

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ
Σχολή Χρηματοοικονομικής και Στατιστικής



Τμήμα Στατιστικής και Ασφαλιστικής Επιστήμης

ΜΕΤΑΠΤΥΧΙΑΚΟ ΠΡΟΓΡΑΜΜΑ ΣΠΟΥΔΩΝ
ΣΤΗΝ ΕΦΑΡΜΟΣΜΕΝΗ ΣΤΑΤΙΣΤΙΚΗ

Ο αλγόριθμος της Προσομοιωμένης Ανόπτησης
(Simulated Annealing) με χρηματοοικονομικές
εφαρμογές.

Άννα Κουρούκλη

Διπλωματική Εργασία

που υποβλήθηκε στο Τμήμα Στατιστικής και Ασφαλιστικής
Επιστήμης του Πανεπιστημίου Πειραιώς ως μέρος των
απαιτήσεων για την απόκτηση του Μεταπτυχιακού
Διπλώματος Ειδίκευσης στην *Εφαρμοσμένη Στατιστική*

Πειραιάς
Νοέμβριος 2018

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ
Σχολή Χρηματοοικονομικής και Στατιστικής



Τμήμα Στατιστικής και Ασφαλιστικής Επιστήμης

ΜΕΤΑΠΤΥΧΙΑΚΟ ΠΡΟΓΡΑΜΜΑ ΣΠΟΥΔΩΝ
ΣΤΗΝ ΕΦΑΡΜΟΣΜΕΝΗ ΣΤΑΤΙΣΤΙΚΗ

Ο αλγόριθμος της Προσομοιωμένης Ανόπτησης
(Simulated Annealing) με χρηματοοικονομικές
εφαρμογές.

Άννα Κουρούκλη

Διπλωματική Εργασία

που υποβλήθηκε στο Τμήμα Στατιστικής και Ασφαλιστικής
Επιστήμης του Πανεπιστημίου Πειραιώς ως μέρος των
απαιτήσεων για την απόκτηση του Μεταπτυχιακού
Διπλώματος Ειδίκευσης στην *Εφαρμοσμένη Στατιστική*

Πειραιάς
Νοέμβριος 2018

Η παρούσα Διπλωματική Εργασία εγκρίθηκε ομόφωνα από την Τριμελή Εξεταστική Επιτροπή που ορίστηκε από τη ΓΣΕΣ του Τμήματος Στατιστικής και Ασφαλιστικής Επιστήμης του Πανεπιστημίου Πειραιώς στην υπ' αριθμ. συνεδρίασή του σύμφωνα με τον Εσωτερικό Κανονισμό Λειτουργίας του Προγράμματος Μεταπτυχιακών Σπουδών στην Εφαρμοσμένη Στατιστική

Τα μέλη της Επιτροπής ήταν:

- (Επιβλέπων)
-
-

Η έγκριση της Διπλωματικής Εργασίας από το Τμήμα Στατιστικής και Ασφαλιστικής Επιστήμης του Πανεπιστημίου Πειραιώς δεν υποδηλώνει αποδοχή των γνώμων του συγγραφέα.

UNIVERSITY OF PIRAEUS
School of Finance and Statistics



Department of Statistics and Insurance Science

**POSTGRADUATE PROGRAM IN
APPLIED STATISTICS**

**THE SIMULATED ANNEALING
ALGORITHM WITH FINANCIAL
APPLICATIONS**

By

Anna Kouroukli

MSc Dissertation

submitted to the Department of Statistics and Insurance
Science of the University of Piraeus in partial fulfilment
of the requirements for the degree of Master of Science in
Applied Statistics

Piraeus, Greece
November 2018

Περίληψη

Αντικείμενο της παρούσας εργασίας είναι η παρουσίαση του αλγορίθμου της προσομοιωμένης ανόπτωσης και η εφαρμογή του σε χρηματοοικονομικά προβλήματα βελτιστοποίησης. Ο αλγόριθμος της προσομοιωμένης ανόπτωσης ανήκει στην κατηγορία των Monte Carlo τεχνικών βελτιστοποίησης. Το βασικό χαρακτηριστικό του αλγορίθμου είναι ότι, κατά την εξερεύνηση του χώρου λύσεων του προβλήματος, αποδέχεται με κάποια πιθανότητα και μεταβάσεις που δεν βελτιώνουν την αντικειμενική συνάρτηση. Με αυτόν τον τρόπο, ευελπιστούμε ότι θα αποφύγουμε τον εγκλωβισμό σε περιοχές τοπικών ελαχίστων.

Στην εργασία αυτή περιλαμβάνονται πέντε κεφάλαια:

Στο πρώτο κεφάλαιο γίνεται μία σύντομη εισαγωγή στις μεθόδους προσομοίωσης. Καταρχάς, αναφερόμαστε στη σημασία των τυχαίων αριθμών από τη $U(0,1)$ (Uniform Distribution in $(0,1)$) για την αναπαράσταση ενός στοχαστικού φαινομένου και παρουσιάζουμε μία από τις πιο απλές μεθόδους παραγωγής τυχαίων αριθμών, τη γραμμική αναλογική γεννήτρια. Στη συνέχεια, αναφερόμαστε σε κάποιες από τις τεχνικές, μέσω των οποίων μπορούμε να προσομοιώσουμε τυχαίες μεταβλητές που προέρχονται από μία συγκεκριμένη κατανομή. Το κεφάλαιο ολοκληρώνεται με την παρουσίαση του αλγορίθμου τοπικής αναζήτησης.

Στο δεύτερο κεφάλαιο παρουσιάζεται ο αλγόριθμος της προσομοιωμένης ανόπτωσης. Ξεκινάμε περιγράφοντας τη διαδικασία της ανόπτωσης στα φυσικά συστήματα. Ο λόγος που συμβαίνει αυτό είναι ότι η ιδέα στην οποία βασίζεται ο αλγόριθμος είναι η μίμηση της διαδικασίας της ανόπτωσης για την επίλυση προβλημάτων βελτιστοποίησης. Στη συνέχεια, παρουσιάζεται η μαθηματική μοντελοποίηση του αλγορίθμου μέσω Μαρκοβιανών αλυσίδων και δίνονται αποτελέσματα σχετικά με τη σύγκλιση του αλγορίθμου. Το κεφάλαιο ολοκληρώνεται με ένα παράδειγμα: εφαρμόζουμε τον αλγόριθμο για την εύρεση του μέγιστου ύψους μιας επιφάνειας που έχει ένα τοπικό και ένα ολικό μέγιστο.

Στο τρίτο κεφάλαιο γίνεται μία συνοπτική παρουσίαση των αυτοπαλίνδρομων υποδειγμάτων κινητού μέσου (*ARMA models*), καθώς και των υποδειγμάτων δεσμευμένης ετεροσκεδαστικότητας (*ARCH, GARCH models*).

Στο τέταρτο κεφάλαιο εφαρμόζεται ο αλγόριθμος της προσομοιωμένης ανόπτωσης για την εκτίμηση των παραμέτρων ενός $GARCH(1,1)$ υποδείγματος. Πιο συγκεκριμένα, μεγιστοποιείται η συνάρτηση πιθανοφάνειας του υποδείγματος με βάση προσομοιωμένα δεδομένα. Η συγκεκριμένη διαδικασία επαναλαμβάνεται για διάφορες τιμές των παραμέτρων του αλγορίθμου. Επιπλέον, εκτός του αλγορίθμου της προσομοιωμένης ανόπτωσης, εφαρμόζεται και η μέθοδος *Quasi Newton* για τη μεγιστοποίηση της συνάρτησης πιθανοφάνειας, καθώς θέλουμε να συγκρίνουμε την αποτελεσματικότητα των δύο μεθόδων. Η υλοποίηση του αλγορίθμου της προσομοιωμένης ανόπτωσης γίνεται μέσω ενός κώδικα, ο οποίος αναπτύχθηκε στη γλώσσα προγραμματισμού *R*.

Στο πέμπτο κεφάλαιο (Παράρτημα) παραθέτουμε κάποιους βασικούς ορισμούς και θεωρήματα από τη θεωρία των Μαρκοβιανών αλυσίδων.

Η εργασία ολοκληρώνεται με τη σύνοψη των συμπερασμάτων που προέκυψαν από την εκπόνηση της συγκεκριμένης εργασίας.

Abstract

The purpose of this MSc dissertation is the presentation of the simulated annealing algorithm and its implementation in financial optimization problems. Simulated annealing is a Monte Carlo optimization method. The key feature of the algorithm is that, during the exploitation of the solution space, it may also allow moves that worsen the objective function value, in hopes of escaping local optima.

This dissertation includes five chapters.

