{(n)}$ και $C_{i0}^{(n)}$, $i \geq 0$, ο αναμενόμενος χρόνος και το αναμενόμενο κόστος, αντίστοιχα, μέχρι η διαδικασία να βρεθεί στην κατάσταση 0, υπό τον έλεγχο της πολιτικής P_n, $n \geq 1$, αν η αρχική κατάσταση είναι η κατάσταση i. Έστω επίσης G_n το μακροπρόθεσμο αναμενόμενο μέσο κόστος ανά μονάδα χρόνου, υπό τον έλεγχο της πολιτικής P_n. Η διαδικασία, υπό τον έλεγχο της πολιτικής P_n, είναι μία διαδικασία η οποία αναγεννάται οποτεδήποτε βρεθεί στην κατάσταση 0. Αν η διαδικασία βρεθεί στην κατάσταση 0 ένας κύκλος ολοκληρώνεται όταν η διαδικασία ξαναβρεθεί στην κατάσταση 0. Σύμφωνα με το Θεώρημα 5.9 στη σελ. 98 του βιβλίου του Ross (1992), το μακροπρόθεσμο αναμενόμενο μέσο κόστος G_n ανά μονάδα χρόνου, υπό τον έλεγχο της πολιτικής P_n, είναι ίσο με το αναμενόμενο κόστος ενός κύκλου διαιρούμενο με τον αναμενόμενο χρόνο ενός κύκλου. Επομένως, ισχύει ότι:

$$G_n = \frac{\frac{r}{\lambda} + \sum_{i=1}^{\infty} g_i C_{i0}^{(n)}}{T_{00}^{(n)}}. \tag{6.8}$$

Έστω $h_i^{(n)}$, $i \ge 0$, οι σχετικές τιμές της πολιτικής P_n, $n \ge 1$, οι οποίες ορίζονται ως εξής (βλέπε σχέση (2.5) στη σελ. 23 της παρούσας διατριβής):

$$h_i^{(n)} = C_{i0}^{(n)} - G_n T_{i0}^{(n)}. \tag{6.9}$$

Ισχύει ότι:

$$h_0^{(n)} = 0, \tag{6.10}$$

διότι, $G_n = C_{00}^{(n)} / T_{00}^{(n)}$. Ισχύει επίσης ότι:

$$T_{i0}^{(n)} = \mu^{-1}, \quad i \ge n. \tag{6.11}$$

Η επόμενη πρόταση παρέχει μία αναγκαία και ικανή συνθήκη έτσι ώστε η μονότονη πολιτική P_n να είναι βέλτιστη.

Πρόταση 6.3. Η πολιτική P_n, $n \ge 1$, είναι βέλτιστη αν και μόνο αν

$$\frac{c_i - G_n}{\lambda} + \sum_{j=1}^{\infty} g_j h_{i+j}^{(n)} \le \frac{c_i + k - G_n + \lambda \sum_{j=1}^{\infty} g_j h_{i+j}^{(n)} + \mu h_0^{(n)}}{\lambda + \mu}, \quad 1 \le i \le n-1, \tag{6.12}$$

και

$$\frac{c_i + k - G_n + \lambda \sum_{j=1}^{\infty} g_j h_{i+j}^{(n)} + \mu h_0^{(n)}}{\lambda + \mu} \leq \frac{c_i - G_n}{\lambda} + \sum_{j=1}^{\infty} g_j h_{i+j}^{(n)}, \quad i \geq n. \tag{6.13}$$

Απόδειξη. Υποθέτουμε ότι η πολιτική P_n είναι βέλτιστη. Σύμφωνα με τη σχέση (1) της εργασίας των Nobel and Tijms (1999), οι αριθμοί $h_i^{(n)}$, $i \geq 0$, και G_n ικανοποιούν το ακόλουθο σύστημα των γραμμικών εξισώσεων:

$$h_i^{(n)} = \frac{c_i - G_n}{\lambda} + \sum_{j=1}^{\infty} g_j h_{i+j}^{(n)}, \quad 0 \leq i \leq n-1,$$

$$h_i^{(n)} = \frac{c_i + k - G_n + \lambda \sum_{j=1}^{\infty} g_j h_{i+j}^{(n)} + \mu h_0^{(n)}}{\lambda + \mu}, \quad i \geq n.$$

Υποθέτουμε ότι για κάποια κατάσταση $\hat{i}$, $0 \leq \hat{i} \leq n-1$, η ανισότητα (6.12) δεν ισχύει. Τότε σύμφωνα με τη σχέση (2) της εργασίας των Nobel and Tijms (1999), το μακροπρόθεσμο αναμενόμενο μέσο κόστος της στάσιμης πολιτικής $\tilde{f} \equiv \{\tilde{f}_i\}$, $i \geq 0$, η οποία, ορίζεται ως εξής:

$$\tilde{f}_i = 0, \quad 0 \leq i \leq n-1, \quad i \neq \hat{i},$$

$$\tilde{f}_{\hat{i}} = 1,$$

$$\tilde{f}_i = 1, \quad i \geq n,$$

είναι μικρότερο από το μακροπρόθεσμο αναμενόμενο μέσο κόστος της πολιτικής P_n. Αυτό είναι μία αντίφαση. Συνεπώς η ανισότητα (6.12) ισχύει. Η ανισότητα (6.13) μπορεί να αποδειχθεί με παρόμοιο τρόπο.

Υποθέτουμε ότι η πολιτική P_n ικανοποιεί τις ανισότητες (6.12) και (6.13). Τότε οι αριθμοί $h_i^{(n)}$, $i \geq 0$, και G_n ικανοποιούν τις παρακάτω εξισώσεις:

$$h_i^{(n)} = \min\left\{ \frac{c_i - G_n}{\lambda} + \sum_{j=1}^{\infty} g_j h_{i+j}^{(n)},\ \frac{c_i + k - G_n + \lambda \sum_{j=1}^{\infty} g_j h_{i+j}^{(n)} + \mu h_0^{(n)}}{\lambda + \mu} \right\}, \quad i \geq 1. \tag{6.14}$$

Διακρίνουμε τις δύο ακόλουθες περιπτώσεις.

Περίπτωση 1. Υποθέτουμε ότι η ακολουθία $\{c_i\}$ είναι άνω φραγμένη. Τότε, οι αριθμοί $h_i^{(n)}$, $i \geq 0$, είναι άνω φραγμένοι. Από το Θεώρημα 2.4 για τις ημι-Μαρκοβιανές διαδικασίες αποφάσεων (βλέπε (γ) στη σελ. 27 της παρούσας διατριβής) προκύπτει ότι η πολιτική P_n είναι βέλτιστη.

Περίπτωση 2. Υποθέτουμε ότι $c_i \to \infty$, καθώς $i \to \infty$. Από την Πρόταση 6.2 έπεται ότι υπάρχει ένας θετικός ακέραιος αριθμός u τέτοιος ώστε η πολιτική P_u είναι βέλτιστη.

Έστω X_N και a_N η κατάσταση και η ενέργεια που επιλέγεται, αντίστοιχα, μετά από την $N-$οστή μετάβαση της διαδικασίας. Θεωρούμε ότι X_0 και a_0 είναι η αρχική κατάσταση και η αρχική ενέργεια, αντίστοιχα. Έστω επίσης $T(i,a)$ και $C(i,a)$ ο αναμενόμενος χρόνος και το αναμενόμενο κόστος, αντίστοιχα, μέχρι να πραγματοποιηθεί μία μετάβαση της διαδικασίας όταν επιλέγεται η ενέργεια a στην κατάσταση i. Από την απόδειξη του Θεωρήματος 7.6 στη σελ. 162 του βιβλίου του Ross (1992), έχουμε ότι:

$$G_n \leq \frac{E_u[h_{X_N}^{(n)} - h_{X_0}^{(n)}] + E_u\left[\sum_{i=1}^{N} C(X_{i-1}, a_{i-1})\right]}{E_u\left[\sum_{i=1}^{N} T(X_{i-1}, a_{i-1})\right]}, \tag{6.15}$$

όπου, E_u αναπαριστά την υπό-συνθήκη αναμενόμενη τιμή, δοθέντος ότι η πολιτική P_u έχει υιοθετηθεί για τον έλεγχο της διαδικασίας. Η σχέση (6.15) ισχύει ως ισότητα για $n = u$. Χρησιμοποιούμε τον ορισμό της σχετικής τιμής $h_{X_N}^{(n)}$, και γράφουμε τη σχέση (6.15) με την ακόλουθη μορφή:

$$G_n \leq \frac{E_u[C_{X_N 0}^{(n)}] - G_n E_u[T_{X_N 0}^{(n)}] - h_{X_0}^{(n)} + E_u\left[\sum_{i=1}^{N} C(X_{i-1}, a_{i-1})\right]}{E_u\left[\sum_{i=1}^{N} T(X_{i-1}, a_{i-1})\right]},$$

ή ισοδύναμα,

$$G_n \leq \frac{\sum_{j=0}^{\infty} p_{X_0 j}^{(N)} C_{j0}^{(n)} - G_n \sum_{j=0}^{\infty} p_{X_0 j}^{(N)} T_{j0}^{(n)} - h_{X_0}^{(n)} + E_u\left[\sum_{i=1}^{N} C(X_{i-1}, a_{i-1})\right]}{E_u\left[\sum_{i=1}^{N} T(X_{i-1}, a_{i-1})\right]}, \tag{6.16}$$

όπου, $p_{X_0 j}^{(N)}$ είναι η πιθανότητα η κατάσταση j να είναι η κατάσταση της διαδικασίας μετά από N μεταβάσεις, υπό τον έλεγχο της πολιτικής P_u.

Αφήνουμε τον αριθμό N να τείνει στο άπειρο και χρησιμοποιούμε το Θεώρημα Φραγμένης Σύγκλισης με το οποίο είναι δυνατόν να εναλλάξουμε τη σειρά του ορίου και του αθροίσματος στον αριθμητή του δεξιού μέλους της ανισότητας (6.16). Η εναλλαγή είναι εφικτή διότι ισχύουν οι ακόλουθες σχέσεις:

$$\lim_{N\to\infty} \sum_{j=0}^{\infty} p_{X_0 j}^{(N)} C_{j0}^{(n)} = \sum_{j=0}^{\infty} p_j C_{j0}^{(n)}, \tag{6.17}$$

και

$$\lim_{N\to\infty} \sum_{j=0}^{\infty} p_{X_0 j}^{(N)} T_{j0}^{(n)} = \sum_{j=0}^{\infty} p_j T_{j0}^{(n)}, \tag{6.18}$$

όπου, $p_j = \lim_{N\to\infty} p_{X_0 j}^{(N)}, \quad j \geq 0,$ είναι οι στάσιμες πιθανότητες της εμφυτευμένης Μαρκοβιανής αλυσίδας της διαδικασίας υπό τον έλεγχο της πολιτικής P_u. Οι σχέσεις (6.17) και (6.18) είναι συνέπειες του ακόλουθου θεωρήματος (βλέπε σελ. 351 του βιβλίου του Tijms (1994)) το οποίο παραθέτουμε παρενθετικά. Αυτό το θεώρημα είναι μία ειδική περίπτωση του Θεωρήματος Φραγμένης Σύγκλισης.

Θεώρημα 6.1. Έστω $\{p_m, m = 0, 1, ...\}$ μία ακολουθία μη-αρνητικών αριθμών και έστω ότι οι αριθμοί a_{nm}, $n, m = 0, 1, ...$ είναι τέτοιοι ώστε το όριο $\lim_{n\to\infty} a_{nm} = a_m$ υπάρχει για κάθε $m = 0, 1, ...$. Αν υπάρχει μία σταθερά $M > 0$ τέτοια ώστε $|a_{nm}| \le M$ για κάθε n, m και αν $\sum_{m=0}^{\infty} p_m < \infty$, τότε:

$$\lim_{n\to\infty} \sum_{m=0}^{\infty} a_{nm} p_m = \sum_{m=0}^{\infty} a_m p_m .$$

Θα εξηγήσουμε τον τρόπο με τον οποίο εξάγεται η σχέση (6.17). Για να υπολογίσουμε το αναμενόμενο κόστος $C_{j0}^{(n)}$, $j \ge n$, δεσμευόμαστε ως προς το χρόνο μέχρι να πραγματοποιηθεί μία καταστροφή του πληθυσμού των παρασίτων. Ισχύει ότι:

$$C_{j0}^{(n)} = \int_0^{\infty} \left[\int_0^t E\{c_{X(s)} + k \mid X(0) = j\} ds \right] \mu e^{-\mu t} dt,$$

όπου, $X(s)$ αναπαριστά το μέγεθος του πληθυσμού τη χρονική στιγμή s, για τη σύνθετη διαδικασία μετανάστευσης. Επειδή ισχύει ότι $c_i \le A i^m$, για μία σταθερά $A > 0$ και έναν ακέραιο αριθμό m, από την παραπάνω εξίσωση, προκύπτει η ακόλουθη ανισότητα:

$$C_{j0}^{(n)} \le A \int_0^{\infty} \left[\int_0^t E\{(X(s))^m \mid X(0) = j\} ds \right] \mu e^{-\mu t} dt + k\mu^{-1}.$$

Από τον ορισμό της σύνθετης διαδικασίας Poisson έπεται ότι η παραπάνω ανισότητα μπορεί ισοδύναμα να γραφεί ως εξής:

$$C_{j0}^{(n)} \le A\int_0^\infty \left[\int_0^t E\{(j+Y_1+\cdots+Y_{N(s)})^m\}ds \right] \mu e^{-\mu t}dt + k\mu^{-1},$$

όπου, $N(s)$ είναι μία τυχαία μεταβλητή η οποία ακολουθεί την κατανομή Poisson με παράμετρο λs και Y_i, $i \ge 1$, είναι ανεξάρτητες τυχαίες μεταβλητές τέτοιες ώστε: $P(Y_i = \ell) = g_\ell$, $\ell = 1, 2, \ldots$. Λόγω της Συνθήκης 2, το δεξιό μέλος της παραπάνω ανισότητας είναι ένα πολυώνυμο $P(j)$ ως προς j βαθμού m. Το άπειρο άθροισμα στο αριστερό μέλος της εξίσωσης (6.17) μπορεί να γραφεί με την ακόλουθη μορφή:

$$\sum_{j=0}^\infty p_{X_0 j}^{(N)} j^2 P(j) \frac{C_{j0}^{(n)}}{j^2 P(j)}. \tag{6.19}$$

Από το Θεώρημα 3.1 του Economou (2003) προκύπτει ότι, για αρκετά μεγάλες τιμές των N και j, οι στάσιμες πιθανότητες $p_{X_0 j}^{(N)}$ είναι κατά προσέγγιση ίσες με $c\theta^{-j}$, όπου $c > 0$ και $0 < \theta < 1$. Συνεπώς η ποσότητα $p_{X_0 j}^{(N)} j^2 P(j)$, για αρκετά μεγάλες τιμές των N και j, είναι φραγμένη. Επειδή $C_{j0}^{(n)} < P(j)$, έχουμε επίσης ότι: $\sum_j C_{j0}^{(n)} j^{-2}[P(j)]^{-1} < \infty$. Επομένως, το Θεώρημα 6.1 μπορεί να εφαρμοστεί αν πάρουμε το όριο καθώς $N \to \infty$ της έκφρασης (6.19) και εναλλάξουμε τη σειρά του ορίου και του αθροίσματος. Τότε παίρνουμε το δεξιό μέλος της εξίσωσης (6.17). Η εξίσωση (6.18) προκύπτει με παρόμοιο τρόπο.