In the first chapter, we present a brief review of Monte Carlo simulation methods. Initially, we refer to the importance of generating uniform (pseudo-)random variables on the interval $(0,1)$ for the simulation of a stochastic phenomenon, and we also describe the method of linear congruential generator. Moreover, we briefly present techniques for the generation of (pseudo-)random variables from various probability distributions. We close this chapter with a short review of the local search algorithm.

In the second chapter, we present the simulated annealing algorithm. We begin by describing the physical process of the annealing of solids, given that the algorithm is based on the idea of imitating that process, in order to solve optimization problems. We continue by describing the mathematical model of the algorithm by means of the Markov chain theory. Moreover, we provide results regarding the convergence of the algorithm. The chapter closes with an example. We use the algorithm in order to find the maximum height of a surface that has a local and a global maximum..

In the third chapter, we provide a brief review on the autoregressive moving average models (ARMA) and the conditional heteroskedastic models (GARCH).

In the fourth chapter, we implement the simulated annealing algorithm for the estimation of the parameters of a $GARCH(1,1)$ model. In particular, we maximize the likelihood function of the model based on simulated data. This procedure is repeated for various values of the algorithm parameters in order to study its performance under various scenarios. We numerically compare the accuracy of the the simulated annealing algorithm with the *Quasi Newton* method. The simulated annealing algorithm is implemented using *R* statistical software .

In the fifth chapter (Appendix), we present a short review of Markov chain theory.

We finally present and discuss the main conclusions that have emerged from our empirical study concerning the performance of the simulated annealing algorithm.

Περιεχόμενα

Περίληψη	5
Abstract	7
Κεφάλαιο 1: Μέθοδοι Προσομοίωσης	11
1.1 Εισαγωγή στις μεθόδους προσομοίωσης	11
1.2 Παραγωγή τυχαίων αριθμών	12
1.2.1 Γραμμική αναλογική γεννήτρια	13
1.3 Μέθοδος αντιστροφής	14
1.4 Παραγωγή τ.μ. από την κανονική κατανομή.....	16
1.4.1 Μέθοδος Box Muller.....	17
1.4.2 Παραγωγή τ.μ. από την πολυδιάστατη κανονική κατανομή.....	18
1.5 Monte Carlo μέθοδοι βελτιστοποίησης	19
1.5.1 Τοπική αναζήτηση.....	20
Κεφάλαιο 2: Ο αλγόριθμος της Προσομοιωμένης Ανόπτωσης	22
2.1 Εισαγωγή στον αλγόριθμο της προσομοιωμένης ανόπτωσης.....	22
2.2 Ανόπτωση	22
2.3 Αλγόριθμος Metropolis	25
2.4 Περιγραφή του αλγορίθμου της προσομοιωμένης ανόπτωσης	25
2.5 Σύγκλιση του αλγορίθμου της προσομοιωμένης ανόπτωσης	28
2.6 Χρονοδιάγραμμα ψύξης.....	34
2.7 Παράδειγμα	35
Κεφάλαιο 3: Χρονολογικές Σειρές	44
3.1 Γραμμικές διαδικασίες και αυτοπαλίνδρομα υποδείγματα κινητού μέσου.....	44
3.2 Το Αυτοπαλίνδρομο Υπόδειγμα	44

3.3 Το υπόδειγμα κινητού μέσου	45
3.4 Το Αυτοπαλίνδρομο Υπόδειγμα Κινητού Μέσου $ARMA(p,q)$	46
3.5 Υποδείγματα Δεσμευμένης Ετεροσκεδαστικότητας.....	46
3.5.1 Το υπόδειγμα ARCH.....	47
3.5.2 Το υπόδειγμα GARCH	48
3.5.3 Εκτίμηση παραμέτρων του υποδείγματος $GARCH(1,1)$	52
Κεφάλαιο 4: Εφαρμογή της προσομοιωμένης ανόπτωσης με την βοήθεια της R54	
4.1 Εισαγωγή.....	54
4.2 Εκτίμηση των παραμέτρων μέσω της προσομοιωμένης ανόπτωσης.....	55
4.3 Παραδείγματα	57
4.3.1 Εκτίμηση των παραμέτρων ενός $GARCH(1,1)$ υποδείγματος για διάφορες τιμές της παραμέτρου α	57
4.3.2 Εκτίμηση των παραμέτρων ενός $GARCH(1,1)$ υποδείγματος για διάφορες τιμές της παραμέτρου s	64
4.3.3 Σύγκριση μεταξύ των μεθόδων Quasi-Newton και προσομοιωμένης ανόπτωσης	66
Παράρτημα: Μαρκοβιανές αλυσίδες	77
Σύνοψη-Συζήτηση Αποτελεσμάτων	822
Βιβλιογραφία.....	84

Κεφάλαιο 1

Μέθοδοι Προσομοίωσης

1.1 Εισαγωγή στις μεθόδους προσομοίωσης

Οι μέθοδοι προσομοίωσης πρωτοεμφανίστηκαν στο τέλος της δεκαετίας του '40 και, έκτοτε, έχουν καταστεί αναπόσπαστο υπολογιστικό εργαλείο σε διάφορες επιστήμες, όπως είναι η στατιστική, τα οικονομικά κ.α. Πολλές φορές, είναι ο πιο απλός τρόπος να επιλύσουμε (αριθμητικά) ένα πρόβλημα ενώ, σε αρκετές περιπτώσεις σύνθετων προβλημάτων, ο μόνος εφικτός τρόπος να τα μελετήσουμε.

Οι μέθοδοι προσομοίωσης σχετίζονται με τη διαδικασία μοντελοποίησης και αναπαράστασης ενός στοχαστικού φαινομένου. Δηλαδή, μέσω της παραγωγής πολλών σεναρίων του στοχαστικού φαινομένου, συλλέγουμε διάφορα μεγέθη που μας επιτρέπουν την εξαγωγή «εμπειρικών» συμπερασμάτων. Όπως γίνεται εύκολα αντιληπτό, όσο αυξάνεται το πλήθος των επαναλήψεων του στοχαστικού φαινομένου, θα αυξάνεται και η ποιότητα των αποτελεσμάτων.

Σημαντική ώθηση στην ανάπτυξη των μεθόδων προσομοίωσης προσέδωσε η ανάπτυξη της επιστήμης των υπολογιστών, αφού η τεράστια υπολογιστική ισχύς τους μας παρείχε τη δυνατότητα για την παρακολούθηση πολλαπλών επαναλήψεων του στοχαστικού φαινομένου. Γι' αυτόν τον λόγο, όταν αναφερόμαστε στις μεθόδους προσομοίωσης, πρακτικά εννοούμε την «εικονική» αναπαράσταση ενός στοχαστικού φαινομένου μέσω H/Y και την παρακολούθηση χιλιάδων ή και εκατομμυρίων επαναλήψεών του.

Ενδεικτικές κατηγορίες προβλημάτων στις οποίες μπορούμε να εφαρμόσουμε τις μεθόδους προσομοίωσης είναι οι ακόλουθες:

- Προβλήματα βελτιστοποίησης
- Υπολογισμοί ολοκληρωμάτων
- Προβλήματα στατιστικής φύσεως

Για παράδειγμα, με τη μέθοδο της προσομοίωσης μπορούμε να υπολογίσουμε τη μέση τιμή (όταν αυτή υπάρχει) μίας τυχαίας μεταβλητής, με συνάρτηση κατανομής F και σ.π.π. f . Ως γνωστόν, η μέση τιμή $E(X)$ της f είναι ίση με

$$\int_{-\infty}^{\infty} xf(x)dx.$$

Όμως, σε πολλές περιπτώσεις, είναι πολύ δύσκολο να υπολογιστεί το παραπάνω ολοκλήρωμα αναλυτικά ή με μεθόδους αριθμητικής ανάλυσης. Σε αυτές τις περιπτώσεις, καταφεύγουμε στις μεθόδους προσομοίωσης.

Η τεχνική, πάνω στην οποία θα βασιστούμε, είναι να χρησιμοποιήσουμε τυχαίους αριθμούς από την κατανομή F , ώστε να προσδιορίσουμε τα χαρακτηριστικά της, εν προκειμένω τη μέση τιμή. Έστω ότι έχουμε X_1, X_2, \dots, X_n ανεξάρτητες, τυχαίες μεταβλητές από την κατανομή F . Τότε, από τον Νόμο των Μεγάλων Αριθμών ισχύει:

$$\frac{X_1 + X_2 + \dots + X_n}{n} \xrightarrow{n \rightarrow \infty} E(X).$$

Όπως φαίνεται από το παράδειγμα, οι μέθοδοι προσομοίωσης μάς δίνουν τη δυνατότητα, με έναν απλό και άμεσο τρόπο, να εκτιμήσουμε την ποσότητα που μας ενδιαφέρει. Ωστόσο, θα πρέπει να προσέξουμε τα ακόλουθα σημεία προκειμένου να διασφαλίσουμε ότι τα αποτελέσματα που προκύπτουν από την εφαρμογή της μεθόδου είναι αξιόπιστα:

- Αρχικά, θα πρέπει να είμαστε σίγουροι ότι χρησιμοποιούμε το κατάλληλο μοντέλο για το στοχαστικό μοντέλο που θέλουμε να μελετήσουμε.
- Οι εκτιμήσεις που προκύπτουν με την εφαρμογή μεθόδων προσομοίωσης προσεγγίζουν με σχετικά μικρή ακρίβεια. Γι' αυτό, πολλές φορές, χρειάζεται να αυξηθεί σημαντικά το μέγεθος του δείγματος, ώστε να μειωθεί η διακύμανση της εκτίμησης.
- Θα πρέπει να αναπτύξουμε κατάλληλες μεθόδους παραγωγής τυχαίων μεταβλητών.

1.2 Παραγωγή τυχαίων αριθμών

Οι τυχαίοι αριθμοί είναι χρήσιμοι σε εφαρμογές διαφόρων επιστημονικών πεδίων. Ειδικότερα στην προσομοίωση, οι τυχαίοι αριθμοί είναι το πρωταρχικό στοιχείο που χρησιμοποιούμε ως βάση για να αναπαραστήσουμε ένα στοχαστικό φαινόμενο.

Παλιότερα, αρκετές ήταν οι μέθοδοι παραγωγής τυχαίων αριθμών όπως: η ρίψη ενός νομίσματος, τα ζάρια και το ανακάτεμα της τράπουλας. Εξαιτίας, όμως, της

μηχανικής φύσης αυτών των τεχνικών, η παραγωγή μεγάλων ακολουθιών τυχαίων αριθμών ήταν αρκετά δύσκολη διαδικασία. Γι' αυτό είχε επικρατήσει η χρήση πινάκων τυχαίων αριθμών με τα αποτελέσματα τέτοιων πειραμάτων.

Πλέον, η παραγωγή μίας ακολουθίας τυχαίων αριθμών επιτυγχάνεται με τη χρήση η/υ. Οι λόγοι που προτιμούμε τον συγκεκριμένο τρόπο έναντι των υπολοίπων είναι: η μικρή υπολογιστική ισχύ που απαιτεί, η ταχύτητα και η επαναληψιμότητα. Ο τελευταίος λόγος -και ο πιο σημαντικός- είναι ότι μέσω του η/υ μπορούμε να παράγουμε την ίδια ακολουθία σε πολλά «τρεξίματα» της ίδιας εφαρμογής, και έτσι γίνεται ευκολότερη η διόρθωση σφαλμάτων. Αν θεωρήσουμε π.χ. ότι σε κάποιο σημείο του αλγορίθμου πιθανόν υπάρχει λάθος, λογικά θα πρέπει να επαναλάβουμε όλα τα βήματα μέχρι αυτό το σημείο. Προφανώς, μία διαφορετική ακολουθία τυχαίων αριθμών θα ήταν μη αποδεκτή.

1.2.1 Γραμμική αναλογική γεννήτρια

Υπάρχουν πάρα πολλές μέθοδοι παραγωγής τυχαίων αριθμών. Σ' αυτήν την παράγραφο, θα παρουσιάσουμε ως παράδειγμα την απλούστερη μέθοδο: τη γραμμική αναλογική γεννήτρια (linear congruential generators). Με τη συγκεκριμένη γεννήτρια κάθε αριθμός προκύπτει από τον προηγούμενο μέσω μίας αναδρομικής σχέσης, αφού πρώτα ορίσουμε την αρχική τιμή X_0 .

Μέσω της συγκεκριμένης μεθόδου μπορεί να παραχθεί μία ακολουθία ψευδοτυχαίων αριθμών, αν χρησιμοποιήσουμε τον ακόλουθο αναδρομικό τύπο:

$$X_{i+1} = (aX_i + c) \bmod m, \quad (1)$$

όπου το X_0 λέγεται αρχική τιμή (seed), ενώ τα a, c και m είναι ο πολλαπλασιαστής, η αυξητική σταθερά και το modulus, αντίστοιχα. Ειδικότερα, το mod υποδηλώνει το υπόλοιπο της διαίρεσης του $aX_i + c$ με το m . Συνεπώς, κάθε X_i παίρνει τιμές στο σύνολο $\{0, 1, \dots, m - 1\}$, ενώ οι ποσότητες

$$U_i = \frac{X_i}{m}$$

αναφέρονται ως ψευδοτυχαίοι αριθμοί και μπορεί να θεωρηθεί ότι αποτελούν προσεγγιστικά μία ακολουθία τυχαίων αριθμών από την ομοιόμορφη κατανομή στο $(0,1)$.

Φυσικά, είναι προφανές ότι οι αριθμοί που προκύπτουν από τη μέθοδο δεν είναι τυχαίοι. Γι' αυτόν τον λόγο, άλλωστε, τους αποκαλούμε *ψευδοτυχαίους* αριθμούς. Στην πραγματικότητα, η ακολουθία των αριθμών που παράγουμε είναι περιοδική με περίοδο k , όπου k ακέραιος μικρότερος ή ίσος του m . Δηλαδή, μετά από k βήματα θα εμφανισθεί ο X_0 για δεύτερη φορά. Από εκεί κι έπειτα, επαναλαμβάνεται η ίδια ακολουθία.

Είναι πολύ πιθανό, για αυθαίρετες επιλογές των X_0, a, c και m , η ακολουθία που θα παράγουμε να μην έχει τις επιθυμητές στατιστικές ιδιότητες. Κατά την εφαρμογή της μεθόδου, το m πρέπει να είναι μεγάλο για να λειτουργεί καλά η γεννήτρια. Γι' αυτόν τον λόγο, συνήθως το m επιλέγεται ως ένας μεγάλος πρώτος αριθμός που ταιριάζει με το μέγεθος λέξεων που χειρίζεται ο υπολογιστής. Διότι, ιδανικά, θα θέλαμε το μήκος της περιόδου να είναι ίσο με m ή τουλάχιστον να το πλησιάζει. Σύμφωνα με τα αποτελέσματα της θεωρίας των αριθμών, αν ισχύουν οι ακόλουθες συνθήκες, τότε μπορούμε να επιτύχουμε πλήρη περίοδο.

- Οι παράμετροι m, c είναι μεταξύ τους πρώτοι.
- Αν το 4 διαιρεί το m , τότε θα πρέπει να διαιρεί και το $a - 1$.
- Αν q είναι πρώτος αριθμός που διαιρεί τον m , τότε θα πρέπει να διαιρεί και τον $a - 1$.

Οι γραμμικές αναλογικές γεννήτριες μάς βοηθούν να κατανοήσουμε τη δομή αλλά και τους περιορισμούς της διαδικασίας παραγωγής τυχαίων αριθμών. Ωστόσο, σε πειράματα προσομοίωσης που απαιτούν μία μεγάλου μήκους ακολουθία τυχαίων αριθμών χρησιμοποιούνται πιο εξεζητημένες γεννήτριες, οι οποίες, όμως, έχουν ως βάση τις γραμμικές αναλογικές γεννήτριες.

Πλέον, οι περισσότερες γλώσσες προγραμματισμού αλλά και τα στατιστικά πακέτα περιέχουν τις δικές τους γεννήτριες τυχαίων αριθμών, πάνω στις οποίες μπορούμε να βασιστούμε. Ο χρήστης το μόνο που πρέπει να κάνει είναι να ορίσει την αρχική τιμή X_0 . Για παράδειγμα, στο Mathematica μπορούμε να παράγουμε τυχαίους αριθμούς μέσω της `RandomReal`.

1.3 Μέθοδος αντιστροφής

Υπάρχουν διάφορες τεχνικές, μέσω των οποίων μπορούμε να προσομοιώσουμε τυχαίες μεταβλητές που προέρχονται από μία συγκεκριμένη κατανομή F , βασιζόμενοι

σε τυχαίους αριθμούς από την ομοιόμορφη κατανομή στο $(0,1)$. Μία από αυτές τις μεθόδους είναι και η μέθοδος της αντιστροφής, η οποία βασίζεται στην παρατήρηση ότι η τυχαία μεταβλητή $X = F^{-1}(U)$, όπου $U \sim U(0,1)$, ακολουθεί την επιθυμητή κατανομή F .