Λόγω των σχέσεων (6.17) και (6.18), η ανισότητα (6.16) γράφεται με την ακόλουθη μορφή:

$$G_n \le \frac{\sum_{j=0}^\infty p_j C_{j0}^{(n)} - G_n \sum_{j=0}^\infty p_j T_{j0}^{(n)} - h_{X_0}^{(n)}}{\lim_{N\to\infty} E_u\left[\sum_{i=1}^N T(X_{i-1}, a_{i-1})\right]} + \lim_{N\to\infty} \frac{E_u\left[\sum_{i=1}^N C(X_{i-1}, a_{i-1})\right]}{E_u\left[\sum_{i=1}^N T(X_{i-1}, a_{i-1})\right]}.$$

Ο αριθμητής του πρώτου όρου του αθροίσματος στο δεξιό μέλος της παραπάνω ανισότητας είναι πεπερασμένος διότι οι στάσιμες πιθανότητες p_j, $j \ge 0$, ακολουθούν ασυμπτωτικά τη Γεωμετρική κατανομή (βλέπε Θεώρημα 3.1 του Economou (2003)). Επειδή, ο παρονομαστής του πρώτου όρου του αθροίσματος στο δεξιό μέλος της παραπάνω ανισότητας, καθώς $N \to \infty$, τείνει στο άπειρο, συμπεραίνουμε ότι αυτός ο όρος είναι ίσος με το μηδέν. Ο δεύτερος όρος του αθροίσματος στο δεξιό μέλος της παραπάνω ανισότητας είναι το μακροπρόθεσμο αναμενόμενο μέσο κόστος G_u ανά μονάδα χρόνου, υπό τον έλεγχο της πολιτικής P_u. Συνεπώς ισχύει ότι $G_n \le G_u$, και η πολιτική P_n είναι βέλτιστη. ■

Τα αποτελέσματα των Λημμάτων 6.1, 6.2 και 6.3 θα χρησιμοποιηθούν στην απόδειξη του Θεωρήματος 6.2.

Λήμμα 6.1. Για κάθε σταθερό $n \ge 1,$ η ακολουθία $C_{i0}^{(n)}, \; i \ge n,$ είναι αύξουσα ως προς $i.$

Απόδειξη. Έστω $X(s), \; s \ge 0,$ το μέγεθος του πληθυσμού τη χρονική στιγμή $s,$ για τη (μη ελεγχόμενη) σύνθετη διαδικασία μετανάστευσης. Δεσμευόμαστε ως προς το χρόνο μέχρι την πραγματοποίηση μιας καταστροφής του πληθυσμού των παρασίτων και έχουμε ότι:

$$C_{i0}^{(n)} = \int_0^\infty \left[\int_0^t E\{c_{X(s)} + k \mid X(0) = i\} ds \right] \mu e^{-\mu t} dt$$

$$= \int_0^\infty \left[\int_0^t E\{c_{X(s)+i} \mid X(0) = 0\} ds \right] \mu e^{-\mu t} dt + k\mu^{-1}. \tag{6.20}$$

Η πρώτη ισότητα είναι συνέπεια των ορισμών που δώσαμε στο Εδάφιο 6.1 για τα κόστη του προβλήματος και η δεύτερη ισότητα είναι συνέπεια του γεγονότος ότι η κατανομή της τυχαίας μεταβλητής $X(s),$ αν $X(0) = i,$ συμπίπτει με την κατανομή της τυχαίας μεταβλητής $X(s) + i,$ αν $X(0) = 0.$ Από τη μονοτονία της ακολουθίας $\{c_i\}$ και την εξίσωση (6.20), συμπεραίνουμε ότι η ακολουθία $C_{i0}^{(n)}$ είναι αύξουσα ως προς $i, \; i \ge n.$ ■

Λήμμα 6.2. Έστω ότι η πολιτική $P_n, \; n \ge 1,$ είναι βέλτιστη για κάποια σταθερή τιμή R της παραμέτρου $r.$ Τότε, είναι αδύνατον η πολιτική P_n να είναι βέλτιστη για κάθε $r \ge R.$

Απόδειξη. Υποθέτουμε ότι η πολιτική P_n είναι βέλτιστη για κάθε $r \geq R.$ Τότε οι αριθμοί $h_i^{(n)},$ $i \geq 0,$ ικανοποιούν τη σχέση (6.13) για $i = n$ και για κάθε $r \geq R.$ Επομένως, ισχύει ότι:

$$\frac{c_n + k - G_n + \mu h_0^{(n)}}{\lambda + \mu} + \lambda \sum_{j=1}^{\infty} \frac{g_j}{\lambda + \mu} h_{n+j}^{(n)} \leq \frac{c_n - G_n}{\lambda} + \sum_{j=1}^{\infty} g_j h_{n+j}^{(n)}.$$

Χρησιμοποιώντας τις σχέσεις (6.9), (6.10) και (6.11), η παραπάνω ανισότητα μπορεί να γραφεί ως εξής:

$$(\lambda + \mu) G_n \leq \mu c_n + \lambda \mu \sum_{j=1}^{\infty} g_j C_{n+j,0}^{(n)} - \lambda k.$$

Από τη εξίσωση (6.8), προκύπτει ότι $G_n \to \infty,$ καθώς $r \to \infty.$ Αυτό έρχεται σε αντίφαση με την παραπάνω ανισότητα διότι το δεξιό μέλος της δεν εξαρτάται από την παράμετρο r και ο συντελεστής του μέσου κόστους G_n στο αριστερό μέλος της είναι θετικός. ■

Έστω $P_{n(r)}$ η βέλτιστη μονότονη πολιτική, όπου $n(r)$ είναι η βέλτιστη κρίσιμη τιμή. Σύμφωνα με τις Προτάσεις 6.1 και 6.2 η πολιτική $P_{n(r)}$ είναι η βέλτιστη πολιτική στο σύνολο όλων των δυνατών πολιτικών.

Λήμμα 6.3. Η βέλτιστη κρίσιμη τιμή $n(r)$ είναι αύξουσα ως προς την παράμετρο $r \in (-\infty, \infty).$

Απόδειξη. Υποθέτουμε ότι

$$r_1 < r_2. \tag{6.21}$$

Έστω n_1, n_2 οι βέλτιστες κρίσιμες τιμές που αντιστοιχούν στις παραμέτρους r_1 και r_2, αντίστοιχα. Έστω επίσης G_{n_1} και G_{n_2} τα μέσα κόστη των πολιτικών P_{n_1} και P_{n_2}, αντίστοιχα. Η σχέση (6.21) υποδηλώνει ότι:

$$G_{n_1} < G_{n_2}. \tag{6.22}$$

Θέλουμε να δείξουμε ότι $n_1 \le n_2$. Έστω ότι, αντιθέτως, ισχύει $n_1 > n_2$. Από τη σχέση (6.12) για $i = n_1 - 1$, έχουμε ότι:

$$\frac{c_{n_1-1} - G_{n_1}}{\lambda} + \sum_{j=1}^{\infty} g_j h_{n_1-1+j}^{(n_1)} \le \frac{c_{n_1-1} + k - G_{n_1} + \lambda \sum_{j=1}^{\infty} g_j h_{n_1-1+j}^{(n_1)} + \mu h_0^{(n_1)}}{\lambda + \mu},$$

διότι η πολιτική P_{n_1} είναι βέλτιστη. Από τις εξισώσεις (6.9), (6.10) και (6.11), η παραπάνω ανισότητα μπορεί ισοδύναμα να γραφεί με την ακόλουθη μορφή:

$$G_{n_1}(\lambda + \mu) \ge \mu c_{n_1-1} + \lambda\mu \sum_{j=1}^{\infty} g_j C_{n_1-1+j,0}^{(n_1)} - \lambda k. \tag{6.23}$$

Από τη σχέση (6.13) για $i = n_2$, έχουμε ότι:

$$\frac{c_{n_2} + k - G_{n_2} + \lambda \sum_{j=1}^{\infty} g_j h_{n_2+j}^{(n_2)} + \mu h_0^{(n_2)}}{\lambda + \mu} \le \frac{c_{n_2} - G_{n_2}}{\lambda} + \sum_{j=1}^{\infty} g_j h_{n_2+j}^{(n_2)},$$

διότι η πολιτική P_{n_2} είναι βέλτιστη. Από τις εξισώσεις (6.9), (6.10) και (6.11), η παραπάνω ανισότητα μπορεί ισοδύναμα να γραφεί με την ακόλουθη μορφή:

$$G_{n_2}(\lambda + \mu) \le \mu c_{n_2} + \lambda\mu \sum_{j=1}^{\infty} g_j C_{n_2+j,0}^{(n_2)} - \lambda k. \tag{6.24}$$

Από τις ανισότητες (6.22), (6.23) και (6.24) προκύπτει ότι:

$$\mu c_{n_2} + \lambda\mu\sum_{j=1}^{\infty} g_j C_{n_2+j,0}^{(n_2)} - \lambda k > \mu c_{n_1-1} + \lambda\mu\sum_{j=1}^{\infty} g_j C_{n_1-1+j,0}^{(n_1)} - \lambda k.$$

Το αριστερό μέλος της παραπάνω σχέσης είναι μικρότερο ή ίσο με το δεξιό μέλος της διότι $c_{n_2} \le c_{n_1-1}$ και $C_{n_2+j,0}^{(n_2)} \le C_{n_1-1+j,0}^{(n_1)}$. Αυτό είναι μία αντίφαση. ■

Στο ακόλουθο Θεώρημα 6.2 αποδεικνύεται ότι, για κάθε $n = 1, 2, \ldots$ η πολιτική P_n είναι βέλτιστη όταν η πλασματική παράμετρος r ανήκει σε ένα κατάλληλο διάστημα των πραγματικών αριθμών.

Θεώρημα 6.2. Υπάρχει μία αύξουσα ακολουθία $\{R_n\}$, $n \ge 1$, με $R_1 = -\infty$, $R_2 > -\infty$, και $\lim_{n\to\infty} R_n = \infty$, τέτοια ώστε η πολιτική P_n, $n \ge 1$, είναι βέλτιστη για κάθε $r \in [R_n, R_{n+1}]$, όπου

$$R_{n+1} = \sup\{w : w \ge R_n \text{ και η πολιτική } P_n \text{ είναι βέλτιστη για κάθε } r \in [R_n, w]\}.$$

Απόδειξη. Η απόδειξη γίνεται με επαγωγή ως προς n. Αρχικά αποδεικνύουμε ότι υπάρχει ένας αριθμός $R > -\infty$ τέτοιος ώστε η πολιτική P_1 είναι βέλτιστη για κάθε $r \le R$. Σύμφωνα με την Πρόταση 6.3 αρκεί να δείξουμε ότι οι αριθμοί $h_i^{(1)}$, $i \ge 1$, και G_{n_1} ικανοποιούν την ανισότητα:

$$\frac{c_i + k - G_{n_1} + \lambda\sum_{j=1}^{\infty} g_j h_{i+j}^{(1)} + \mu h_0^{(1)}}{\lambda + \mu} \le \frac{c_i - G_{n_1}}{\lambda} + \sum_{j=1}^{\infty} g_j h_{i+j}^{(1)}, \quad i \ge 1.$$

Χρησιμοποιώντας τις εξισώσεις (6.9), (6.10) και (6.11), η παραπάνω ανισότητα γράφεται με την ακόλουθη μορφή:

$$(\lambda+\mu)G_{n_1} \le \mu c_i + \lambda\mu\sum_{j=1}^{\infty} g_j C_{i+j,0}^{(1)} - \lambda k, \quad i \ge 1.$$

Από τη σχέση (6.8), έχουμε ότι $G_{n_1} \to -\infty,$ καθώς $r \to -\infty.$ Επομένως, υπάρχει ένας αριθμός $R > -\infty$ τέτοιος ώστε οι παραπάνω ανισότητες ισχύουν για κάθε $r \le R.$ Από το Λήμμα 6.2, προκύπτει ότι $R_2 < \infty,$ όπου

$$R_2 = \sup\{w : w \ge R \text{ και η πολιτική } P_1 \text{ είναι βέλτιστη για κάθε } r \le w\}.$$

Υποθέτουμε ότι υπάρχει μία ακολουθία $R_1 < R_2 < \cdots < R_n,$ τέτοια ώστε η πολιτική $P_s,$ $1 \le s \le n,$ είναι βέλτιστη για κάθε $r \in [R_s, R_{s+1}],$ με $R_{n+1} = \sup\{w : w \ge R_n$ και η πολιτική P_n είναι βέλτιστη για κάθε $r \in [R_n, w]\} < \infty.$ Θα δείξουμε ότι η πολιτική P_{n+1} είναι βέλτιστη για $r = R_{n+1}.$ Θεωρούμε κάποιο $\varepsilon > 0$ τέτοιο ώστε η πολιτική P_n δεν είναι βέλτιστη για $r = R_{n+1} + \varepsilon.$ Από την Πρόταση 6.3 έπεται ότι μία από τις δύο ακόλουθες περιπτώσεις συμβαίνουν.

Περίπτωση 1. Για κάποια κατάσταση i τέτοια ώστε $1 \le i \le n-1,$ ισχύει ότι:

$$\frac{c_i + k - G_n + \lambda\sum_{j=1}^{\infty} g_j h_{i+j}^{(n)} + \mu h_0^{(n)}}{\lambda+\mu} < \frac{c_i - G_n}{\lambda} + \sum_{j=1}^{\infty} g_j h_{i+j}^{(n)}.$$

Χρησιμοποιώντας τις εξισώσεις (6.9) και (6.10), η παραπάνω ανισότητα είναι ισοδύναμη με την $\psi_i(R_{n+1} + \varepsilon) > 0,$ όπου

$$\psi_i(r) = \mu c_i - \mu G_n + \lambda\mu\sum_{j=1}^{\infty} g_j (C_{i+j,0}^{(n)} - G_n T_{i+j,0}^{(n)}) - \lambda k.$$

Επειδή το μέσο κόστος G_n, όπως δίνεται από την εξίσωση (6.8), είναι μία αύξουσα συνάρτηση ως προς r, συμπεραίνουμε ότι η συνάρτηση $\psi_i(r)$ είναι φθίνουσα ως προς r. Επομένως ισχύει ότι:

$$0 < \psi_i(R_{n+1} + \varepsilon) < \psi_i(R_{n+1}) \leq 0,$$

όπου η τελευταία ανισότητα προκύπτει από το γεγονός ότι η πολιτική P_n είναι βέλτιστη για $r = R_{n+1}$. Αυτό είναι προφανώς μία αντίφαση και συνεπώς ισχύει η Περίπτωση 2.