Κατά την εφαρμογή της μεθόδου, θα αρκεστούμε στην έννοια της γενικευμένης αντίστροφης της συνάρτησης κατανομής, διότι έτσι μπορούμε να εφαρμόσουμε τη μέθοδο σε διακριτές κατανομές αλλά και σε συνεχείς, όπου η F δεν είναι γνησίως αύξουσα. Ενώ, αν χρησιμοποιήσουμε την αντίστροφη, θα έπρεπε η συνάρτηση κατανομής να είναι γνησίως αύξουσα.

Ορισμός 1.1 Γενικευμένη αντίστροφη της συνάρτησης κατανομής F

Έστω F μία σ.κ. Η γενικευμένη αντίστροφη της ορίζεται ως εξής:

$$F^{-1}(y) = \inf F^{-1}([y, 1]) = \inf\{x: F(x) \in [y, 1]\}, y \in [0,1]$$

Η γενικευμένη αντίστροφη, όταν η F είναι γνησίως αύξουσα, ταυτίζεται με την αντίστροφη συνάρτηση F^{-1} .

Στην περίπτωση που η F είναι συνάρτηση κατανομής, ως γνωστόν είναι δεξιά συνεχής. Οπότε το infimum επιτυγχάνεται εντός του συνόλου $\{x : F(x) \in [y, 1]\}$ και, επομένως, η γενικευμένη αντίστροφη μπορεί να γραφεί στην ακόλουθη μορφή:

$$F^{-1}(y) = \min\{x: F(x) \in [y, 1]\}.$$

Πρόταση 1.1 Μέθοδος της αντιστροφής

Αν $U \sim U(0,1)$ και F είναι μία οποιαδήποτε συνάρτηση κατανομής, τότε η τυχαία μεταβλητή $X = F^{-1}(U)$ έχει συνάρτηση κατανομής F .

Απόδειξη: Η συνάρτηση κατανομής της τυχαίας μεταβλητής X είναι ίση με

$$P(X \leq x) = P(F^{-1}(U) \leq x) = P(\min\{t: F(t) \geq U\} \leq x).$$

Διαπιστώνοντας τώρα ότι $\min\{t: F(t) \geq U\} \leq x \Leftrightarrow U \leq F(x)$, διότι

- $\min\{t: F(t) \geq U\} \leq x \Rightarrow \exists t_0 \leq x: F(t_0) \geq U$
και επειδή
 $F(t_0) \leq F(x) \Rightarrow F(x) \geq U$
- $F(x) \geq U \Rightarrow x \in \{t: F(t) \geq U\} \Rightarrow x \geq \min\{t: F(t) \geq U\},$

έπεται τελικά ότι

$$P(X \leq x) = P(U \leq F(x)) = F(x).$$

Ως εκ τούτου, με τη μέθοδο της αντιστροφής μπορούμε να παράγουμε μία τυχαία μεταβλητή X από μία κατανομή F , σύμφωνα με τον ακόλουθο αλγόριθμο.

Αλγόριθμος 1.1 Μέθοδος της αντιστροφής

1. Παράγουμε έναν τυχαίο αριθμό $U \sim U(0,1)$
2. Θέτουμε $X = F^{-1}(U)$

Η μέθοδος της αντιστροφής χρησιμοποιείται για την παραγωγή τυχαίων αριθμών από διάφορες κατανομές, όπως για παράδειγμα εκθετική, ομοιόμορφη, Weibull, Cauchy και λογιστική. Όμως, σε αρκετές περιπτώσεις, είναι υπολογιστικά δύσκολο να βρούμε την αντίστροφη της συνάρτησης κατανομής. Γι' αυτόν τον λόγο, καταφεύγουμε σε εναλλακτικές τεχνικές.

1.4 Παραγωγή τ.μ. από την κανονική κατανομή

Αν η τυχαία μεταβλητή X ακολουθεί την κανονική κατανομή $N(\mu, \sigma^2)$, η συνάρτηση πυκνότητας πιθανότητας έχει την ακόλουθη μορφή:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), -\infty < x < \infty$$

όπου μ είναι η μέση τιμή και σ^2 η διασπορά της κατανομής.

Σε πολλές εφαρμογές είναι απαραίτητο να προσομοιώσουμε τυχαίες μεταβλητές από την κανονική κατανομή $N(\mu, \sigma^2)$. Από τις ιδιότητες της κανονικής κατανομής γνωρίζουμε ότι αν $Z \sim N(0,1)$, τότε $\mu + \sigma Z \sim N(\mu, \sigma^2)$. Επομένως, για την παραγωγή τυχαίων αριθμών από τη $N(\mu, \sigma^2)$ μπορούμε να αρκεστούμε στην παραγωγή τυχαίων αριθμών από τη $N(0,1)$. Δεδομένου ότι δεν υπάρχει κλειστός τύπος για την Φ^{-1} (Φ σ.κ της $N(0,1)$), η μέθοδος της αντιστροφής δεν είναι η πλέον κατάλληλη τεχνική (αν και χρησιμοποιείται προσεγγιστικά, όταν έχουμε μία πολύ καλή προσεγγιστική έκφραση για την Φ^{-1}). Ως εκ τούτου, για την παραγωγή τυχαίων αριθμών από την κανονική κατανομή χρησιμοποιούμε εναλλακτικές τεχνικές. Μία από αυτές είναι και η ακόλουθη, η οποία αναπτύχθηκε από τους Box και Muller.

1.4.1 Μέθοδος Box Muller

Μέσω της μεθόδου Box Muller παράγουμε δύο ανεξάρτητες τ.μ. από την $N(0,1)$ χρησιμοποιώντας δύο ανεξάρτητες τ.μ., τις U_1, U_2 , που ακολουθούν την $U(0,1)$. Για να το επιτύχει αυτό, η συγκεκριμένη μέθοδος εκμεταλλεύεται την αναπαράσταση των σημείων του επιπέδου από τις πολικές συντεταγμένες.

Ως γνωστόν, η θέση ενός σημείου (X, Y) του επιπέδου προσδιορίζεται μονοσήμαντα όχι μόνο από τις καρτεσιανές συντεταγμένες, αλλά και από τις αντίστοιχες πολικές (R^2, θ) . Οι σχέσεις που συνδέουν τις καρτεσιανές με τις πολικές συντεταγμένες είναι οι εξής:

$$R = \sqrt{X^2 + Y^2} \quad \theta = \arctan\left(\frac{Y}{X}\right) \quad (1.2)$$

Αν θεωρήσουμε ένα τυχαίο σημείο του επιπέδου (X, Y) , όπου X, Y ανεξάρτητες τ.μ. από την $N(0,1)$, τότε για την από κοινού σ.π.π. ισχύει:

$$f_{(x,y)}(x, y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} = \frac{1}{2\pi} e^{-\frac{x^2+y^2}{2}} \quad x, y \in R$$

Ας θεωρήσουμε τον μετασχηματισμό:

$$x = \sqrt{r} \cos\theta \quad \text{και} \quad y = \sqrt{r} \sin\theta.$$

Τότε,

$$J(r, \theta) = \begin{vmatrix} \frac{1}{2} r^{-\frac{1}{2}} \cos\theta & -r^{\frac{1}{2}} \sin\theta \\ \frac{1}{2} r^{-\frac{1}{2}} \sin\theta & r^{\frac{1}{2}} \cos\theta \end{vmatrix} = \frac{1}{2}$$

Άρα, η από κοινού σ.π.π. της (R^2, θ) είναι η εξής:

$$\begin{aligned} f_{R^2, \theta}(r, \theta) &= f_{(x,y)}(\sqrt{r} \cos\theta, \sqrt{r} \sin\theta) \frac{1}{2} \\ &= \frac{1}{2} e^{-\frac{r}{2}} \frac{1}{2\pi} = f_{R^2}(r) f_{\theta}(\theta) \quad r > 0, \theta \in [0, 2\pi). \end{aligned}$$

Επομένως, οι τ.μ. R^2 και θ είναι ανεξάρτητες, με $R^2 \sim \text{Exp}\left(\frac{1}{2}\right)$ και $\theta \sim U(0, 2\pi)$. Ως εκ τούτου, για κάθε σημείο του επιπέδου με πολικές συντεταγμένες (R, θ) , έτσι ώστε $R^2 \sim \text{Exp}\left(\frac{1}{2}\right)$ και $\theta \sim U(0, 2\pi)$ ανεξάρτητες τ.μ., τότε οι αντίστοιχες καρτεσιανές συντεταγμένες (X, Y) ακολουθούν την δισδιάστατη τυπική κανονική κατανομή.

Συνοψίζοντας τα παραπάνω, μπορούμε χρησιμοποιώντας τον ακόλουθο αλγόριθμο να παράγουμε τυχαίους αριθμούς από τη $N(\mu, \sigma^2)$.