Περίπτωση 2. Για κάποια κατάσταση i τέτοια ώστε $i \geq n$, ισχύει ότι:

$$\frac{c_i - G_n}{\lambda} + \sum_{j=1}^{\infty} g_j h_{i+j}^{(n)} < \frac{c_i + k - G_n + \lambda \sum_{j=1}^{\infty} g_j h_{i+j}^{(n)} + \mu h_0^{(n)}}{\lambda + \mu}.$$

Χρησιμοποιώντας τις εξισώσεις (6.9), (6.10) και (6.11), η παραπάνω ανισότητα είναι ισοδύναμη με την $\psi_i(R_{n+1} + \varepsilon) < 0$, όπου

$$\psi_i(r) = \mu c_i - (\lambda + \mu) G_n + \lambda \mu \sum_{j=1}^{\infty} g_j C_{i+j,0}^{(n)} - \lambda k.$$

Από το Λήμμα 6.1 και τη μονοτονία της ακολουθίας $\{c_i\}$ προκύπτει ότι η ακολουθία $\psi_i(r)$, $i \geq n$, είναι αύξουσα ως προς i. Επομένως, ισχύει ότι:

$$\psi_n(R_{n+1} + \varepsilon) \leq \psi_i(R_{n+1} + \varepsilon) < 0.$$

Από τον ορισμό του R_{n+1} και την παραπάνω ανάλυση, υπάρχει μία ακολουθία $\{\varepsilon_k\} \downarrow 0$ τέτοια ώστε $\psi_n(R_{n+1} + \varepsilon_k) < 0$. Επειδή η συνάρτηση $\psi_n(r)$ είναι συνεχής ως προς r, έπεται ότι $\psi_n(R_{n+1}) \leq 0$. Επειδή η πολιτική P_n είναι βέλτιστη για $r = R_{n+1}$, ισχύει ότι $\psi_n(R_{n+1}) \geq 0$. Άρα,

$\psi_n(R_{n+1}) = 0.$ Η τελευταία ισότητα σημαίνει ότι για $r = R_{n+1}$ το δεξιό μέλος της εξίσωσης (6.14) έχει την ίδια τιμή στην κατάσταση $i = n,$ και για τις δύο δυνατές ενέργειες, 0 και 1. Επομένως, για κάθε κατάσταση $i \geq 0,$ η ενέργεια η οποία επιλέγεται από την πολιτική P_{n+1} ελαχιστοποιεί το δεξιό μέλος της εξίσωσης (6.14). Συνεπώς η πολιτική P_{n+1} είναι βέλτιστη για $r = R_{n+1}.$ Ορίζουμε:

$$R_{n+2} = \sup\{w : w \geq R_{n+1} \text{ και η πολιτική } P_{n+1} \text{ είναι βέλτιστη για κάθε } r \in [R_{n+1}, w]\}.$$

Από το Λήμμα 6.2, έχουμε ότι $R_{n+2} < \infty.$ Αρκεί να δείξουμε ότι $\lim_{n\to\infty} R_n = \infty.$ Υποθέτουμε, αντιθέτως, ότι $\lim_{n\to\infty} R_n = R^* < \infty.$ Επιλέγουμε $n^* > n(R^*),$ όπου $n(R^*)$ είναι η βέλτιστη κρίσιμη τιμή δοθέντος ότι η τιμή της πλασματικής παραμέτρου είναι ίση με $R^*.$ Από το Λήμμα 6.3 προκύπτει ότι $n(R^*) \geq n(R_{n^*}).$ Θεωρούμε μία κατάσταση i τέτοια ώστε $n^* > i \geq n(R_{n^*}).$ Για $r = R_{n^*}$ στην κατάσταση i η βέλτιστη πολιτική επιλέγει την ενέργεια 1. Αυτό έρχεται σε αντίφαση με το γεγονός ότι η πολιτική P_{n^*} είναι βέλτιστη για $r = R_{n^*}.$ ■

Σημείωση 6.1. Τονίζουμε ότι, όταν η ακολουθία $\{c_i\},\ i \geq 0,$ είναι άνω φραγμένη, το Θεώρημα 6.2 ισχύει χωρίς να είναι απαραίτητο να ισχύουν οι Συνθήκες 1 και 2, οι οποίες έχουν εισαχθεί στο Εδάφιο 6.1 για τη συνάρτηση πιθανότητας $\{g_j,\ j = 1, 2, ...\},$ διότι, στην περίπτωση αυτή, οι Συνθήκες 1 και 2 δεν είναι αναγκαίες για την απόδειξη της Πρότασης 6.3.

Λήμμα 6.4. Το μέσο κόστος G_n μιας μονότονης πολιτικής $P_n,\ n \geq 1,$ μπορεί να γραφεί ως εξής:

$$G_n = rA_n + \widetilde{G}_n,$$

όπου, η ποσότητα A_n είναι φθίνουσα ως προς n και η ποσότητα $\widetilde{G}_n$ δεν εξαρτάται από την παράμετρο $r.$

Απόδειξη. Από την εξίσωση (6.8), το μέσο κόστος G_n μπορεί να γραφεί ως εξής:

$$G_n = \frac{r}{\lambda T_{00}^{(n)}} + \widetilde{G}_n,$$

όπου, η ποσότητα $\widetilde{G}_n$ δεν εξαρτάται από την παράμετρο r. Έστω $T_{iE_n}^{(n)}$, $0 \le i \le n-1$, ο αναμενόμενος χρόνος μέχρι την είσοδο της διαδικασίας στο σύνολο των καταστάσεων $E_n = \{n, n+1, ...\}$, αν η αρχική κατάσταση είναι η κατάσταση i και η πολιτική P_n έχει υιοθετηθεί για τον έλεγχο της διαδικασίας. Αν η διαδικασία βρεθεί στην κατάσταση 0 ένας κύκλος ολοκληρώνεται όταν η διαδικασία ξαναβρεθεί στην κατάσταση 0. Ο αναμενόμενος χρόνος ενός κύκλου, υπό τον έλεγχο της πολιτικής P_n, δίνεται από τη σχέση:

$$T_{00}^{(n)} = T_{0E_n}^{(n)} + \mu^{-1}.$$

Αρκεί να δείξουμε ότι η ακολουθία $T_{0E_n}^{(n)}$ είναι αύξουσα ως προς n. Η απόδειξη γίνεται εύκολα με επαγωγή ως προς n. Για $n=1$, η ανισότητα $T_{0E_1}^{(1)} \le T_{0E_2}^{(2)}$ είναι ισοδύναμη με την ανισότητα $\lambda^{-1} \le \lambda^{-1} + g_1 T_{1E_2}^{(2)}$, η οποία προφανώς ισχύει. Υποθέτουμε ότι:

$$T_{0E_j}^{(j)} \le T_{0E_{j+1}}^{(j+1)}, \quad 1 \le j \le n-1. \tag{6.25}$$

Θέλουμε να δείξουμε την ανισότητα $T_{0E_n}^{(n)} \le T_{0E_{n+1}}^{(n+1)}$, η οποία είναι ισοδύναμη με την ανισότητα:

$$\lambda^{-1} + \sum_{\ell=1}^{n-1} g_\ell T_{\ell E_n}^{(n)} \le \lambda^{-1} + \sum_{\ell=1}^{n-1} g_\ell T_{\ell E_{n+1}}^{(n+1)} + g_n T_{nE_{n+1}}^{(n+1)}. \tag{6.26}$$

Ισχύει ότι:

$$T_{\ell E_n}^{(n)} = T_{0E_{n-\ell}}^{(n-\ell)}, \quad 1 \le \ell \le n-1, \tag{6.27}$$

και

$$T^{(n+1)}_{\ell E_{n+1}} = T^{(n+1-\ell)}_{0E_{n+1-\ell}}, \quad 1 \le \ell \le n-1. \tag{6.28}$$

Από τη σχέση (6.25) έχουμε ότι:

$$T^{(n-\ell)}_{0E_{n-\ell}} \le T^{(n+1-\ell)}_{0E_{n+1-\ell}}, \quad 1 \le \ell \le n-1. \tag{6.29}$$

Από τις σχέσεις (6.27), (6.28) και (6.29) προκύπτει η ανισότητα (6.26). ■

Χρησιμοποιούμε το Λήμμα 6.4 και παίρνουμε το ακόλουθο αποτέλεσμα το οποίο είναι χρήσιμο για τον υπολογισμό μιας βέλτιστης μονότονης πολιτικής.

Πρόταση 6.4. Για κάθε σταθερή τιμή της παραμέτρου r, το μέσο κόστος G_n, $n \ge 1$, είναι μία μονοκόρυφη συνάρτηση ως προς n.

Απόδειξη. Πρέπει να δείξουμε ότι για κάθε σταθερή τιμή των r και n, $n \ge 1$, ισχύει ότι:

(i) αν $G_n < G_{n+1}$ τότε $G_n < G_{n+s}$, για κάθε $s \ge 2$, και

(ii) αν $G_{n+1} \le G_n$ τότε $G_{n+1} \le G_{n-s}$, για κάθε $s \ge 1$.

Θα αποδείξουμε το (i). Από το Λήμμα 6.4 ισχύει ότι $G_n = rA_n + \widetilde{G}_n$, όπου η ποσότητα A_n είναι φθίνουσα ως προς n και η ποσότητα $\widetilde{G}_n$ δεν εξαρτάται από την παράμετρο r. Επομένως, αν για κάποια σταθερή τιμή της παραμέτρου r, έστω την τιμή r_0, και κάθε σταθερή τιμή του n, ισχύει ότι $G_n < G_{n+1}$, τότε $G_n < G_{n+1}$ για κάθε $r < r_0$. Για να αποδείξουμε το (i), υποθέτουμε, αντιθέτως, ότι $G_n > G_{n+s}$ και επομένως $G_{n+1} > G_{n+s}$, για κάθε $s \ge 2$. Από τα παραπάνω προκύπτει επίσης ότι $G_{n+1} > G_{n+s}$, για κάθε $r \ge r_0$ και κάθε $s \ge 2$. Συνεπώς η πολιτική P_{n+1} δεν είναι βέλτιστη για καμία τιμή της παραμέτρου r. Αυτό έρχεται σε αντίφαση με το Θεώρημα 6.2. Η απόδειξη του (ii) γίνεται με παρόμοιο τρόπο. ■

6.4 Ο υπολογισμός της βέλτιστης πολιτικής

Η Πρόταση 6.4 μας επιτρέπει να σχεδιάσουμε δύο αποδοτικούς αλγόριθμους για να υπολογίσουμε τη βέλτιστη πολιτική. Στο παρόν εδάφιο υποθέτουμε ότι η παράμετρος r παίρνει την τιμή 0, δηλαδή θεωρούμε το μοντέλο που περιγράψαμε στο Εδάφιο 6.1. Η κρίσιμη τιμή n^* της βέλτιστης μονότονης πολιτικής P_{n^*} μπορεί να βρεθεί με τη συνήθη μέθοδο της διχοτόμησης η οποία υπολογίζει την ελάχιστη τιμή του συνόλου των τιμών $\{G_n, n \geq 1\}$.

Έστω N ένας θετικός ακέραιος αριθμός τέτοιος ώστε $G_N < G_{N+1}$. Η μέθοδος της διχοτόμησης παρουσιάζεται παρακάτω.

Μέθοδος της Διχοτόμησης

Βήμα 1. Αρχικά θέτουμε $n_1 = 1$ και $n_2 = N$.

Βήμα 2. Έστω n το ακέραιο μέρος του πηλίκου $(n_1 + n_2)/2$. Αν $G_n < G_{n+1}$, θέτουμε $n_2 = n$; αν $G_n > G_{n+1}$, θέτουμε $n_1 = n+1$; αν $G_n = G_{n+1}$ και $G_{n-1} \leq G_n$, θέτουμε $n_2 = n$; αν $G_n = G_{n+1}$ και $G_{n-1} > G_n$, θέτουμε $n_1 = n+1$. Αν $n_2 - n_1 > 1$, επαναλαμβάνουμε το Βήμα 2. Διαφορετικά, επιλέγουμε n^* τέτοιο ώστε $G_{n^*} = \min\{G_{n_1}, G_{n_2}\}$. Η βέλτιστη πολιτική είναι η πολιτική P_{n^*} με μέσο κόστος ίσο με G_{n^*}.

Χρησιμοποιούμε το Θεώρημα 5.10 στη σελ. 98 του βιβλίου του Ross (1992) και εκφράζουμε τα μέσα κόστη G_n, υπό τον έλεγχο της πολιτικής P_n, στο Βήμα 2 της μεθόδου της διχοτόμησης, ως εξής:

$$G_n = \sum_{i=0}^{n-1} c_i p_i + \sum_{i=n}^{\infty} (c_i + k) p_i,$$

όπου p_i, $i \geq 0$, είναι οι στάσιμες πιθανότητες της διαδικασίας. Ο Economou (2003) πρότεινε έναν αλγόριθμο ο οποίος υπολογίζει με προσεγγιστικό τρόπο τις στάσιμες πιθανότητες p_i, $i \geq 0$. Μία μέθοδος για τον ακριβή υπολογισμό των στάσιμων πιθανοτήτων παρουσιάζεται παρακάτω.

Δεσμευόμαστε ως προς την πρώτη μετάβαση της διαδικασίας, υπό τον έλεγχο της πολιτικής P_n, αν η αρχική κατάσταση είναι η κατάσταση i, $0 \le i \le n-1$, και έχουμε ότι:

$$T_{n-1,0}^{(n)} = \lambda^{-1} + \mu^{-1}, \tag{6.30}$$

$$T_{i0}^{(n)} = \lambda^{-1} + \sum_{j=1}^{n-i-1} g_j T_{i+j,0}^{(n)} + \mu^{-1} \sum_{j=n-i}^{\infty} g_j, \quad 0 \le i \le n-2. \tag{6.31}$$

Υπό τον έλεγχο της πολιτικής P_n, η διαδικασία αναγεννάται οποτεδήποτε βρεθεί στην κατάσταση 0. Αν η διαδικασία βρεθεί στην κατάσταση 0 ένας κύκλος ολοκληρώνεται όταν η διαδικασία ξαναβρεθεί στην κατάσταση 0. Σύμφωνα με το Θεώρημα 5.8 στη σελ. 95 του βιβλίου του Ross (1992), η στάσιμη πιθανότητα p_0 δίνεται από τον αναμενόμενο χρόνο κατά τον οποίον η διαδικασία βρίσκεται στην κατάσταση 0 στη διάρκεια ενός κύκλου διαιρούμενο με τον αναμενόμενο χρόνο ενός κύκλου. Επομένως, ισχύει ότι:

$$p_0 = \frac{\lambda^{-1}}{T_{00}^{(n)}}.$$

Ο παρονομαστής της παραπάνω σχέσης μπορεί να υπολογιστεί από τις εξισώσεις (6.30) και (6.31). Οι άλλες στάσιμες πιθανότητες p_i, $i \ge 1$, μπορούν να βρεθούν από τις ακόλουθες εξισώσεις ισορροπίας:

$$p_i = \sum_{j=0}^{i-1} g_{i-j} p_j, \quad 1 \le i \le n-1,$$

$$p_i = \frac{\lambda}{\lambda + \mu} \sum_{j=0}^{i-1} g_{i-j} p_j, \quad i \ge n.$$

Ένας εναλλακτικός τρόπος για τον υπολογισμό της βέλτιστης πολιτικής είναι ο σχεδιασμός ενός κατάλληλου αλγορίθμου βελτίωσης των πολιτικών ο οποίος παράγει μία ακολουθία βελτιωμένων πολιτικών που έχουν τη μονότονη μορφή. Η Πρόταση 6.4 εγγυάται ότι αυτή η ακολουθία συγκλίνει στη βέλτιστη πολιτική. Ο σχεδιασμός του αλγορίθμου παρουσιάζεται παρακάτω.

Η εξίσωση (6.20) δίνει το αναμενόμενο κόστος $C_{i0}^{(n)}$ μέχρι η διαδικασία να βρεθεί στην κατάσταση 0, αν η αρχική κατάσταση είναι η κατάσταση $i,$ $i \geq n,$ και η πολιτική P_n έχει υιοθετηθεί για τον έλεγχο της διαδικασίας. Σύμφωνα με μία γνωστή ιδιότητα των μετασχηματισμών Laplace (βλέπε π.χ. σχέση (C.2) στη σελ. 362 του βιβλίου του Tijms (1994)), έχουμε ότι:

$$C_{i0}^{(n)} = \int_0^\infty E\{c_{X(t)+i} \mid X(0)=0\} e^{-\mu t} dt + k\mu^{-1}.$$

Έστω $p_{0i}(t),$ $i \geq 0,$ η πιθανότητα το μέγεθος του πληθυσμού $X(t)$ τη χρονική στιγμή $t,$ για τη (μη ελεγχόμενη) σύνθετη διαδικασία μετανάστευσης να είναι ίσο με $i,$ αν $X(0)=0.$ Η παραπάνω έκφραση παίρνει την ακόλουθη μορφή:

$$C_{i0}^{(n)} = \int_0^\infty \left\{ \sum_{j=0}^\infty p_{0j}(t) c_{j+i} \right\} e^{-\mu t} dt + k\mu^{-1}.$$

Εναλλάσσοντας τη σειρά του ολοκληρώματος και του αθροίσματος στο δεξιό μέλος της παραπάνω εξίσωσης, έχουμε ότι:

$$C_{i0}^{(n)} = \sum_{j=0}^\infty p_{0j}^* c_{j+i} + k\mu^{-1}, \tag{6.32}$$

όπου, $p_{0j}^* = \int_0^\infty p_{0j}(t) e^{-\mu t} dt,$ είναι ο μετασχηματισμός Laplace της πιθανότητας $p_{0j}(t)$ στο σημείο $\mu.$ Έστω $N(t)$ ο αριθμός των αφίξεων των παρασιτικών ομάδων μέχρι τη χρονική στιγμή $t \geq 0,$ και έστω

$$\phi_n(j) = P[X(t) = j \mid N(t) = n,\ X(0) = 0], \quad j \geq 0, \quad n \geq 1.$$

Δεσμευόμαστε ως προς την τυχαία μεταβλητή $N(t)$ και υπολογίζουμε τις πιθανότητες $p_{0j}(t)$ από την ακόλουθη εξίσωση:

$$p_{0j}(t) = e^{-\lambda t} I(j=0) + \sum_{n=1}^{\infty} \phi_n(j) e^{-\lambda t} \frac{(\lambda t)^n}{n!},$$

όπου, $I(j=0)=0,$ αν $j \neq 0,$ και $I(j=0)=1,$ αν $j=0.$ Ολοκληρώνοντας την παραπάνω εξίσωση παίρνουμε την ακόλουθη έκφραση για την ποσότητα p_{0j}^* μετά από κάποιες απλοποιήσεις:

$$p_{0j}^* = \frac{I(j=0)}{\lambda+\mu} + \sum_{n=1}^{\infty} \phi_n(j) \frac{\lambda^n}{(\lambda+\mu)^{n+1}}. \tag{6.33}$$

Οι υπό-συνθήκη πιθανότητες $\phi_n(j)$ μπορούν να υπολογιστούν από τον παρακάτω αναδρομικό τύπο:

$$\phi_n(j) = \sum_{\ell=1}^{j-1} \phi_{n-1}(\ell) g_{j-\ell}, \quad j \geq n, \quad n \geq 2,$$

όπου, $\phi_n(j)=0,$ $j<n,$ $n \geq 1,$ και $\phi_1(j) = g_j,$ $j \geq 1.$ Είναι προφανές ότι τα αναμενόμενα κόστη $C_{i0}^{(n)},$ $i \geq n,$ μπορούν να υπολογιστούν από τις εξισώσεις (6.32) και (6.33). Δεσμευόμαστε ως προς την πρώτη μετάβαση της διαδικασίας, υπό τον έλεγχο της πολιτικής P_n, αν η αρχική κατάσταση είναι η κατάσταση $i,$ $0 \leq i \leq n-1,$ και έχουμε ότι:

$$C_{i0}^{(n)} = \lambda^{-1} c_i + \sum_{j=1}^{n-i-1} g_j C_{i+j,0}^{(n)} + \sum_{j=n-i}^{\infty} g_j C_{i+j,0}^{(n)}, \quad 0 \leq i \leq n-1. \tag{6.34}$$

Το μέσο κόστος G_n, υπό τον έλεγχο της πολιτικής P_n, υπολογίζεται από την εξίσωση $G_n = C_{00}^{(n)} / T_{00}^{(n)}$.

Θεωρούμε την ποσότητα $Q_n(i;a)$ η οποία ορίζεται ως εξής:

$$Q_n(i;a) = C(i,a) - G_n T(i,a) + \sum_{j=0}^{\infty} p_{ij}(a) h_j^{(n)}, \quad i \geq 0, \quad a \in \{0,1\},$$

όπου, $p_{ij}(a)$ είναι η πιθανότητα η επόμενη κατάσταση της διαδικασίας να είναι η κατάσταση j, αν η παρούσα κατάσταση είναι η κατάσταση i και επιλέγεται η ενέργεια a και $T(i,a)$, $C(i,a)$ είναι ο αναμενόμενος χρόνος και το αναμενόμενο κόστος, αντίστοιχα, για τη μετάβαση της διαδικασίας στην επόμενη κατάσταση, αν η αρχική κατάσταση είναι η κατάσταση i και επιλέγεται η ενέργεια a. Αν σε μία κατάσταση της διαδικασίας, επιλέγεται η ενέργεια η οποία θέτει σε λειτουργία το μηχανισμό της καταστροφής, τότε $a = 1$, ενώ αν δεν επιλέγεται αυτή η ενέργεια, τότε $a = 0$.

Υποθέτουμε ότι υπάρχει ένας ακέραιος αριθμός $\tilde{n}$ τέτοιος ώστε $1 \leq \tilde{n} < n$ και $Q_n(i;1) < h_i^{(n)}$, $\tilde{n} \leq i < n$. Σύμφωνα με το Θεώρημα 2.7 για τις ημι-Μαρκοβιανές διαδικασίες αποφάσεων (βλέπε (ε) στη σελ. 28 της παρούσας διατριβής), το μακροπρόθεσμο αναμενόμενο μέσο κόστος της μονότονης πολιτικής $P_{\tilde{n}}$, η οποία χαρακτηρίζεται από το κρίσιμο σημείο $\tilde{n}$, είναι μικρότερο από το μακροπρόθεσμο αναμενόμενο μέσο κόστος της μονότονης πολιτικής P_n. Παρόμοια, αν υπάρχει ένας ακέραιος αριθμός $\tilde{n}$ τέτοιος ώστε $n < \tilde{n}$ και $Q_n(i;0) < h_i^{(n)}$, $n \leq i < \tilde{n}$, σύμφωνα πάλι με το Θεώρημα 2.7 για τις ημι-Μαρκοβιανές διαδικασίες αποφάσεων, το μακροπρόθεσμο αναμενόμενο μέσο κόστος της μονότονης πολιτικής $P_{\tilde{n}}$, η οποία χαρακτηρίζεται από το κρίσιμο σημείο $\tilde{n}$, είναι μικρότερο από το μακροπρόθεσμο αναμενόμενο μέσο κόστος της μονότονης πολιτικής P_n. Σύμφωνα με την Πρόταση 6.4, όταν δεν είναι δυνατόν να βρεθεί ένας ακέραιος αριθμός $\tilde{n}$ με τις παραπάνω ιδιότητες, το μέσο κόστος G_n είναι η ελάχιστη τιμή του συνόλου των τιμών $\{G_i, i \geq 1\}$.

Οι παραπάνω παρατηρήσεις μας οδηγούν να σχεδιάσουμε τον ακόλουθο αλγόριθμο βελτίωσης των πολιτικών ο οποίος παράγει μία ακολουθία βελτιωμένων μονότονων πολιτικών.

Αλγόριθμος

Βήμα 1. *(Έναρξη).* Αρχικά επιλέγουμε έναν ακέραιο αριθμό $n,\ n \geq 1,$ ο οποίος χαρακτηρίζει τη μονότονη μορφή μιας πολιτικής P_n.

Βήμα 2. *(Εύρεση μέσου κόστους).* Για τη μονότονη πολιτική P_n υπολογίζουμε το μέσο κόστος G_n από την εξίσωση $G_n = C_{00}^{(n)} / T_{00}^{(n)}$.

Βήμα 3. *(Βελτίωση των πολιτικών).*

(3α) Βρίσκουμε, αν υπάρχει, το μικρότερο ακέραιο αριθμό $\tilde{n}$ τέτοιο ώστε $1 \leq \tilde{n} < n$ και $Q_n(i;1) < h_i^{(n)}$ για κάθε $i \in \{\tilde{n}, \tilde{n}+1, ..., n-1\}$ και επιστρέφουμε στο Βήμα 2 αντικαθιστώντας τον αριθμό n με τον αριθμό $\tilde{n}$. Διαφορετικά,

(3β) βρίσκουμε, αν υπάρχει, το μεγαλύτερο ακέραιο αριθμό $\tilde{n}$ τέτοιο ώστε $n < \tilde{n}$ και $Q_n(i;0) < h_i^{(n)}$ για κάθε $i \in \{n, n+1, ..., \tilde{n}-1\}$ και επιστρέφουμε στο Βήμα 2 αντικαθιστώντας τον αριθμό n με τον αριθμό $\tilde{n}$.

Βήμα 4. *(Κριτήριο Σύγκλισης).* Αν δεν είναι δυνατόν να βρεθεί ένας ακέραιος αριθμός $\tilde{n}$ ο οποίος να ικανοποιεί τις ιδιότητες του Βήματος (3α) ή του Βήματος (3β), ο αλγόριθμος σταματά. Η τελική πολιτική που παράγει ο αλγόριθμος είναι η πολιτική P_n με μακροπρόθεσμο αναμενόμενο μέσο κόστος G_n ανά μονάδα χρόνου.

Η ποσότητα $Q_n(i;1)$ στο Βήμα (3α) δίνεται από την εξίσωση:

$$Q_n(i;1) = \frac{c_i + k - G_n + \lambda \sum_{j=1}^{\infty} g_j h_{i+j}^{(n)} + \mu h_0^{(n)}}{\lambda + \mu},$$

και η ποσότητα $Q_n(i;0)$ στο Βήμα (3β) δίνεται από την εξίσωση:

$$Q_n(i;0) = \frac{c_i - G_n}{\lambda} + \sum_{j=1}^{\infty} g_j h_{i+j}^{(n)}.$$

Οι σχετικές τιμές $h_i^{(n)},\ i \geq 0,$ μπορούν να υπολογιστούν από την εξίσωση $h_i^{(n)} = C_{i0}^{(n)} - G_n T_{i0}^{(n)},$ όπου οι αναμενόμενοι χρόνοι $T_{i0}^{(n)}$ και τα αναμενόμενα κόστη $C_{i0}^{(n)}$ μπορούν να βρεθούν από τις εξισώσεις (6.11), (6.30), (6.31), (6.32) και (6.34).

Πολλά αριθμητικά παραδείγματα μας οδηγούν στο συμπέρασμα ότι ο αλγόριθμος επιτυγχάνει σημαντική βελτίωση της αρχικής πολιτικής στην πρώτη επανάληψη. Για κάθε αρχική μονότονη πολιτική $P_n,\ n \geq 1,$ τα μέσα κόστη των πολιτικών που δίνει ο αλγόριθμος μετά από την πρώτη επανάληψη δεν διαφέρουν σημαντικά από το μέσο κόστος της τελικής βέλτιστης πολιτικής. Επιπλέον υπάρχει ισχυρή ένδειξη ότι ο αριθμός των επαναλήψεων του αλγορίθμου δεν επηρεάζεται ιδιαίτερα από την αρχική κρίσιμη τιμή η οποία χαρακτηρίζει την αρχική μονότονη πολιτική κατά την έναρξη του αλγορίθμου στο Βήμα 1.

Δύο αριθμητικά παραδείγματα παρουσιάζονται παρακάτω. Στο πρώτο παράδειγμα θεωρούμε ότι $\lambda = 2,\ \ \mu = 5$ και ότι η συνάρτηση πιθανότητας του πληθυσμιακού μεγέθους των ομαδικών αυξήσεων είναι ίση με $(g_1, g_2, g_3) = (0.6, 0.2, 0.2).$ Θεωρούμε επίσης ότι $c_i = \sqrt{i},\ \ i \geq 0.$ Η κρίσιμη τιμή $n = 20$ επιλέγεται ως η αρχική κρίσιμη τιμή κατά την έναρξη του αλγορίθμου στο Βήμα 1. Στον Πίνακα 6.1 παρουσιάζονται η κρίσιμη τιμή $n*$ της βέλτιστης πολιτικής που δίνει ο αλγόριθμος και το αντίστοιχο ελάχιστο μέσο κόστος $G_{n*},$ για διαφορετικές τιμές της παραμέτρου k. Παρουσιάζονται επίσης ο απαιτούμενος αριθμός των επαναλήψεων και ο απαιτούμενος χρόνος (σε δευτερόλεπτα) για την εκτέλεση του αλγορίθμου. Στις παρενθέσεις παρουσιάζονται ο απαιτούμενος αριθμός των επαναλήψεων και ο απαιτούμενος χρόνος για την εκτέλεση της μεθόδου της διχοτόμησης.

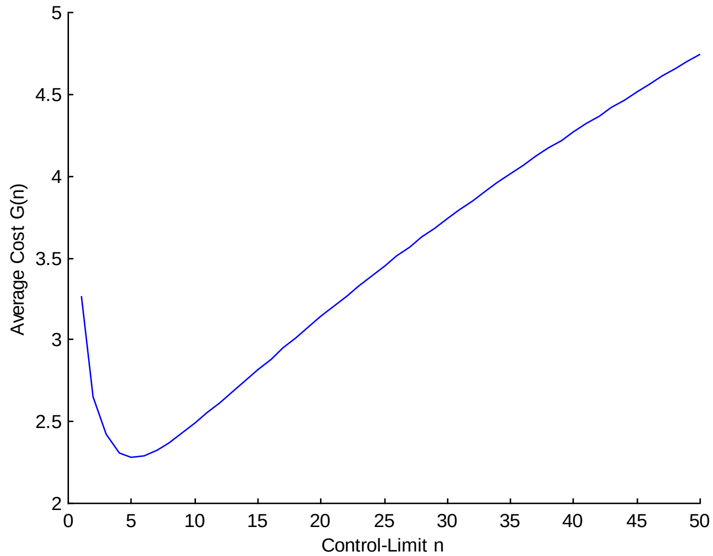
Πίνακας 6.1. Η κρίσιμη τιμή $n*$ της βέλτιστης πολιτικής, το μέσο κόστος $G_{n*},$ ο αριθμός των επαναλήψεων και οι χρόνοι εκτέλεσης (σε δευτερόλεπτα) για διαφορετικές τιμές της παραμέτρου k

k	Κρίσιμη τιμή $n*$	Μέσο κόστος G_{n^*}	Αριθμός επαναλήψεων	Χρόνοι εκτέλεσης
10	5	2.28	4 (5)	2 (3.9)
20	9	3.06	4 (6)	2.16 (5.3)
50	19	4.35	2 (6)	1.08 (6.4)
100	31	5.58	4 (6)	2.5 (5.9)

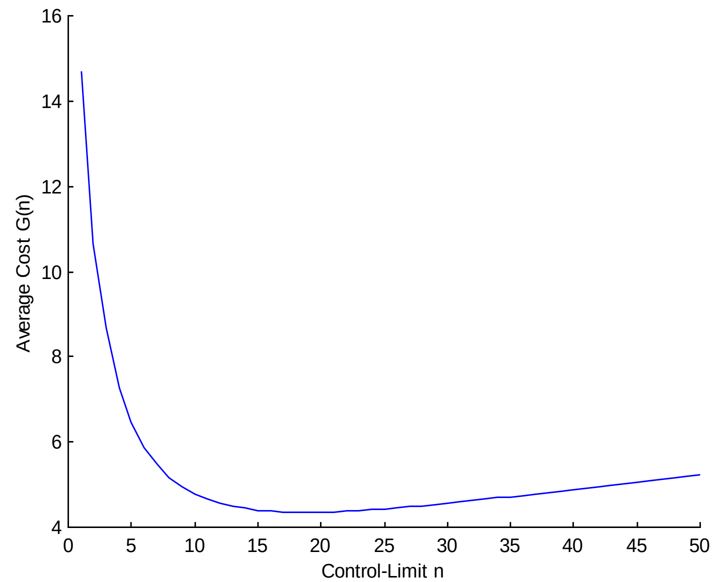
Παρατηρούμε ότι η αύξηση της τιμής της παραμέτρου k έχει ως επακόλουθο την αύξηση της κρίσιμης τιμής n^* της βέλτιστης πολιτικής. Παρατηρούμε επίσης ότι σε όλες τις περιπτώσεις ο απαιτούμενος αριθμός των επαναλήψεων και ο απαιτούμενος χρόνος για την εκτέλεση του αλγορίθμου είναι μικρότερος από τον απαιτούμενο αριθμό των επαναλήψεων και τον απαιτούμενο χρόνο για την εκτέλεση της μεθόδου της διχοτόμησης. Στο δεύτερο παράδειγμα θεωρούμε ότι $\lambda = 10,\ \ \mu = 8$ και ότι η συνάρτηση πιθανότητας του πληθυσμιακού μεγέθους των ομαδικών αυξήσεων είναι ίση με $(g_1, g_2, g_3, g_4, g_5) = (0.2, 0.2, 0.2, 0.2, 0.2)$ (βλέπε Σενάριο 4 της εργασίας του Economou (2003)). Θεωρούμε επίσης ότι $c_i = 0.5i,\ \ i \geq 0$ και $k = 30.$ Στον Πίνακα 6.2 παρουσιάζονται οι διαδοχικές κρίσιμες τιμές και τα αντίστοιχα μέσα κόστη των διαδοχικών μονότονων πολιτικών που δίνει ο αλγόριθμος για τις διαφορετικές αρχικές κρίσιμες τιμές οι οποίες επιλέγονται κατά την έναρξη του αλγορίθμου στο Βήμα 1.