Αλγόριθμος 1.2 Μέθοδος Box -Muller

1. Παράγουμε $U_2, U_1 \sim U(0,1)$

Θέτουμε $X_1 = \mu + \sigma\sqrt{-2\ln U_1}\cos(2\pi U_2)$ και $X_2 = \mu + \sigma\sqrt{-2\ln U_1}\sin(2\pi U_2)$

1.4.2 Παραγωγή τ.μ. από την πολυδιάστατη κανονική κατανομή

Η σ.π.π. της p -διάστατης κανονικής κατανομής έχει την ακόλουθη μορφή:

$$f(x_1, \dots, x_p) = \frac{1}{2\pi^{p/2}|\Sigma|^{1/2}} e^{-\frac{1}{2}(x - \mu)^T \Sigma^{-1}(x - \mu)}$$

όπου μ , είναι το $p \times 1$ διάνυσμα των μέσων και, όπου Σ , είναι ο $p \times p$ πίνακας συνδιακυμάνσεων του τυχαίου διανύσματος x .

Ο πίνακας Σ είναι συμμετρικός. Αν, επιπλέον, έχει γνήσια θετικές ιδιοτιμές, τότε είναι ένας θετικά ορισμένος πίνακας. Επομένως, από το θεώρημα φασματικής ανάλυσης ο Σ μπορεί να γραφεί στην μορφή $P\Lambda P^{-1}$, όπου P είναι ο πίνακας των ιδιοδιανυσμάτων του Σ και Λ είναι ο πίνακας που έχει στη διαγώνιο τις ιδιοτιμές του Σ . Ενώ για τον πίνακα $\Sigma^{1/2}$ ισχύει ότι είναι ίσος με $P\Lambda^{1/2}P^{-1}$.

Για την παραγωγή ενός τυχαίου διανύσματος $X = (X_1, X_2, \dots, X_p)$ από την $N_p(\mu, \Sigma)$ θα βασιστούμε στην ακόλουθη ιδιότητα της πολυδιάστατης κανονικής κατανομής:

αν $Z \sim N_p(\mathbf{0}, I)$, τότε η τυχαία μεταβλητή $X = \mu + \Sigma^{1/2}Z \sim N_p(\mu, \Sigma)$.

Πιο συγκεκριμένα, ισχύει ότι

$$E(X) = \mu + \Sigma^{1/2}E(Z) = \mu + \Sigma^{1/2} \cdot \mathbf{0} = \mu$$

και

$$\begin{aligned} Cov(X) &= E((X - \mu)(X - \mu)') = E\left(\left(\Sigma^{1/2}Z\right)\left(\Sigma^{1/2}Z\right)'\right) = E\left(\Sigma^{1/2}ZZ'\Sigma^{1/2}\right) \\ &= \Sigma^{1/2}E(ZZ')\Sigma^{1/2} = \Sigma^{1/2}I_p\Sigma^{1/2} = \Sigma \end{aligned}$$

Στον ακόλουθο αλγόριθμο περιγράφονται τα βήματα για την παραγωγή ενός τυχαίου διανύσματος $X \sim N_p(\mu, \Sigma)$.

Αλγόριθμος 1.3 Προσομοίωση τυχαίου διανύσματος \mathbf{X} από την $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$

1. Παράγουμε Z_1, Z_2, \dots, Z_p ανεξάρτητες τ.μ. από την $N(0,1)$.
2. Υπολογίζουμε τον $\boldsymbol{\Sigma}^{1/2} = \mathbf{P}\boldsymbol{\Lambda}^{1/2}\mathbf{P}^{-1}$.
3. Θέτουμε $\mathbf{X} = \boldsymbol{\mu} + \boldsymbol{\Sigma}^{1/2}\mathbf{Z}$.

Για περισσότερες πληροφορίες σχετικά με τις μεθόδους παραγωγής τ.μ., ο αναγνώστης παραπέμπεται στα συγγράμματα των C.P. Robert and G. Casella (2010), S. Ross (1997), D.P.Kroese, R.Y.Rubinstein (2007), C.P. Robert and G. Casella (2004), M. Μπούτσικας, (2004).

1.5 Monte Carlo μέθοδοι βελτιστοποίησης

Οι Monte Carlo μέθοδοι βελτιστοποίησης μάς δίνουν τη δυνατότητα να προσεγγίσουμε τη λύση ενός προβλήματος βελτιστοποίησης στις περιπτώσεις που οι ντετερμινιστικές τεχνικές είναι μη αποτελεσματικές. Το βασικό τους χαρακτηριστικό είναι ότι εξερευνούν με τυχαίο τρόπο τον χώρο των δυνατών λύσεων.

Έστω το ακόλουθο πρόβλημα βελτιστοποίησης:

$$\min_{\mathbf{x} \in S} f(\mathbf{x})$$

όπου $f: S \rightarrow R$ είναι η αντικειμενική συνάρτηση, \mathbf{x} είναι το διάνυσμα των μεταβλητών και το σύνολο S ο χώρος των δυνατών λύσεων. Ένα στοιχείο \mathbf{x}^* του S τέτοιο, ώστε $f(\mathbf{x}^*) \leq f(\mathbf{x})$ για κάθε $\mathbf{x} \in S$, αποκαλείται ολικό ελάχιστο της αντικειμενικής συνάρτησης.

Ένα σημαντικό ζήτημα που πρέπει να αντιμετωπίσουμε στα προβλήματα βελτιστοποίησης είναι η εύρεση του ολικού ελαχίστου μίας συνάρτησης που έχει πολλά τοπικά ακρότατα. Οι Monte Carlo μέθοδοι βελτιστοποίησης επιχειρούν να το επιτύχουν αυτό με την εισαγωγή τυχειότητας στον μηχανισμό παραγωγής υποψήφιων λύσεων.

Ένας ακόμη λόγος για τον οποίο στρεφόμαστε στις συγκεκριμένες τεχνικές είναι ότι η εφαρμογή τους δεν θέτει σαν προϋπόθεση να έχει συγκεκριμένες αναλυτικές ιδιότητες η αντικειμενική συνάρτηση του προβλήματος. Το μόνο που απαιτείται κατά την εφαρμογή των Monte Carlo τεχνικών βελτιστοποίησης είναι η εκτίμηση της αντικειμενικής συνάρτησης.

Στη συνέχεια, παρουσιάζουμε συνοπτικά τον αλγόριθμο της τοπικής αναζήτησης που αποτελεί τον πιο γενικό αλγόριθμο από την κλάση των Monte Carlo τεχνικών βελτιστοποίησης.

1.5.1 Τοπική αναζήτηση

Έστω ότι θέλουμε να βρούμε τη βέλτιστη τιμή $x \in S$ που ελαχιστοποιεί την ποσότητα $f(x)$. Η μέθοδος της τοπικής αναζήτησης είναι ίσως η πιο γενική και ευέλικτη μέθοδος που μπορούμε να εφαρμόσουμε.

Προτού προχωρήσουμε στην περιγραφή του αλγορίθμου, πρέπει να αναφέρουμε ότι η εφαρμογή της μεθόδου προϋποθέτει την ύπαρξη ενός μηχανισμού μετάβασης από μία λύση σε μία άλλη: δηλαδή, μία λύση y μπορεί να ληφθεί από μία μικρή διαταραχή μίας άλλης λύσης x . Με αυτόν τον τρόπο, για κάθε λύση x ορίζεται το σύνολο $N(x)$ των γειτονικών λύσεων, το οποίο αποτελείται από όλες τις λύσεις στις οποίες μπορεί να μεταβούμε σε ένα βήμα, μέσω του μηχανισμού μετάβασης.

Σε αυτό το σημείο κι αφού ορίσαμε την έννοια της γειτνίασης των λύσεων, μπορούμε ακολούθως να περιγράψουμε τον τρόπο λειτουργίας του αλγορίθμου τοπικής αναζήτησης. Ξεκινάμε με μία αρχική λύση, την οποία επιλέγουμε συνήθως με τυχαίο τρόπο. Ο αλγόριθμος όλες τις φορές διατηρεί μία τρέχουσα λύση $x \in S$. Δεδομένου ότι η x είναι η τρέχουσα λύση, διαλέγουμε μία υποψήφια λύση y από το σύνολο των γειτονικών λύσεων της x . Αν η υποψήφια λύση y βελτιώνει την τιμή της αντικειμενικής συνάρτησης $f(y) < f(x)$, τότε η τρέχουσα λύση αντικαθίσταται από την υποψήφια, και η διαδικασία επαναλαμβάνεται μέχρι να ικανοποιηθεί το κριτήριο τερματισμού. Συνήθως, ο αλγόριθμος τερματίζεται, όταν η τιμή της αντικειμενικής συνάρτησης της τρέχουσας λύσης είναι μικρότερη από αυτή των γειτόνων της ή όταν συμπληρωθεί προκαθορισμένος αριθμός επαναλήψεων. Επιπλέον, θα πρέπει να επισημάνουμε ότι τα αποτελέσματα που επιστρέφει ο αλγόριθμος εξαρτώνται από την επιλογή του μηχανισμού μετάβασης.

Αλγόριθμος 1.4 Τοπική αναζήτηση

1. Αρχικοποίηση: παράγουμε μία λύση $x \in S$.
2. Επαναλαμβάνουμε:
 - Υπολογισμός της νέας λύσης y ,

αν $(f(y) < f(x))$, τότε $x = y$,

έως ότου ικανοποιηθεί το κριτήριο τερματισμού.

3. Επιστρέφουμε το x .

Το βασικό πλεονέκτημα του αλγορίθμου τοπικής αναζήτησης είναι ότι μπορεί να εφαρμοστεί για οποιοδήποτε υπολογιστικά δύσκολο πρόβλημα. Εντούτοις, παρουσιάζει και αρκετά μειονεκτήματα, τα οποία πρέπει να λάβουμε υπόψη. Πρώτον, η λύση που μας επιστρέφει ο αλγόριθμος επηρεάζεται από την αρχική λύση που επιλέγουμε για να ξεκινήσουμε. Ακόμη, εξαιτίας της δομής του αλγόριθμου, είναι προφανές ότι αυτή η λύση ενδέχεται να είναι τοπικά ελάχιστη, χωρίς να μας δίνει καμία πληροφορία για το ολικό ελάχιστο.

Συνεπώς, γίνεται εύκολα αντιληπτό ότι ο αλγόριθμος τοπικής αναζήτησης δεν μπορεί να μας βοηθήσει να πούμε κάτι ακριβές για το ολικό ελάχιστο της αντικειμενικής συνάρτησης. Για να ξεπεραστεί αυτό το πρόβλημα, έχουν προταθεί διάφορες εναλλακτικές. Εμείς θα αναφερθούμε σε μία από τις πιο γενικές προσεγγίσεις του συγκεκριμένου θέματος, σύμφωνα με την οποία αποδεχόμαστε, σε περιορισμένο βαθμό, και κάποιες λύσεις που οδηγούν σε αύξηση της τιμής της αντικειμενικής συνάρτησης. Στην παραπάνω ιδέα βασίζεται ο αλγόριθμος της προσομοιωμένης απόπτωσης, αλλά και ο αλγόριθμος Tabu search.

Ο αλγόριθμος της προσομοιωμένης απόπτωσης, εκτός από τις λύσεις που οδηγούν σε μείωση της αντικειμενικής συνάρτησης, αποδέχεται με κάποια μικρή πιθανότητα και λύσεις που οδηγούν σε αύξησή της. Με αυτόν τον τρόπο, ευελπιστούμε ότι θα αποφύγουμε να εγκλωβιστούμε σε κάποιο τοπικό ελάχιστο.

Ο αλγόριθμος Tabu search βελτιώνει την απόδοση του αλγορίθμου τοπικής αναζήτησης εισάγοντας δομές μνήμης. Δηλαδή, χρησιμοποιώντας τον αλγόριθμο τοπικής αναζήτησης μετακινείται από μία λύση x σε μία λύση y στη γειτονία της x , μέχρι να ικανοποιηθεί το κριτήριο τερματισμού. Παράλληλα, όμως, κατασκευάζεται μία λίστα με μη αποδεκτές λύσεις, η οποία χρησιμοποιείται στις επόμενες επαναλήψεις. Με αυτόν τον τρόπο, επιχειρούμε να διευρύνουμε την εξερεύνηση του χώρου των λύσεων και να αποφύγουμε τον εγκλωβισμό σε περιοχές τοπικών ελαχίστων.

Για περισσότερες πληροφορίες σχετικά με τις υπόλοιπες μεθόδους, που ανήκουν στην κλάση των Monte Carlo τεχνικών βελτιστοποίησης, ο αναγνώστης παραπέμπεται στο βιβλίο του. P.Brandimarte : *Handbook in Monte Carlo Simulation*.

Κεφάλαιο 2

Ο αλγόριθμος της Προσομοιωμένης Ανόπτωσης

2.1 Εισαγωγή στον αλγόριθμο της προσομοιωμένης ανόπτωσης

Όπως έχουμε ήδη αναφέρει, ένας από τους τρόπους προσέγγισης ενός προβλήματος βελτιστοποίησης είναι να εξερευνήσουμε τον χώρο των λύσεων του με στοχαστικό τρόπο. Ο αλγόριθμος της προσομοιωμένης ανόπτωσης (simulated annealing) είναι ένας από τους πιο δημοφιλείς αλγορίθμους αυτής της κατηγορίας. Αρχικά, εισήχθη από τους Kirpatrick et al. (1983) και, ανεξαρτήτως, από τον Cerny (1985) ως μέθοδος επίλυσης συνδυαστικών προβλημάτων βελτιστοποίησης, όπως παραδείγματος χάριν του προβλήματος του πλανόδιου πωλητή.

Ο αλγόριθμος της προσομοιωμένης ανόπτωσης, όπως φανερώνει και η ονομασία του, βασίζεται στη μίμηση μίας φυσικής διαδικασίας, της ανόπτωσης. Το βασικό πλεονέκτημα του αλγορίθμου είναι η ικανότητά του να μην εγκλωβίζεται σε τοπικά βέλτιστα. Αυτό γίνεται εφικτό μέσω της αποδοχής με κάποια πιθανότητα και λύσεων που δεν βελτιώνουν την αντικειμενική συνάρτηση.

2.2 Ανόπτωση

Στα φυσικά συστήματα η ανόπτωση είναι μία διαδικασία, κατά την οποία μεταφέρουμε ένα στερεό σε μία κρυσταλλική κατάσταση χαμηλής ενέργειας. Αυτό επιτυγχάνεται σε δύο στάδια: αρχικά, θερμαίνουμε το στερεό (μέταλλο) σε υψηλή θερμοκρασία και, στη συνέχεια, μειώνουμε τη θερμοκρασία σταδιακά, έως ότου φτάσει κοντά στο μηδέν. Ο ακριβής τρόπος με τον οποίο μειώνεται η θερμοκρασία ονομάζεται χρονοδιάγραμμα ψύξης (cooling schedule).

Όταν το στερεό, λόγω της υψηλής θερμοκρασίας, έχει αποκτήσει υγρή μορφή, τότε τα μόριά του μπορούν να κινηθούν με σχετική ελευθερία. Ωστόσο, καθώς η θερμοκρασία μειώνεται, οι κινήσεις των μορίων γίνονται όλο και πιο περιορισμένες, λόγω του ενεργειακού κόστους. Στην τελική φάση, αναμένουμε τα μόρια του στερεού να βρίσκονται στη χαμηλότερη ενεργειακή κατάσταση. Βεβαίως, απαραίτητη

προϋπόθεση για να φθάσει το σύστημα σε αυτήν την κατάσταση είναι: η θερμοκρασία να μειώνεται με πολύ αργό ρυθμό κατά τη διάρκεια της ψύξης.

Πιο συγκεκριμένα, η διαδικασία της ψύξης του μετάλλου γίνεται με τον ακόλουθο τρόπο: ξεκινώντας από τη μέγιστη θερμοκρασία, αφήνουμε το στερεό να προσεγγίσει την κατάσταση ισορροπίας που αντιστοιχεί σε αυτήν τη θερμοκρασία. Στη συνέχεια, σύμφωνα με το χρονοδιάγραμμα ψύξης, μειώνουμε τη θερμοκρασία και επαναλαμβάνουμε την ίδια διαδικασία. Όταν το στερεό προσεγγίσει την κατάσταση θερμοδυναμικής ισορροπίας για τη θερμοκρασία T , τότε από τη στατιστική μηχανική είναι γνωστό ότι η πιθανότητα εύρεσης σε μία κατάσταση x με ενέργεια $E(x)$ είναι ανάλογη της κατανομής Boltzmann,

$$P(E(x)) = \frac{1}{Z(T)} \exp\left(\frac{-E(x)}{kT}\right)$$

όπου k είναι η σταθερά Boltzmann, ενώ η $Z(T)$ είναι μία σταθερά κανονικοποίησης, η οποία δίνεται από τη σχέση

$$Z(T) = \sum_{x \in S} \exp\left(\frac{-E(x)}{kT}\right).$$

Παρατηρούμε ότι η κατανομή για οποιαδήποτε θερμοκρασία T είναι μία γνησίως φθίνουσα συνάρτηση ως προς την ενέργεια $E(x)$. Αυτό σημαίνει ότι η πιθανότητα εύρεσης ενός συστήματος σε καταστάσεις χαμηλής ενέργειας είναι μεγαλύτερη από την πιθανότητα εύρεσης σε καταστάσεις υψηλής ενέργειας.