Πίνακας 6.2. Οι διαδοχικές κρίσιμες τιμές και τα μέσα κόστη των διαδοχικών μονότονων πολιτικών που δίνει ο αλγόριθμος για διαφορετικές αρχικές κρίσιμες τιμές

Αρχική κρίσιμη τιμή	Διαδοχικές κρίσιμες τιμές	Διαδοχικά μέσα κόστη
1	1, 34, 12, 17	18.54, 11.92, 10.36, 10.03
20	20, 16, 17	10.13, 10.04, 10.03
50	50, 9, 19, 17	15.02, 11.12, 10.08, 10.03
70	70, 6, 22, 16, 17	19.44, 12.57, 10.27, 10.04, 10.03
90	90, 5, 24, 15, 17	23.83, 13.59, 10.47, 10.07, 10.03



Σχήμα 6.1. *Γράφημα μέσου κόστους* G_n *έναντι κρίσιμης τιμής* n *για το πρώτο παράδειγμα όταν* $k = 10$.



Σχήμα 6.2. *Γράφημα μέσου κόστους* G_n *έναντι κρίσιμης τιμής* n *για το πρώτο παράδειγμα όταν* $k = 50$.

Στα Σχήματα 6.1 και 6.2 παρουσιάζονται τα γραφήματα του μέσου κόστους G_n έναντι της κρίσιμης τιμής n, για το πρώτο παράδειγμα, όταν η παράμετρος k είναι ίση με 10 και 50, αντίστοιχα. Τα γραφήματα επιβεβαιώνουν το αποτέλεσμα της Πρότασης 6.4.

6.5 Διωνυμικές καταστροφές

Στο παρόν εδάφιο θεωρούμε ένα γενικότερο μοντέλο στο οποίο επιλέγεται μία ενέργεια που θέτει σε λειτουργία ένα μηχανισμό που προξενεί μία διωνυμική καταστροφή αντί μιας ολοκληρωτικής καταστροφής του πληθυσμού των παρασίτων. Οι Brockwell et al. (1982) θεώρησαν διάφορα Μαρκοβιανά μοντέλα για την ανάπτυξη πληθυσμών που υπόκεινται σε καταστροφές. Σε ένα από τα μοντέλα που μελέτησαν, υπέθεσαν ότι οι καταστροφές είναι διωνυμικές. Στην περίπτωση των διωνυμικών καταστροφών, οι μεταβάσεις της διαδικασίας, υπό τον έλεγχο της στάσιμης πολιτικής $f \equiv \{f_i\}$, $i \geq 0$, σε ένα χρονικό διάστημα $(t, t+\delta t)$, καθώς $\delta t \to 0$, είναι οι εξής:

$$\begin{array}{lll} i \to i+j & \text{με πιθανότητα} & \lambda g_j \delta t + o(\delta t), \quad i \geq 0, \quad j \geq 1, \\ i \to j & \text{με πιθανότητα} & \binom{i}{j} p^j (1-p)^{i-j} f_i \delta t + o(\delta t), \quad i \geq 1, \quad 0 \leq j \leq i. \end{array}$$

Η συνάρτηση $o(\cdot)$ είναι τέτοια ώστε $o(h)/h \to 0$, καθώς $h \to 0$. Ο αριθμός p, $0 \leq p < 1$, αναπαριστά την πιθανότητα ένα παράσιτο να επιβιώσει κατά τη διάρκεια μιας διωνυμικής καταστροφής. Αν $p = 0$, το μοντέλο συμπίπτει με αυτό που περιγράψαμε στο Εδάφιο 6.1. Στην περίπτωση των διωνυμικών καταστροφών, όπως στην περίπτωση των ολοκληρωτικών καταστροφών, φαίνεται διαισθητικά λογικό ότι η βέλτιστη πολιτική είναι επίσης μονότονη. Μία αναλυτική απόδειξη αυτού του αποτελέσματος φαίνεται δύσκολη.

Η βέλτιστη πολιτική μπορεί να υπολογιστεί με την εφαρμογή της προσεγγιστικής μεθόδου της Sennott (βλέπε Εδάφιο 2.6 της παρούσας διατριβής και Κεφάλαιο 10 του βιβλίου της Sennott (1999)). Βάσει αυτής της μεθόδου, μία Μαρκοβιανή διαδικασία αποφάσεων με άπειρο χώρο καταστάσεων προσεγγίζεται από μία ακολουθία Μαρκοβιανών διαδικασιών αποφάσεων με πεπερασμένο χώρο καταστάσεων. Ο αλγόριθμος των διαδοχικών προσεγγίσεων χρησιμοποιείται για τον υπολογισμό της βέλτιστης πολιτικής κάθε Μαρκοβιανής διαδικασίας αποφάσεων της ακολουθίας. Πρέπει να εισαχθούν κατάλληλες συνθήκες οι οποίες εγγυώνται τη σύγκλιση του αλγορίθμου των διαδοχικών προσεγγίσεων σε μία βέλτιστη πολιτική της αρχικής Μαρκοβιανής διαδικασίας αποφάσεων με τον άπειρο χώρο καταστάσεων.

Η προσεγγιστική μέθοδος της Sennott (1997) εφαρμόστηκε στο τροποποιημένο μοντέλο των διωνυμικών καταστροφών. Πολλά αριθμητικά παραδείγματα παρέχουν ισχυρή ένδειξη ότι η βέλτιστη πολιτική είναι μονότονη.

Ένα αριθμητικό παράδειγμα παρουσιάζεται παρακάτω. Θεωρούμε ότι $\lambda = 3,\ \ p = 0.3$ και ότι η συνάρτηση πιθανότητας του πληθυσμιακού μεγέθους των ομαδικών αυξήσεων είναι ίση με $(g_1, g_2, g_3) = (0.4, 0.2, 0.4)$. Θεωρούμε επίσης ότι $c_i = i,\ \ i \geq 0$ και $k = 5$. Η κρίσιμη τιμή $n*$ της βέλτιστης πολιτικής βρέθηκε ίση με 4. Το μέσος κόστος G της βέλτιστης πολιτικής βρέθηκε ίσο με 12.7 και ο απαιτούμενος αριθμός των επαναλήψεων για την εκτέλεση του αλγορίθμου ήταν ίσος με 61.

ΚΕΦΑΛΑΙΟ 7

Υπολογισμός της βέλτιστης πολιτικής για τον έλεγχο μιας απλής διαδικασίας μετανάστευσης με την εισαγωγή ενός αρπακτικού

7.1 Εισαγωγή

Θεωρούμε έναν πληθυσμό παρασίτων ο οποίος αναπτύσσεται στοχαστικά σε ένα βιότοπο σύμφωνα με μία απλή διαδικασία Poisson με ρυθμό ίσο με $\nu > 0.$ Υποθέτουμε ότι το κόστος της βλάβης που προξενούν τα παράσιτα είναι ίσο με $c_i,\ \ i \geq 0,$ για κάθε μονάδα χρόνου κατά την οποία το μέγεθος του πληθυσμού των παρασίτων είναι ίσο με $i.$ Είναι λογικό να υποθέσουμε ότι η ακολουθία $\{c_i\}$ είναι αύξουσα ως προς i και ότι $c_0 = 0.$ Επιπλέον θεωρούμε ότι $c_i \to \infty,$ καθώς $i \to \infty$ και ότι $c_i \leq Ai^m,\ \ i \geq 1,$ για κάποιο θετικό αριθμό A και κάποιο θετικό ακέραιο αριθμό $m.$

Θεωρούμε ότι η ανάπτυξη του πληθυσμού των παρασίτων μπορεί να ελεγχθεί μέσω μιας ενέργειας που εισάγει ένα αρπακτικό στο βιότοπο μετά από τυχαίο χρόνο ο οποίος είναι εκθετικά κατανεμημένος. Η παρουσία του αρπακτικού στο βιότοπο διακόπτει αμέσως τις μεταναστεύσεις των παρασίτων. Όταν το αρπακτικό εισαχθεί στο βιότοπο, εξοντώνει ένα-ένα τα παράσιτα με σταθερό ρυθμό ίσο με $\sigma > 0,$ μέχρι να μηδενιστεί ο πληθυσμός τους. Υποθέτουμε ότι το αρπακτικό μπορεί να αποδημήσει από το βιότοπο με ρυθμό ίσο με $\theta > 0,$ προτού εξοντώσει όλα τα παράσιτα. Η μονάδα του χρόνου επιλέγεται έτσι ώστε ο ρυθμός με τον οποίον το αρπακτικό εισάγεται στο βιότοπο να είναι ίσος με τη μονάδα. Συνεπώς, όταν επιλέγεται η ενέργεια που εισάγει το αρπακτικό στο βιότοπο, ο απαιτούμενος χρόνος για την εισαγωγή του αρπακτικού είναι μία τυχαία μεταβλητή η οποία ακολουθεί την Εκθετική κατανομή με μέση τιμή ίση με τη μονάδα. Υποθέτουμε επίσης ότι η εισαγωγή του αρπακτικού στον πληθυσμό των παρασίτων επιφέρει ένα κόστος ίσο με $k > 0,$ ανά μονάδα χρόνου και ότι $\sigma > \nu\theta.$

Έστω i και i' οι καταστάσεις της διαδικασίας στις οποίες το μέγεθος του πληθυσμού των παρασίτων είναι ίσο με $i,\ \ i \geq 0,$ και το αρπακτικό απουσιάζει από το βιότοπό τους ή είναι παρόν, αντίστοιχα. Μία στάσιμη πολιτική f ορίζεται από μία ακολουθία $\{f_i\},\ \ i \geq 0,$ όπου f_i είναι η ενέργεια που επιλέγεται όταν η διαδικασία βρίσκεται στην κατάσταση $i.$ Υποθέτουμε ότι $f_i = 1,$ όταν επιλέγεται η ενέργεια η οποία εισάγει το αρπακτικό στο βιότοπο και $f_i = 0,$ όταν δεν επιλέγεται αυτή η ενέργεια. Αν η στάσιμη πολιτική $f \equiv \{f_i\}, \ i \geq 0,$ υιοθετηθεί για τον έλεγχο της διαδικασίας, η ανάπτυξη του πληθυσμού των παρασίτων μπορεί να περιγραφεί από ένα Μαρκοβιανό μοντέλο σε συνεχή χρόνο με χώρο καταστάσεων $S = \{0, 0', 1, 1', ...\}$ και με τις ακόλουθες μεταβάσεις σε ένα χρονικό διάστημα $(t, t+\delta t),$ καθώς $\delta t \to 0$:

$$
\begin{array}{lll}
i \to i+1 & \text{με πιθανότητα} & \nu\delta t + o(\delta t), \quad i \geq 0, \\
i \to i' & \text{με πιθανότητα} & f_i\delta t + o(\delta t), \quad i \geq 0, \\
i' \to (i-1)' & \text{με πιθανότητα} & \sigma\delta t + o(\delta t), \quad i \geq 1, \\
i' \to i & \text{με πιθανότητα} & \theta\delta t + o(\delta t), \quad i \geq 0.
\end{array}
$$

Η συνάρτηση $o(\cdot)$ είναι τέτοια ώστε $o(h)/h \to 0,$ καθώς $h \to 0.$ Το πρόβλημα είναι η εύρεση της πολιτικής, η οποία, ανάμεσα σε όλες τις στάσιμες πολιτικές, ελαχιστοποιεί το μακροπρόθεσμο αναμενόμενο μέσο κόστος ανά μονάδα χρόνου. Οι χρονικές στιγμές που επιλέγεται μία ενέργεια είναι οι χρονικές στιγμές κατά τις οποίες ένα παράσιτο μεταναστεύει στο βιότοπο και οι χρονικές στιγμές κατά τις οποίες το αρπακτικό αποδημεί από το βιότοπο. Ο Kyriakidis (2003) μελέτησε το πρόβλημα στην περίπτωση κατά την οποία το αρπακτικό έχει τη δυνατότητα να αποδημήσει από το βιότοπο μόνο αφού έχει εξοντώσει όλα τα παράσιτα.

Όπως αναφέρεται στην εργασία του Kyriakidis (2003) μπορεί να θεωρηθεί ότι ο πληθυσμός των παρασίτων αποτελείται από γλάρους οι οποίοι ενδεχομένως να γίνουν επικίνδυνοι για ένα αεροπλάνο που απογειώνεται ή προσγειώνεται σε ένα αεροδρόμιο, το οποίο βρίσκεται κοντά σε μία παραθαλάσσια περιοχή. Το αρπακτικό μπορεί να είναι ένα γεράκι κατάλληλα εκπαιδευμένο ώστε να κυνηγήσει και να απομακρύνει τους γλάρους από την παραθαλάσσια περιοχή.

Στο παρόν πρόβλημα φαίνεται διαισθητικά λογικό ότι η βέλτιστη πολιτική ανήκει στην κατηγορία των μονότονων πολιτικών $\{P_n, n = 1, 2, ...\},$ όπου P_n είναι η στάσιμη πολιτική, υπό τον έλεγχο της οποίας, επιλέγεται η ενέργεια που εισάγει το αρπακτικό στο βιότοπο αν και μόνο αν το μέγεθος του πληθυσμού των παρασίτων είναι μεγαλύτερο ή ίσο με $n.$

Μία μέθοδος η οποία, σε πολλές περιπτώσεις, μας οδηγεί να αποδείξουμε ότι μία μονότονη πολιτική είναι βέλτιστη είναι η μέθοδος των διαδοχικών προσεγγίσεων (βλέπε π.χ. Derman (1970), Ross (1992)). Για μία περιληπτική αναφορά στη μέθοδο των διαδοχικών προσεγγίσεων παραπέμπουμε στη σελ. 125 της παρούσας διατριβής.

Οι Federgruen and So (1989) επινόησαν μία διαφορετική τεχνική για να αποδείξουν ότι μία μονότονη πολιτική είναι βέλτιστη. Η τεχνική βασίζεται σε μία παραμετρική ανάλυση. Αρχικά, αποδεικνύεται ότι υπάρχει μία βέλτιστη μονότονη πολιτική, όταν μία παράμετρος του μοντέλου, ενδεχομένως πλασματική, παίρνει αρκετά μικρές τιμές. Στη συνέχεια, αυτός ο ισχυρισμός αποδεικνύεται με επαγωγή για τιμές της παραμέτρου που ανήκουν σε κατάλληλα διαστήματα των πραγματικών αριθμών. Η τεχνική των Federgruen and So εφαρμόστηκε στο Εδάφιο 6.3 της παρούσας διατριβής και στην εργασία του Kyriakidis (2003). Στο Μαρκοβιανό μοντέλο αποφάσεων που μελετήθηκε από τον Kyriakidis, ως πλασματική παράμετρος θεωρήθηκε το κόστος για κάθε μονάδα χρόνου κατά την οποία η διαδικασία βρίσκεται στην κατάσταση $0'$.