Όσον αφορά τη θερμοκρασία παρατηρούμε τα εξής: για υψηλές τιμές της θερμοκρασίας είναι εξίσου πιθανό το σύστημα να βρίσκεται είτε σε καταστάσεις υψηλής ενέργειας, είτε σε καταστάσεις χαμηλής ενέργειας. Ωστόσο, καθώς μειώνεται η θερμοκρασία, είναι πιθανότερο το σύστημα να βρίσκεται σε καταστάσεις χαμηλής ενέργειας. Ενώ, όταν η θερμοκρασία γίνει πολύ χαμηλή $T \rightarrow 0$, τότε η κατανομή Boltzmann τείνει να συγκεντρωθεί στις καταστάσεις που ελαχιστοποιούν την ενέργεια.

Έστω E μία συνάρτηση ενέργειας ορισμένη σε έναν πεπερασμένο χώρο καταστάσεων S , την οποία θέλουμε να ελαχιστοποιήσουμε. Συμβολίζουμε με M το σύνολο των ολικών ελαχίστων της E , δηλαδή

$$M = \{x \in S : E(x) = E^*\}$$

όπου

$$E^* = \min_{x \in S} E(x).$$

Η βασική ιδιότητα της κατανομής Boltzmann, όπως έχουμε αναφέρει και προηγουμένως, είναι ότι καθώς $T \rightarrow 0$, τότε η κατανομή τείνει να συγκεντρωθεί στο σύνολο M . Αυτή η ιδιότητα μπορεί εύκολα να καταδειχθεί, σύμφωνα με την ακόλουθη διαδικασία:

Θεωρούμε την κατανομή Boltzmann

$$P(E(x)) = \frac{\exp\left(\frac{-E(x)}{T}\right)}{\sum_{x \in S} \exp\left(\frac{-E(x)}{T}\right)}$$

και, πολλαπλασιάζοντας αριθμητή και παρονομαστή με $\exp\left(\frac{E^*}{T}\right)$, προκύπτει ότι

$$\begin{aligned} P(E(x)) &= \frac{\exp\left(-\frac{1}{T}(E(x) - E^*)\right)}{\sum_{x \in S} \exp\left(-\frac{1}{T}(E(x) - E^*)\right)} \\ &= \frac{\exp\left(-\frac{1}{T}(E(x) - E^*)\right)}{\sum_{x \in M} \exp\left(-\frac{1}{T}(E(x) - E^*)\right) + \sum_{x \notin M} \exp\left(-\frac{1}{T}(E(x) - E^*)\right)} \\ &= \frac{\exp\left(-\frac{1}{T}(E(x) - E^*)\right)}{|M| + \sum_{x \notin M} \exp\left(-\frac{1}{T}(E(x) - E^*)\right)} \end{aligned}$$

καθώς ισχύει $E(x) - E^* > 0$, όταν $x \notin M$, τότε για $T \rightarrow 0$,

$$P(E(x)) \rightarrow \frac{\delta(x, M)}{|M|}$$

όπου

$$\delta(x, M) = \begin{cases} 1 & x \in M \\ 0 & x \notin M \end{cases}$$

Άρα, συνοψίζοντας, βλέπουμε ότι, αφήνοντας το στερεό να προσεγγίσει την κατάσταση ισορροπίας για κάθε ενδιάμεση θερμοκρασία, αναμένουμε τελικά να έχει προσεγγίσει μία κατάσταση ελάχιστης ενέργειας.

2.3 Αλγόριθμος Metropolis

Βασισμένοι στην κατανομή Boltzmann, οι Metropolis et. al (1953) πρότειναν έναν απλό αλγόριθμο για την προσέγγιση της κατάστασης ισορροπίας ενός συστήματος σε μία δεδομένη θερμοκρασία. Αυτό επιτυγχάνεται μέσω της προσομοίωσης της συμπεριφοράς του συστήματος, με τον ακόλουθο τρόπο: ας υποθέσουμε ότι η τρέχουσα κατάσταση x_0 ενός συστήματος έχει ενέργεια $E(x_0)$. Τότε, εφαρμόζοντας μία τυχαία διαταραχή στην τρέχουσα κατάσταση του συστήματος, προκύπτει μία νέα κατάσταση x_1 με ενέργεια $E(x_1)$. Αν $E(x_1) < E(x_0)$, τότε το σύστημα μεταβαίνει στην κατάσταση x_1 . Ενώ, αν $E(x_1) \geq E(x_0)$, τότε η κατάσταση x_1 γίνεται αποδεκτή με πιθανότητα

$$\exp\left(-\frac{E(x_1) - E(x_0)}{kT}\right).$$

Αλλιώς, το σύστημα παραμένει στην κατάσταση x_0 . Έπειτα από έναν μεγάλο αριθμό επαναλήψεων της παραπάνω διαδικασίας αναμένουμε το σύστημα να έχει προσεγγίσει την κατάσταση ισορροπίας για τη συγκεκριμένη θερμοκρασία.

Ο αλγόριθμος Metropolis μπορεί επίσης να χρησιμοποιηθεί ως μέθοδος επίλυσης συνδυαστικών προβλημάτων βελτιστοποίησης. Σε αυτή την περίπτωση, χρησιμοποιούμε τον αλγόριθμο για να παράγουμε την ακολουθία των εφικτών λύσεων του προβλήματος βελτιστοποίησης. Πιο συγκεκριμένα: οι καταστάσεις του στερεού αντιστοιχούν στις εφικτές λύσεις του προβλήματος βελτιστοποίησης. Ενώ, από την άλλη πλευρά, η ενέργεια και η θερμοκρασία αντιστοιχούν στην αντικειμενική συνάρτηση και στην παράμετρο ελέγχου T .

2.4 Περιγραφή του αλγορίθμου της προσομοιωμένης ανόπτωσης

Η ιδέα, πάνω στην οποία βασίζεται ο αλγόριθμος της προσομοιωμένης ανόπτωσης, είναι η σύνδεση μεταξύ της διαδικασίας της ανόπτωσης και ενός συνδυαστικού προβλήματος βελτιστοποίησης. Με άλλα λόγια, οι Kirpatrick et. al. σκέφτηκαν ότι, εκτελώντας τον αλγόριθμο Metropolis, ενώ σταδιακά μειώνεται η θερμοκρασία, μπορούν να μιμηθούν υπολογιστικά την διαδικασία της ανόπτωσης, για να λύσουν προβλήματα βελτιστοποίησης. Ο τρόπος με τον οποίο μειώνεται η ακολουθία των τιμών της θερμοκρασίας ονομάζεται χρονοδιάγραμμα ψύξης (cooling schedule), και

η κατάλληλη επιλογή χρονοδιαγράμματος είναι εξαιρετικά σημαντική για την αποτελεσματικότητα της μεθόδου.

Σε αυτή την παράγραφο περιγράφουμε τον τρόπο με τον οποίο κατασκευάζουμε μία ακολουθία λύσεων X_0, X_1, \dots , μέσω του αλγορίθμου της προσομοιωμένης ανόπτωσης. Έστω f μία συνάρτηση ορισμένη στο χώρο λύσεων S , την οποία θέλουμε να ελαχιστοποιήσουμε. Υποθέτουμε ότι για κάθε στοιχείο x του S ορίζεται το σύνολο $N(x)$ των γειτονικών λύσεων. Επιπλέον, υποθέτουμε ότι υπάρχει ένας πίνακας πιθανοτήτων μετάβασης R τέτοιος, ώστε $R(x, y) > 0$, αν και μόνο αν $y \in N(x)$.