Στο παρόν μοντέλο φαίνεται δύσκολο να αποδείξουμε ότι μία μονότονη πολιτική είναι βέλτιστη χρησιμοποιώντας είτε τη μέθοδο των διαδοχικών προσεγγίσεων είτε την τεχνική των Federgruen and So (1989). Είναι ωστόσο δυνατόν να υπολογίσουμε αριθμητικά τη βέλτιστη πολιτική. Αυτό μπορεί να επιτευχθεί, αν μετατρέψουμε το αρχικό Μαρκοβιανό μοντέλο αποφάσεων σε συνεχή χρόνο σε ένα ισοδύναμο Μαρκοβιανό μοντέλο αποφάσεων σε διακριτό χρόνο και στη συνέχεια εφαρμόσουμε την προσεγγιστική μέθοδο της Sennott (1997) στο μοντέλο σε διακριτό χρόνο. Η προσεγγιστική μέθοδος της Sennott παρουσιάζεται με συνοπτικό τρόπο στο Εδάφιο 2.6 της παρούσας διατριβής. Σύμφωνα με τη μέθοδο αυτή ορίζεται μία κατάλληλη ακολουθία Μαρκοβιανών μοντέλων αποφάσεων με πεπερασμένους χώρους καταστάσεων τέτοια ώστε, τα ελάχιστα μέσα κόστη και οι βέλτιστες πολιτικές των μοντέλων της ακολουθίας να συγκλίνουν στο ελάχιστο μέσο κόστος και σε μία βέλτιστη πολιτική του αρχικού μοντέλου. Ένα σύνολο κατάλληλων υποθέσεων εγγυάται τη σύγκλιση. Πολλά αριθμητικά παραδείγματα παρέχουν ισχυρή ένδειξη ότι η βέλτιστη πολιτική στο μοντέλο μας είναι μονότονη.

Στο επόμενο εδάφιο μετατρέπουμε το αρχικό Μαρκοβιανό μοντέλο αποφάσεων σε συνεχή χρόνο σε ένα ισοδύναμο Μαρκοβιανό μοντέλο αποφάσεων σε διακριτό χρόνο και υπολογίζουμε την πολιτική η οποία ελαχιστοποιεί το μακροπρόθεσμο αναμενόμενο μέσο κόστος ανά μονάδα χρόνου χρησιμοποιώντας την προσεγγιστική μέθοδο της Sennott (1997). Δύο αριθμητικά παραδείγματα επίσης παρουσιάζονται. Στο Εδάφιο 7.3 παρουσιάζουμε και επαληθεύουμε τις υποθέσεις που εγγυώνται τη σύγκλιση του αλγορίθμου στη βέλτιστη πολιτική.

Τα αποτελέσματα των Εδαφίων 7.2 και 7.3 περιλαμβάνονται στην εργασία των Kyriakidis and Dimitrakos (2005a).

7.2 Ο υπολογισμός της βέλτιστης πολιτικής

Μετατρέπουμε το Μαρκοβιανό μοντέλο αποφάσεων σε συνεχή χρόνο που εισάγαμε στο προηγούμενο εδάφιο σε μία ισοδύναμη Μαρκοβιανή διαδικασία αποφάσεων σε διακριτό χρόνο με τον ίδιο χώρο καταστάσεων S. Αυτό θα επιτευχθεί χρησιμοποιώντας την τεχνική της ομοιομορφοποίησης (βλέπε σελ. 30-31 της παρούσας διατριβής).

Έστω ένας θετικός αριθμός T τέτοιος ώστε $T < \min\{(\nu+1)^{-1}, (\sigma+\theta)^{-1}\}$. Έστω $p_{ij}(a)$, $a \in \{0,1\}$, η πιθανότητα η επόμενη κατάσταση της διαδικασίας σε διακριτό χρόνο να είναι η κατάσταση j, αν η παρούσα κατάσταση είναι η κατάσταση i και επιλέγεται η ενέργεια $a \in \{0,1\}$. Έστω επίσης $C(i,a)$ το αναμενόμενο κόστος ενός βήματος της διαδικασίας σε διακριτό χρόνο για τη μετάβαση στην επόμενη κατάσταση, αν η παρούσα κατάσταση είναι η κατάσταση i και επιλέγεται η ενέργεια a. Σύμφωνα με την τεχνική της ομοιομορφοποίησης, όπως αυτή περιγράφεται στις σελ. 30-31 της παρούσας διατριβής, ισχύει ότι:

$$
\begin{aligned}
&p_{i,i+1}(a) = T\nu, \quad i \ge 0, \quad a \in \{0,1\},\\
&p_{ii}(0) = 1 - T\nu, \quad i \ge 0,\\
&p_{ii'}(1) = T, \quad i \ge 0,\\
&p_{ii}(1) = 1 - T(\nu+1), \quad i \ge 0,\\
&p_{i'i}(0) = T\theta, \quad i \ge 0,\\
&p_{i',(i-1)'}(0) = T\sigma, \quad i \ge 1,\\
&p_{i'i'}(0) = 1 - T(\sigma+\theta), \quad i \ge 1,\\
&p_{0'0'}(0) = 1 - T\theta,\\
&C(i,0) = C(i',0) = c_i, \quad i \ge 0, \qquad (7.1)\\
&C(i,1) = c_i + k, \quad i \ge 0. \qquad (7.2)
\end{aligned}
$$

Υπό τον έλεγχο οποιασδήποτε στάσιμης πολιτικής, το Μαρκοβιανό μοντέλο αποφάσεων σε διακριτό χρόνο έχει το ίδιο μακροπρόθεσμο αναμενόμενο μέσο κόστος ανά μονάδα χρόνου με το αρχικό Μαρκοβιανό μοντέλο αποφάσεων σε συνεχή χρόνο (βλέπε σελ. 222 του βιβλίου του

Tijms (1994)). Συνεπώς τα δύο μοντέλα έχουν την ίδια βέλτιστη πολιτική. Χρησιμοποιώντας την προσεγγιστική μέθοδο της Sennott (1997) θεωρούμε μία ακολουθία Μαρκοβιανών μοντέλων αποφάσεων με χώρους καταστάσεων $G_N = \{0, 0', 1, 1', ..., N, N'\}$, $N \geq 1$. Οι πιθανότητες μετάβασης $\tilde{p}_{ij}(a)$, $i, j \in G_N$, $a \in \{0, 1\}$, σε ένα μοντέλο της ακολουθίας συμπίπτουν με τις πιθανότητες μετάβασης $p_{ij}(a)$ του μοντέλου με τον άπειρο χώρο καταστάσεων S, εκτός από τις ακόλουθες πιθανότητες:

$$\tilde{p}_{NN'}(0) = Tv,$$
$$\tilde{p}_{NN'}(1) = T(v+1).$$

Τα αναμενόμενα κόστη μιας μετάβασης $C(i,a)$, $i \in G_N$, $a \in \{0, 1\}$, σε ένα μοντέλο της ακολουθίας συμπίπτουν με αυτά του μοντέλου με τον άπειρο χώρο καταστάσεων. Για τις διάφορες τιμές του $N = 1, 2, ...,$ η βέλτιστη πολιτική και το ελάχιστο μέσο κόστος για κάθε μοντέλο της ακολουθίας υπολογίζονται χρησιμοποιώντας τον αλγόριθμο των διαδοχικών προσεγγίσεων (βλέπε Κεφάλαιο 3 του βιβλίου του Tijms (1994)). Από το Θεώρημα 2.4 της Sennott (1997), έπεται ότι, καθώς $N \to \infty$, τα ελάχιστα μέσα κόστη και οι βέλτιστες πολιτικές της ακολουθίας συγκλίνουν στο ελάχιστο μέσο κόστος και σε μία βέλτιστη πολιτική του μοντέλου με τον άπειρο χώρο καταστάσεων. Οι υποθέσεις που εγγυώνται τη σύγκλιση θα παρουσιαστούν και θα επαληθευτούν στο επόμενο εδάφιο. Κατά την εφαρμογή της μεθόδου, ο αριθμός N αυξάνεται μέχρι η αλλαγή στην τιμή του ελάχιστου μέσου κόστους να είναι πολύ μικρή και η μορφή της βέλτιστης πολιτικής να παραμείνει αμετάβλητη. Πολλά αριθμητικά παραδείγματα που έχουμε εξετάσει μας οδηγούν να εικάσουμε ότι η βέλτιστη πολιτική ανήκει στην κατηγορία των μονότονων πολιτικών $\{P_n, n = 1, 2, ...\}$. Αυτό σημαίνει ότι υπάρχει μία κρίσιμη τιμή $n*$ τέτοια ώστε η βέλτιστη πολιτική εισάγει το αρπακτικό στο βιότοπο αν και μόνο αν το μέγεθος του πληθυσμού των παρασίτων είναι μεγαλύτερο ή ίσο με $n*$.

Παρουσιάζονται δύο αριθμητικά παραδείγματα. Στο πρώτο παράδειγμα υποθέτουμε ότι $\sigma = 8$, $v = 3$, $\theta = 2$, $c_i = \sqrt{i}$, $i \geq 0$ και $k = 10$. Επιλέγουμε $T = 0.05$ και τον αριθμό 5×10^{-5} ως το μέγιστο αριθμό της σχετικής διαφοράς ανάμεσα στο ανώτερο και στο κατώτερο φράγμα για το ελάχιστο μέσο κόστος στον αλγόριθμο των διαδοχικών προσεγγίσεων (βλέπε σελ. 208 του βιβλίου του Tijms (1994)). Στον Πίνακα 7.1 παρουσιάζονται η βέλτιστη κρίσιμη τιμή n^* και το ελάχιστο μέσο κόστος g_{n^*} για διάφορες τιμές του αριθμού N, στο μοντέλο με χώρο καταστάσεων G_N.

Πίνακας 7.1. Οι κρίσιμες τιμές n^* και το ελάχιστο μέσο κόστος g_{n^*} για διάφορες τιμές του N

N	Κρίσιμη τιμή n^*	Μέσο κόστος g_{n^*}
70	0	8.1362
72	7	8.2506
75	6	8.4110
80	5	8.6797
85	4	8.9075
200	4	9.3674
250	4	9.3691
290	4	9.3692
300	4	9.3693
1000	4	9.3693

Παρατηρούμε ότι η κρίσιμη τιμή n^* δεν αλλάζει και το μέσο κόστος g_{n^*} δεν αλλάζει σημαντικά αν ο αριθμός N είναι μεγαλύτερος ή ίσος με 200. Συμπεραίνουμε ότι η βέλτιστη πολιτική για το αρχικό μοντέλο με τον άπειρο χώρο καταστάσεων είναι η πολιτική P_4 με μακροπρόθεσμο αναμενόμενο μέσο κόστος ανά μονάδα χρόνου ίσο με 9.3693.

Στο δεύτερο παράδειγμα υποθέτουμε ότι $\sigma = 10$, $\nu = 4$, $\theta = 2$, $c_i = \sqrt{i+1}$, $i \geq 0$ και $k = 8$. Επιλέγουμε τον ίδιο αριθμό T και τον ίδιο μέγιστο αριθμό της σχετικής διαφοράς των δύο φραγμάτων με το πρώτο παράδειγμα. Στον Πίνακα 7.2 παρουσιάζονται η βέλτιστη κρίσιμη τιμή n^* και το ελάχιστο μέσο κόστος g_{n^*} για διάφορες τιμές του αριθμού N, στο μοντέλο με χώρο καταστάσεων G_N.

Πίνακας 7.2. Οι κρίσιμες τιμές $n*$ και το ελάχιστο μέσο κόστος g_{n*} για διάφορες τιμές του N

N	Κρίσιμη τιμή $n*$	Μέσο κόστος g_{n*}
70	6	7.7870
72	5	8.2017
75	4	8.3574
85	3	8.7843
200	3	9.2746
250	3	9.2814
300	3	9.2828
700	3	9.2830
1000	3	9.2830

Παρατηρούμε ότι η κρίσιμη τιμή $n*$ δεν αλλάζει και το μέσο κόστος g_{n*} δεν αλλάζει σημαντικά αν ο αριθμός N είναι μεγαλύτερος ή ίσος με 250. Συμπεραίνουμε ότι η βέλτιστη πολιτική για το αρχικό μοντέλο με τον άπειρο χώρο καταστάσεων είναι η πολιτική P_3 με μακροπρόθεσμο αναμενόμενο μέσο κόστος ανά μονάδα χρόνου ίσο με 9.2830.

7.3 Η σύγκλιση του αλγορίθμου

Σύμφωνα με την Πρόταση 4.1 της Sennott (1997), στο πρόβλημά μας τα ελάχιστα μέσα κόστη και οι βέλτιστες πολιτικές των Μαρκοβιανών μοντέλων αποφάσεων με τους πεπερασμένους χώρους καταστάσεων G_N, $N \geq 1$, συγκλίνουν στο ελάχιστο μέσο κόστος και σε μία βέλτιστη πολιτική του μοντέλου με τον άπειρο χώρο καταστάσεων S, αν ικανοποιούνται οι δύο ακόλουθες συνθήκες.

Συνθήκη 1: Στο μοντέλο με τον άπειρο χώρο καταστάσεων S υπάρχει μία στάσιμη πολιτική $\tilde{f} \equiv \{\tilde{f}_i\}$, $i \in S$, υπό τον έλεγχο της οποίας, ορίζεται μία ανάγωγη έμμονη θετική Μαρκοβιανή αλυσίδα στο S με πεπερασμένο μακροπρόθεσμο αναμενόμενο μέσο κόστος g ανά μονάδα χρόνου για κάθε αρχική κατάσταση της διαδικασίας. Επιπλέον υπάρχει ένας αριθμός $\varepsilon > 0$

τέτοιος ώστε το σύνολο: $D = \{i \in S,$ υπάρχει ενέργεια a τέτοια ώστε $C(i,a) \leq g + \varepsilon\}$ να είναι πεπερασμένο.

Συνθήκη 2: Ισχύει ότι: $\lim_{N\to\infty} \pi_i(N) = \pi_i,\ i \in S,$ και $\lim_{N\to\infty} g_{f|N}^{N} = g_f$, όπου

$f \equiv \{f_i\},\ i \in S,$ είναι η βέλτιστη πολιτική του Μαρκοβιανού μοντέλου αποφάσεων με τον άπειρο χώρο καταστάσεων S,

$f \mid N,$ είναι η στάσιμη πολιτική του Μαρκοβιανού μοντέλου αποφάσεων με τον πεπερασμένο χώρο καταστάσεων G_N τέτοια ώστε: $(f \mid N)(i) = f_i,\ i \in G_N$,

$\pi_i,\ i \in S,$ είναι η στάσιμη κατανομή υπό τον έλεγχο της πολιτικής f,

$\pi_i(N),\ i \in G_N$, είναι η στάσιμη κατανομή υπό τον έλεγχο της πολιτικής $f \mid N,$

g_f, είναι το μακροπρόθεσμο αναμενόμενο μέσο κόστος ανά μονάδα χρόνου υπό τον έλεγχο της πολιτικής f,

$g_{f|N}^{N}$, είναι το μακροπρόθεσμο αναμενόμενο μέσο κόστος ανά μονάδα χρόνου υπό τον έλεγχο της πολιτικής $f \mid N.$

Επισημαίνουμε ότι η Συνθήκη 1 έχει ως επακόλουθο την ύπαρξη μιας βέλτιστης πολιτικής f, υπό τον έλεγχο της οποίας, η Μαρκοβιανή αλυσίδα που ορίζεται στο χώρο καταστάσεων S, είναι έμμονη θετική (βλέπε Πρόταση 4.1 της Sennott (1997)).

Για να αποδείξουμε τη Συνθήκη 1 επιλέγουμε τη στάσιμη πολιτική $\tilde{f} = P_0$. Υπό τον έλεγχο της πολιτικής P_0, οι στάσιμες πιθανότητες $\tilde{\pi}_i$ και $\tilde{\pi}_{i'}$, $i \geq 0$, της διαδικασίας με τον άπειρο χώρο καταστάσεων ικανοποιούν τις ακόλουθες εξισώσεις ισορροπίας και τη συνθήκη κανονικοποίησης:

$$\tilde{\pi}_0 = [1 - T(\nu + 1)]\tilde{\pi}_0 + T\theta\tilde{\pi}_{0'},$$
$$\tilde{\pi}_{0'} = (1 - T\theta)\tilde{\pi}_{0'} + T\tilde{\pi}_0 + T\sigma\tilde{\pi}_{1'},$$
$$\tilde{\pi}_i = [1 - T(\nu + 1)]\tilde{\pi}_i + T\nu\tilde{\pi}_{i-1} + T\theta\tilde{\pi}_{i'}, \quad i \geq 1,$$

$$\tilde{\pi}_{i'} = [1 - T(\sigma + \theta)]\tilde{\pi}_{i'} + T\tilde{\pi}_i + T\sigma\tilde{\pi}_{(i+1)'}, \quad i \geq 1,$$

$$\sum_{i=0}^{\infty} \tilde{\pi}_i + \sum_{i=0}^{\infty} \tilde{\pi}_{i'} = 1.$$

Οι παραπάνω εξισώσεις οδηγούν στις ακόλουθες εκφράσεις για τις στάσιμες πιθανότητες:

$$\tilde{\pi}_i = \left[\frac{\nu(\sigma + \theta)}{\sigma(\nu + 1)}\right]^i \frac{\theta(\sigma - \nu\theta)}{\sigma(\theta + 1)(\nu + 1)}, \quad i \geq 0,$$

$$\tilde{\pi}_{i'} = \left[\frac{\nu(\sigma + \theta)}{\sigma(\nu + 1)}\right]^{i-1} \frac{\nu\theta(\sigma - \nu\theta)}{\sigma^2(\theta + 1)(\nu + 1)}, \quad i \geq 1,$$

$$\tilde{\pi}_{0'} = \frac{\sigma - \nu\theta}{\sigma(\theta + 1)}.$$

Παρατηρούμε ότι η συνθήκη $\sigma > \nu\theta$, που εισάγαμε στο Εδάφιο 7.1, εγγυάται την ύπαρξη της στάσιμης κατανομής. Υπό τον έλεγχο της πολιτικής P_0, η διαδικασία είναι προφανώς ανάγωγη. Επιπλέον είναι έμμονη θετική διότι έχει στάσιμη κατανομή (βλέπε Θεώρημα 3 στη σελ. 205 του βιβλίου των Grimmett and Stirzaker (1992)). Για κάθε αρχική κατάσταση της διαδικασίας, το μακροπρόθεσμο αναμενόμενο μέσο κόστος g_0 ανά μονάδα χρόνου, υπό τον έλεγχο της πολιτικής P_0, δίνεται από την ακόλουθη εξίσωση (βλέπε σχέση (3.1.3) στη σελ. 186 του βιβλίου του Tijms (1994)):

$$g_0 = \sum_{i=1}^{\infty} \tilde{\pi}_i (c_i + k) + \sum_{i=1}^{\infty} \tilde{\pi}_{i'} c_i.$$

Οι παραπάνω εκφράσεις για τις στάσιμες πιθανότητες $\tilde{\pi}_i$, $\tilde{\pi}_{i'}$, $i \geq 1$, και η συνθήκη $c_i \leq Ai^m$, $i \geq 1$, που εισάγαμε στο Εδάφιο 7.1, έχουν ως επακόλουθο ότι $g_0 < \infty$. Σύμφωνα με τις σχέσεις

(7.1), (7.2) και την υπόθεση $c_i \to \infty,$ καθώς $i \to \infty,$ το σύνολο $D = \{i \in S,$ υπάρχει ενέργεια a τέτοια ώστε $C(i,a) \le g_0 + \varepsilon\},$ είναι πεπερασμένο για κάθε τιμή του $\varepsilon > 0.$

Για να αποδείξουμε τη Συνθήκη 2, παρατηρούμε ότι οι εξισώσεις ισορροπίας για τις καταστάσεις $0, 1, ..., N-1, N, 0', ..., (N-1)'$ είναι ίδιες στους χώρους καταστάσεων S και G_N. Διαφέρουν μόνο στην κατάσταση N'. Αυτό σημαίνει ότι:

$$\pi_i(N) = \frac{\pi_i}{\sum_{j \in G_N} \pi_j}, \quad i \in G_N.$$

Από την παραπάνω εξίσωση συμπεραίνουμε ότι: $\lim_{N \to \infty} \pi_i(N) = \pi_i,\ i \in S.$ Επιπλέον έχουμε ότι:

$$g^N_{f|N} = \sum_{i \in G_N} \pi_i(N) C(i, (f \mid N)(i)) = \frac{\sum_{i \in G_N} \pi_i C(i, (f \mid N)(i))}{\sum_{j \in G_N} \pi_j} \to \sum_{i \in S} \pi_i C(i, f_i) = g_f,$$

καθώς $N \to \infty.$

ΠΑΡΑΡΤΗΜΑ

Παρουσιάζουμε αποσπάσματα του προγράμματος Matlab για τον αλγόριθμο του Εδαφίου 4.5. Ο αλγόριθμος υπολογίζει τη βέλτιστη πολιτική για τον έλεγχο ενός συστήματος παραγωγής. Ο σχεδιασμός του αλγορίθμου βασίζεται στην τεχνική εμφύτευσης του Tijms (1994).

{**Ορισμοί πινάκων του προγράμματος**}

```
fold=zeros(m+1,K+1); fnew=zeros(m+1,K+1); s=zeros(1,K+1);
s1=zeros(1,K+1); p=zeros(m+2,m+2); c1=zeros(1,m+1);
c2=zeros(1,m+1); t=zeros(m+1,K+1); c=zeros(m+1,K+1);
p1=zeros(m+1,K+1,m+1,K+1); rel=zeros(m+1,K+1); z1=zeros(1,K+1);
z2=zeros(1,K+1);
```

{**Οι παράμετροι του μοντέλου**}

`m=input('Give m:');` {`m`: μέγιστος βαθμός επιδείνωσης του μηχανισμού}

`K=input('Give K:');` {`K`: μέγιστη χωρητικότητα του αποθηκευτικού χώρου}

`d=input('Give d:');` {`d`: ρυθμός της ζήτησης του ακατέργαστου υλικού}

`Cp=input('Give Cp:');` {`Cp`: ρυθμός του κόστους της προληπτικής συντήρησης}

`Cf=input('Give Cf:');` {`Cf`: ρυθμός του κόστους της επισκευής}

`a=input('Give a:');` {`a`: πιθανότητα επιτυχίας της προληπτικής συντήρησης}

`b=input('Give b:');` {`b`: πιθανότητα επιτυχίας της επισκευής}

`h=input('Give h:');` {`h`: κόστος της αποθήκευσης του ακατέργαστου υλικού}

{Υπολογισμός των πιθανοτήτων μετάβασης p_{ij}, $0 \le i,\ j \le m+1$}

```
for i=1:m+2
    for j=1:m+2
       if (i<=j)
        p(i,j)=1/(m+2-i+1);
       else
        p(i,j)=0;
       end
    end
end
```

{Τα κόστη της λειτουργίας του μηχανισμού}

{Ο αποθηκευτικός χώρος δεν είναι πλήρης}

```
for i=1:m+1
   c1(i)=0.1*i;
end
```

{Ο αποθηκευτικός χώρος είναι πλήρης}

```
for i=1:m+1
   c2(i)=0.05*i;
end
```

{Η αρχική μονότονη πολιτική του Βήματος 1}

```
for i=1:m+1
   for x=1:K+1
      fold(i,x)=0;
   end
end
```

{Έναρξη της επαναληπτικής διαδικασίας}

```
flag=1;
```

```
time0=cputime;
```

`it=0;` {`it`: αριθμός επαναλήψεων του αλγορίθμου}

```
while (flag==1)

   flag=0;
```

{**Υπολογισμός της διάστασης του γραμμικού συστήματος του Βήματος 2**}

```
    number=0;

     for x=1:K+1
        if (fold(m+1,x)==0)
           s1(x)=m+1;
        end
        if (fold(m+1,x)~=0)
           s1(x)=0;
           for i=1:m+1
               if (fold(i,x)==0)
                 s1(x)=s1(x)+1;
               end
           end
           if (s1(x)<m+1)
              s1(x)=s1(x)+1;
           end
        end
        number=number+s1(x);
     end
```

{**Υπολογισμός της κρίσιμης τιμής** $i(x)$ **για κάθε περιεχόμενο** x **του αποθηκευτικού χώρου**}

```
     flag1=1;

     for x=1:K+1
        s(x)=0;
        for i=1:m+1
           if (fold(i,x)==0)
              s(x)=s(x)+1;
           end
           if (fold(i,x)==1 & flag1==1)
              i1=i;
              flag1=0;
              s(x)=i1;
           end
        end
        flag1=1;
     end
```

{**Υπολογισμός** $T^{E}_{(i,x)}$ **για τις καταστάσεις** (i,x) **τέτοιες ώστε** $0 \le x \le K-1,\ 0 \le i < i(x)$}

```
for x=1:K
   if (fold(m+1,x+1)~=0)
      for i=1:s(x)-1
         sum1=0;
         for j=s(x+1)+1:m+1
            sum1=sum1+p(i,j);
         end
         t(i,x)=1+(1/a)*sum1+p(i,m+2)*(1/b);
      end
   end
end
```

{**Υπολογισμός** $C^{E}_{(i,x)}$ **για τις καταστάσεις** (i,x) **τέτοιες ώστε** $i(x) \le i \le m, 0 \le x \le K$}

```
for x=1:K
   if (fold(m+1,x+1)~=0)
      for i=s(x)+1:m+1
         sum2=0;
         for j=1:fix((x-1)/d)
           sum2=sum2+(1-a)^(j-1)*(2*j*(x-1)-d*j^2+d*j);
         end
c(i,x)=(1/a)*(Cp+(1-a)^(fix((x-1)/d))*(d-a*(x-1)+a*d*fix((x-1)/d)))
        +(h/2)*(1-a)^(fix((x-1)/d))*(1+fix((x-1)/d))*(2*(x-1)-d*fix((x-1)/d))
        +((h*a)/2)*sum2;
      end
   end
end


if (fold(m+1,K+1)~=0)
   for i=s(K+1)+1:m+1
      sum3=0;
      for j=1:fix(K/d)
         sum3=sum3+(1-a)^(j-1)*(2*j*K-d*j^2+d*j);
      end
c(i,K+1)=(1/a)*(Cp+(1-a)^(fix(K/d))*(d-a*K+a*d*fix(K/d)))
          +(h/2)*(1-a)^(fix(K/d))*(1+fix(K/d))*(2*K-d*fix(K/d))
          +((h*a)/2)*sum3;
   end
end
```

{**Υπολογισμός** $p^E_{(i,x)(j,K)}$, $0 \le i < i(K),\ 0 \le j \le i(K)$}

```
if (fold(m+1,K+1)~=0)
   for i=1:s(K+1)-1
      for j=1:s(K+1)
         p1(i,K+1,j,K+1)=p(i,j);
      end
   end
end
```

{**Υπολογισμός** $p^E_{(i,x)(0,0)}$, $0 \le i < i(x),\ 0 \le x \le K-1$}

{1η περίπτωση: $(x+1)/d$, ακέραιος αριθμός}

```
for x=1:K
   if (fold(m+1,x+1)~=0 & mod(x,d)==0)
      for i=1:s(x)-1
         sum4=0;
         for j=s(x+1)+1:m+1
            sum4=sum4+p(i,j);
         end
         p1(i,x,1,1)=sum4*(1-a)^((x/d)-1)+p(i,m+2)*(1-b)^((x/d)-1);
      end
   end
end
```

{2η περίπτωση: $(x+1)/d$, δεν είναι ακέραιος αριθμός}

```
for x=1:K
   if (fold(m+1,x+1)~=0 & mod(x,d)~=0)
      for i=1:s(x)-1
         sum5=0;
         for j=s(x+1)+1:m+1
            sum5=sum5+p(i,j);
         end
         p1(i,x,1,1)=sum5*(1-a)^(fix(x/d))+p(i,m+2)*(1-b)^(fix(x/d));
      end
   end
end
```

{**Κατασκευή του συστήματος των γραμμικών εξισώσεων του Βήματος 2**}

```
u=zeros(number+1,1); u1=zeros(number+1,1);

vv=zeros(number+1,number+1); v=zeros(number+1,m+1,K+1);

count1=1;

for x=1:K+1
   for i=1:s(x)
      u(count1)=c(i,x);
      count1=count1+1;
   end
end

for count2=1:number+1
   for x=1:K+1
      for i=1:s(x)
         v(count2,i,x)=0;
      end
   end
end

count3=0;

for x=1:K
   if (fold(m+1,x+1)~=0)
      for i=1:s(x)-1
         count3=count3+1;
         v(count3,i,x)=1;
         for y=1:s(x+1)
            v(count3,y,x+1)=v(count3,y,x+1)-p1(i,x,y,x+1);
         end
         if (mod(x,d)==0)
            for y=1:(x/d)-1
               v(count3,1,x+1-y*d)=v(count3,1,x+1-y*d)-p1(i,x,1,x+1-y*d);
            end
            v(count3,1,1)=v(count3,1,1)-p1(i,x,1,1);
         end
         if (mod(x,d)~=0)
            for y=1:fix(x/d)
               v(count3,1,x+1-y*d)=v(count3,1,x+1-y*d)-p1(i,x,1,x+1-y*d);
            end
            v(count3,1,1)=v(count3,1,1)-p1(i,x,1,1);
         end
      end
      count3=count3+1;
      v(count3,s(x),x)=1;
      if (mod((x-1),d)==0)
         for y=1:((x-1)/d)-1
            v(count3,1,x-y*d)=v(count3,1,x-y*d)-p1(s(x),x,1,x-y*d);
         end
         v(count3,1,1)=v(count3,1,1)-p1(s(x),x,1,1);
      end
      if (mod((x-1),d)~=0)
```

```
         for y=1:fix((x-1)/d)
            v(count3,1,x-y*d)=v(count3,1,x-y*d)-p1(s(x),x,1,x-y*d);
         end
         v(count3,1,1)=v(count3,1,1)-p1(s(x),x,1,1);
      end
   end
end

if (fold(m+1,K+1)~=0)
   for i=1:s(K+1)-1
      count3=count3+1;
      v(count3,i,K+1)=1;
      for y=1:s(K+1)
         v(count3,y,K+1)=v(count3,y,K+1)-p1(i,K+1,y,K+1);
      end
      if (mod(K,d)==0)
         for y=1:(K/d)-1
            v(count3,1,K+1-y*d)=v(count3,y,K+1-y*d)-p1(i,K+1,1,K+1-y*d);
         end
         v(count3,1,1)=v(count3,1,1)-p1(i,K+1,1,1);
      end
      if (mod(K,d)~=0)
         for y=1:fix(K/d)
            v(count3,1,K+1-y*d)=v(count3,1,K+1-y*d)-p1(i,K+1,1,K+1-y*d);
         end
         v(count3,1,1)=v(count3,1,1)-p1(i,K+1,1,1);
      end
   end
   count3=count3+1;
   v(count3,s(K+1),K+1)=1;
   if (mod(K,d)==0)
      for y=1:(K/d)-1
         v(count3,1,K+1-y*d)=v(count3,1,K+1-y*d)-p1(s(K+1),K+1,1,K+1-y*d);
      end
      v(count3,1,1)=v(count3,1,1)-p1(s(K+1),K+1,1,1);
   end
   if (mod(K,d)~=0)
      for y=1:fix(K/d)
         v(count3,1,K+1-y*d)=v(count3,1,K+1-y*d)-p1(s(K+1),K+1,1,K+1-y*d);
      end
      v(count3,1,1)=v(count3,1,1)-p1(s(K+1),K+1,1,1);
   end
end


for n=1:number
   count4=0;
   for x=1:K+1
      for i=1:s(x)
         count4=count4+1;
         vv(n,count4)=v(n,i,x);
      end
   end
end
```

```
count5=1;

for x=1:K+1
   for i=1:s(x)
      u1(count5)=t(i,x);
      count5=count5+1;
   end
end

for n=1:number
   vv(n,number+1)=u1(n);
end

vv(number+1,1)=1;
```

{**Επίλυση του γραμμικού συστήματος του Βήματος 2**}

`lisi=inv(vv)*u;` {`lisi:` λύση του γραμμικού συστήματος}

`g=lisi(number+1);` {`g:` μέσο κόστος των διαδοχικών μονότονων πολιτικών}

{`rel:` πίνακας «αποθήκευσης» των σχετικών τιμών}

```
count6=1;

for x=1:K+1
   for i=1:s(x)
      rel(i,x)=lisi(count6);
      count6=count6+1;
   end
end
```

{**Υπολογισμός των σχετικών τιμών για τις καταστάσεις (i,x) τέτοιες ώστε $i(x) < i \leq m$, $0 \leq x \leq K-1$**}

```
for x=1:K
   if (fold(m+1,x+1)~=0)
      if (mod((x-1),d)==0)
         for i=s(x)+1:m+1
            sum6=0;
            for j=1:((x-1)/d)-1
               sum6=sum6+p1(i,x,1,x-j*d)*rel(1,x-j*d);
            end
            rel(i,x)=c(i,x)-g*(1/a)+sum6+p1(i,x,1,1)*rel(1,1);
         end
      end
      if (mod((x-1),d)~=0)
         for i=s(x)+1:m+1
            sum7=0;
```

```
          for j=1:fix((x-1)/d)
             sum7=sum7+p1(i,x,1,x-j*d)*rel(1,x-j*d);
          end
          rel(i,x)=c(i,x)-g*(1/a)+sum7+p1(i,x,1,1)*rel(1,1);
       end
    end
  end
end
```

{Βήμα 3. Βελτίωση των πολιτικών}

{Βήμα 3(α). Βελτίωση των πολιτικών για τις καταστάσεις (i,x) τέτοιες ώστε $0 \le i < i(x)$, $0 \le x \le K-1$}

```
for x=1:K
    if (fold(m+1,x+1)~=0)
       for i=s(x)-1:-1:1
          if (mod((x-1),d)==0)
             for j=1:((x-1)/d)-1
                sum8=0;
                sum8=sum8+a*(1-a)^(j-1);
                z1(j)=sum8;
             end
             sum9=0;
             for j=1:((x-1)/d)-1
                sum9=sum9+z1(j)*rel(1,x-j*d);
             end
             sum10=sum9+(1-a)^(((x-1)/d)-1)*rel(1,1);
          end
          if (mod((x-1),d)~=0)
             for j=1:fix((x-1)/d)
                sum8=0;
                sum8=sum8+a*(1-a)^(j-1);
                z2(j)=sum8;
             end
             sum9=0;
             for j=1:fix((x-1)/d)
                sum9=sum9+z2(j)*rel(1,x-j*d);
             end
             sum10=sum9+(1-a)^(fix((x-1)/d))*rel(1,1);
          end
          if (rel(i,x)<c(i,x)-g*(1/a)+sum10)
             fnew(i,x)=0;
          else
             fnew(i,x)=1;
          end
      end
    end
end
```

{**Βήμα 4. Κριτήριο Σύγκλισης**}

```
   for i=1:m+1
      for x=1:K+1
         if (fnew(i,x)~=fold(i,x))
            flag=1;
         end
      end
   end


   for i=1:m+1
      for x=1:K+1
         fold(i,x)=fnew(i,x);
      end
   end

   it=it+1;

end

gmin=g;

time1=cputime;
```

{**Παρουσίαση των αποτελεσμάτων**}

{`fnew`: πίνακας «αποθήκευσης» της τελικής βέλτιστης πολιτικής}

{`gmin`: ελάχιστο μέσο κόστος}

{`time1-time0`: χρόνος εκτέλεσης του προγράμματος}

```
disp('The form of the optimal policy is:');
disp(fnew');
disp('The minimum long-run expected average cost is:');
disp(gmin);
disp('The number of iterations of the algorithm was:');
disp(it);
disp('The CPU time of the algorithm was:');
disp(time1-time0);
```

ΒΙΒΛΙΟΓΡΑΦΙΑ

ABAKUKS A. (1973) An optimal isolation policy for an epidemic, *Journal of Applied Probability* **10**, 247-262.

ABAKUKS A. (1974) Optimal immunisation policies for epidemics, *Advances in Applied Probability* **6**, 494-511.

ABAKUKS A. (1979) An optimal hunting policy for a stochastic logistic model, *Journal of Applied Probability* **16**, 319-331.

BAILEY N. T. J. (1950) A simple stochastic epidemic, *Biometrika* **37**, 193-202.

BALL F. G. and O' NEIL P. D. (1993) A modification of the general stochastic epidemic motivated by AIDS modelling, *Advances in Applied Probability* **25**, 39-62.

BARLOW R. E. and PROSCHAN F. (1965) *Mathematical Theory of Reliability*, Wiley, New York.

BATHER J. (1976) Optimal stationary policies for denumerable Markov chains in continuous time, *Advances in Applied Probability* **8**, 144-158.

BATHER J. (2000) *Decision Theory: An Introduction to Dynamic Programming and Sequential Decisions*, Wiley, New York.

BELLMAN R. E. (1957) *Dynamic Programming*, Princeton University Press, Princeton, New Jersey.

BENYAMINI Z. and YECHIALI U. (1999) Optimality of control limit maintenance policies under nonstationary deterioration, *Probability in the Engineering and Informational Sciences* **13**, 55-70.

BERG M. (1976) A proof of optimality for age replacement policies, *Journal of Applied Probability* **13**, 751-759.

BILLARD L., LACAYO H. and LANGBERG N. A. (1979) The symmetric *m*-dimensional simple epidemic process, *Journal of the Royal Statistical Society Series B* **41**, 196-202.

BLACKBURN J. D. (1972) Optimal control of a single-server queue with balking and reneging, *Management Science* **19**, 297-313.

BROCKWELL P. J., GANI J. and RESNICK S. I. (1982) Birth, immigration and catastrophe processes, *Advances in Applied Probability* **14**, 709-731.

CHEN M. and FELDMAN R. M. (1997) Optimal replacement policies with minimal repair and age-dependent costs, *European Journal of Operational Research* **98**, 75-84.

CLANCY D. (1999a) Optimal intervention for epidemic models with general infection and removal rate functions, *Journal of Mathematical Biology* **39**, 309-331.

CLANCY D. (1999b) Outcomes of epidemic models with general infection and removal rate functions at certain stopping times, *Journal of Applied Probability* **36**, 799-813.

DERMAN C. (1970) *Finite State Markovian Decision Processes*, Academic Press, New York.

DIMITRAKOS T. D. and KYRIAKIDIS E. G. (2005a) An algorithm for the computation of the optimal repair/replacement policy under general repairs, Submitted for publication.

DIMITRAKOS T. D. and KYRIAKIDIS E. G. (2005b) A semi-Markov decision algorithm for the optimal preventive maintenance of a production system with buffer capacity and continuous repair times, Submitted for publication.

DOUER N. and YECHIALI U. (1994) Optimal repair and replacement in Markovian systems, *Stochastic Models* **10**, 253-270.

ECONOMOU A. (2003) On the control of a compound immigration process through total catastrophes, *European Journal of Operational Research* **147**, 522-529.

FEDERGRUEN A. and SO K. C. (1989) Optimal time to repair a broken server, *Advances in Applied Probability* **21**, 376-397.

FEDERGRUEN A. and SO K. C. (1990) Optimal maintenance policies for single-server queueing systems subject to breakdowns, *Operations Research* **38**, 330-343.

FEDERGRUEN A. and SO K. C. (1991) Optimality of threshold policies in single-server queueing systems with server vacations, *Advances in Applied Probability* **23**, 388-405.

GLEISSNER W. (1988) The spread of epidemics, *Applied Mathematical Computing* **27**, 167-171.

GRIMMETT G. R. and STIRZAKER D. R. (1992) *Probability and Random Processes*, 2nd edition, Oxford University Press, Oxford.

HEYMAN D. P. and SOBEL M. J. (2004) *Stochastic Models in Operations Research*, vol. 2, Dover, New York.

HOWARD R. A. (1960) *Dynamic Programming and Markov Processes*, Wiley, New York.

KAWAI H. (1983) An optimal ordering and replacement policy of a Markovian degradation system under complete observation-Part I, *Journal of the Operations Research Society of Japan* **26**, 279-290.

KIJIMA M., MORIMURA H. and SUZUKI Y. (1988) Periodical replacement problem without assuming minimal repair, *European Journal of Operational Research* **37**, 194-203.

KYRIAKIDIS E. G. (1993) A Markov decision algorithm for optimal pest control through uniform catastrophes, *European Journal of Operational Research* **64**, 38-44.

KYRIAKIDIS E. G. (1995) Optimal control of two competing diseases or species, *The Mathematical Scientist* **20**, 56-66.

KYRIAKIDIS E. G. (1999a) Characterization of the optimal policy for the control of a simple immigration process through total catastrophes, *Operations Research Letters* **24**, 245-248.

KYRIAKIDIS E. G. (1999b) Optimal control of a truncated general immigration process through total catastrophes, *Journal of Applied Probability* **36**, 461-472.

KYRIAKIDIS E. G. (1999c) Optimal isolation policies for controlling two competing diseases, *The Mathematical Scientist* **24**, 56-67.

KYRIAKIDIS E. G. and DIMITRAKOS T. D. (2002) Optimal control of two competing diseases with state-dependent infection rates, *The Mathematical Scientist* **27**, 36-44.

KYRIAKIDIS E. G. (2003) Optimal control of a simple immigration process through the introduction of a predator, *Probability in the Engineering and Informational Sciences* **17**, 119-135.

KYRIAKIDIS E. G. (2004) Optimal control of a simple immigration-emigration process through total catastrophes, *European Journal of Operational Research* **155**, 198-208.

KYRIAKIDIS E. G. and DIMITRAKOS T. D. (2004) Optimal preventive maintenance of a production system with an intermediate buffer, To appear in *European Journal of Operational Research*.

KYRIAKIDIS E. G. and DIMITRAKOS T. D. (2005a) A pest immigration process controlled by an intermittent predator, To appear in *The Mathematical Scientist*.

KYRIAKIDIS E. G. and DIMITRAKOS T. D. (2005b) Computation of the optimal policy for the control of a compound immigration process through total catastrophes, *Methodology and Computing in Applied Probability* **7**, 97-118.

LOVE C. E., ZHANG Z. G., ZITRON M. A. and GUO R. (2000) A discrete semi-Markov decision model to determine the optimal repair/replacement policy under general repairs, *European Journal of Operational Research* **125**, 398-409.

MAKIS V. and JARDINE A. (1993) A note on optimal replacement policy under general repair, *European Journal of Operational Research* **69**, 75-82.

MELLER R. D. and KIM D. S. (1996) The impact of preventive maintenance on system cost and buffer size, *European Journal of Operational Research* **95**, 577-591.

NAKAGAWA T. and KOWADA M. (1983) Analysis of a system with minimal repair and its application to replacement policy, *European Journal of Operational Research* **12**, 176-182.

NOBEL R. D. and TIJMS H. C. (1999) Optimal control for an $M^X/G/1$ queue with two service modes, *European Journal of Operational Research* **113**, 610-619.

NOBEL R. D. and TIJMS H. C. (2000) Optimal control of a queueing system with heterogeneous servers and setup costs, *IEEE Transactions on Automatic Control* **45**, 780-784.

O' NEIL P. D. (1997) An epidemic model with removal-dependent infection rate, *Annals of Applied Probability* **7**, 90-109.

OZEKICI S. and GUNLUNK N. O. (1992) Maintenance of a device with age-dependent exponential failures, *Naval Research Logistics* **39**, 699-714.

PHELPS R. I. (1981) Replacement policies under minimal repair, *Journal of Operations Research Society* **32**, 549-554.

PHELPS R. I. (1983) Optimal policy for minimal repair, *Journal of Operations Research Society* **34**, 425-427.

PIERSKALA W. P. and VOELKER L. A. (1976) A survey of maintenance model, *Naval Research Logistics Quarterly* **23**, 353-388.

PUTERMAN M. L. (1994) *Markov Decision Processes: Discrete Stochastic Dynamic Programming*, Wiley, New York.

ROSS S. M. (1983) *Introduction to Stochastic Dynamic Programming*, Academic Press, New York.

ROSS S. M. (1992) *Applied Probability Models with Optimization Applications*, Dover, New York.

SALAMEH M. K. and GHATTAS R. E. (2001) Optimal just-in-time buffer inventory for regular preventive maintenance, *International Journal of Production Economics* **74**, 157-161.

SAUNDERS I. W. (1980a) A model for myxomatosis, *Mathematical Biosciences* **48**, 1-15.

SAUNDERS I. W. (1980b) An approximate maximum likelihood estimator for chain binomial models, *Australian Journal of Statistics* **22**, 307-316.

SCARF P. A. (1997) On the application of mathematical models in maintenance, *European Journal of Operational Research* **99**, 493-506.

SCHWEITZER P. J. (1971) Iterative solution of the functional equations of undiscounted Markov renewal programming, *Journal of Mathematical Analysis and Applications* **34**, 495-501.

SENNOTT L. I. (1989) Average cost semi-Markov decision processes and the control of queueing systems, *Probability in the Engineering and Informational Sciences* **3**, 247-272.

SENNOTT L. I. (1997) The computation of average optimal policies in denumerable state Markov decision chains, *Advances in Applied Probability* **29**, 114-137.

SENNOTT L. I. (1999) *Stochastic Dynamic Programming and the Control of Queueing Systems*, Wiley, New York.

SERFOZO R. F. (1979) An equivalence between continuous and discrete time Markov decision processes, *Operations Research* **27**, 616-620.

SEVERO N. C. (1969) Generalizations of some stochastic epidemic models, *Mathematical Biosciences* **4**, 395-402.

SO K. C. (1992) Optimality of control limit policies in replacement models, *Naval Research Logistics* **39**, 685-697.

THOMAS L. C. and STENGOS D. (1985) Finite state approximation algorithms for average cost denumerable state Markov decision processes, *OR Spektrum* **7**, 27-37.

TIJMS H. C. (1994) *Stochastic Models: An Algorithmic Approach*, Wiley, New York.

VALDEZ-FLORES C. and FELDMAN R. M. (1989) A survey of preventive maintenance models for stochastically deteriorating single-unit systems, *Naval Research Logistics* **36**, 419-446.

VAN DER DUYN SCHOUTEN F. A. and VANNESTE S. G. (1995) Maintenance optimization of a production system with buffer capacity, *European Journal of Operational Research* **82**, 323-338.

VANNESTE S. G. (1992) A generalized age-replacement model, *Probability in the Engineering and Informational Sciences* **6**, 525-541.

WANG H. (2002) A survey of maintenance policies of deteriorating systems, *European Journal of Operational Research* **139**, 469-489.

WHITE D. J. (1980) Finite state approximations for denumerable state infinite horizon discounted Markov decision processes: The method of successive approximation, In *Recent Developments in Markov Decision Processes*, edited by R. Hartley, L. C. Thomas and D. J. White, Academic Press, New York, 57-72.