Σε μία τυπική εκτέλεση του αλγορίθμου της προσομοιωμένης ανόπτωσης ξεκινάμε από μία πολύ υψηλή αρχική θερμοκρασία. Ενώ, η εξέλιξη της ακολουθίας των τιμών της θερμοκρασίας $\{T_m\}$ ορίζεται έτσι ώστε:

$$T_1 \geq T_2 \geq \dots \geq T_m$$

$$\lim_{m \rightarrow \infty} T_m = 0.$$

Έτσι, λοιπόν, ξεκινάμε από μία τυχαία λύση x από το σύνολο S , την οποία αποθηκεύουμε ως τρέχουσα λύση. Στη συνέχεια, για κάθε τιμή T_m της θερμοκρασίας επαναλαμβάνουμε N_m φορές τα εξής βήματα: δεδομένου ότι η τρέχουσα λύση στο k -οστό βήμα είναι $X_k = x$, επιλέγουμε μία υποψήφια λύση $Y_k = y$ από το $N(x)$ με κατανομή $P[Y_k = y | X_k = x] = R(x, y)$. Η υποψήφια λύση $Y_k = y$ γίνεται δεκτή με πιθανότητα

$$\alpha_{T_m}(x, y) = \min \left\{ 1, \exp \left(\frac{-\Delta f}{T_m} \right) \right\}$$

όπου

$$\Delta f = f(y) - f(x).$$

Στη συνέχεια, θέτουμε

$$X_{k+1} = \begin{cases} Y_k & \text{με πιθανότητα } \alpha_{T_m} \\ X_k & \text{αλλιώς} \end{cases}$$

Με αυτόν τον τρόπο, για κάθε λύση $x \in S$ και για κάθε m ορίζεται ένας στοχαστικός πίνακας P_{T_m} με τις εξής πιθανότητες μετάβασης πρώτης τάξης:

$$P_{T_m}(x, y) = \begin{cases} 0 & y \notin N(x) \text{ και } y \neq x \\ R(x, y) \alpha_{T_m}(x, y) & y \in N(x) \text{ και } y \neq x \\ 1 - \sum_{z \neq x} P_{T_m}(x, z) & y = x \end{cases} \quad (2.1)$$

Η μη-ομογενής Μαρκοβιανή αλυσίδα που παράγεται από τον αλγόριθμο καθορίζεται από τις παραπάνω πιθανότητες μετάβασης. Μπορεί να δειχθεί ότι, αν κατά την εκτέλεση του αλγορίθμου μειώνουμε την θερμοκρασία σύμφωνα με ένα κατάλληλα επιλεγμένο χρονοδιάγραμμα ψύξης, τότε μπορούμε να προσεγγίσουμε ασυμπτωτικά το ολικό ελάχιστο της συνάρτησης f .

Η θερμοκρασία είναι μία πολύ σημαντική παράμετρος του αλγορίθμου, καθώς συνδέεται άμεσα με την πιθανότητα αποδοχής μίας υποψήφιας λύσης. Η πιθανότητα αποδοχής είναι ίση με 1, όταν $\Delta f < 0$. Δηλαδή, όταν μία υποψήφια λύση βελτιώνει την αντικειμενική συνάρτηση, γίνεται πάντα δεκτή. Ωστόσο, στην περίπτωση που έχουμε ανηφορική κίνηση (uphill move) $\Delta f > 0$, τότε η υποψήφια λύση γίνεται δεκτή με κάποια πιθανότητα που εξαρτάται από την παράμετρο ελέγχου T . Για μεγάλες τιμές του T υπάρχει σχετικά μεγάλη πιθανότητα να γίνει δεκτή μία ανηφορική κίνηση. Καθώς όμως μειώνεται η τιμή του T , μειώνεται η πιθανότητα αποδοχής μίας ανηφορικής κίνησης και καταλήγουμε, για τιμές του T κοντά στο 0, οι κινήσεις προς τα επάνω να μην γίνονται σχεδόν ποτέ αποδεκτές.

Αλγόριθμος 2.1 προσομοιωμένη απόπτωση

1. Επιλέγουμε μία αρχική λύση $X_{(0)} = x$; $x \in S$.
2. Θέτουμε $m = 0$.
3. Επιλέγουμε το χρονοδιάγραμμα ψύξης T_m .
4. Επιλέγουμε μία αρχική θερμοκρασία $T = T_0$.
5. Επιλέγουμε το N_m , δηλαδή το πλήθος των επαναλήψεων που εκτελούμε τον αλγόριθμο για κάθε T_m .
6. Επαναλαμβάνουμε (όσο δεν ικανοποιείται το κριτήριο τερματισμού).
 - 6.1. Θέτουμε $k = 0$.
 - 6.2. Επαναλαμβάνουμε (όσο $k \leq N_m$).
 - 6.3. Παράγουμε μία λύση $y \in N(x)$.
 - 6.4. Υπολογίζουμε το $\Delta f = f(y) - f(x)$.
 - 6.5. Εάν $\Delta f \leq 0$, τότε $x = y$.
 - 6.6. Εάν $\Delta f > 0$, τότε παράγουμε έναν τυχαίο αριθμό $u \sim U(0,1)$.

6.7. Εάν $\exp\left(\frac{-\Delta f}{T_m}\right) > u$, τότε $x = y$.

6.8. Αλλιώς, $x = x$.

6.9. $k = k + 1$.

Τέλος επανάληψης.

7. Θέτουμε $m = m + 1$.

Τέλος επανάληψης.

8. Επιστρέφουμε το x .

Για περισσότερες πληροφορίες σχετικά με τον αλγόριθμο της προσομοιωμένης ανόπτωσης, ο αναγνώστης παραπέμπεται στα συγγράμματα M. Gendreau, J-Y Potvin (2010), Jun S. Liu (2004), O. Haggstrom (2002).

2.5 Σύγκλιση του αλγορίθμου της προσομοιωμένης ανόπτωσης

Ένα από τα βασικά θεωρητικά ζητήματα που προκύπτουν από τη μελέτη του αλγορίθμου είναι το κατά πόσο και κάτω από ποιες προϋποθέσεις εξασφαλίζεται η σύγκλιση του αλγορίθμου, στο σύνολο των βέλτιστων λύσεων. Στη βιβλιογραφία, υπάρχουν διάφορα αποτελέσματα, σχετικά με αυτό το θέμα (βλ., μεταξύ άλλων, Geman and Geman (1984), Mitra et al. (1986), Hajek (1988)) που βασίζονται στην θεωρία των Μαρκοβιανών Αλυσίδων (MA). Έτσι, λοιπόν, στις επόμενες παραγράφους δίνεται ένα σχέδιο της απόδειξης των Mitra et al. (1986).

Αρχικά, θα πρέπει να παρατηρήσουμε ότι, αν κατά τη διάρκεια εκτέλεσης του αλγορίθμου η θερμοκρασία T_m είναι σταθερή και ίση με T , τότε η MA που προκύπτει είναι ομογενής.

Επιπλέον, αν υποθέσουμε ότι ο πίνακας πιθανοτήτων μετάβασης R είναι συμμετρικός, δηλαδή

$$R(x, y) = R(y, x) \quad \forall x, y \in S \quad (2.2),$$

τότε η MA που προκύπτει είναι αντιστρέψιμη, με στάσιμη κατανομή την

$$\pi_T(x) = \frac{\exp(-f(x)/T)}{Z_T} \quad (2.3),$$

όπου

$$Z_T = \sum_{x \in S} \exp\left(\frac{-f(x)}{T}\right).$$

Για να το αποδείξουμε αυτό, αρκεί να δείξουμε ότι $\forall x, y$ ισχύει η ακόλουθη σχέση:

$$\pi_T(x)P(x, y) = \pi_T(y)P(y, x). \quad (2.4)$$

Δεδομένου ότι για την πιθανότητα αποδοχής ισχύει

$$a(x, y) = \min\left\{1, \exp\left(\frac{-(f(y) - f(x))}{T}\right)\right\}$$

τότε, αν $a(x, y) < 1$, έχουμε

$$\exp\left(\frac{-(f(x) - f(y))}{T}\right) > 1$$

κι έτσι προκύπτει ότι $a(y, x) = 1$.

Με τον ίδιο τρόπο, μπορεί να δειχθεί ότι, όταν $a(y, x) < 1$, τότε $a(x, y) = 1$.

Συνεπώς, αν υποθέσουμε ότι $a(x, y) < 1$ και $a(y, x) = 1$, τότε έχουμε

$$a(x, y) = \exp\left(\frac{-(f(y) - f(x))}{T}\right) = \frac{\pi_T(y)}{\pi_T(x)}$$

Πολλαπλασιάζοντας με $R(x, y)$ και, λόγω της (2.2), προκύπτει ότι:

$$\begin{aligned} a(x, y)R(x, y)\pi_T(x) &= \pi_T(y)R(y, x)a(y, x) \\ \pi_T(x)P(x, y) &= \pi_T(y)P(y, x). \end{aligned}$$

Συμπερασματικά, η MA είναι αντιστρέψιμη ως προς την κατανομή π_T . Επομένως, αυτή είναι, επίσης, και η στάσιμη κατανομή της. Επιπλέον, η στάσιμη κατανομή είναι και οριακή κατανομή, αν η MA είναι αδιαχώριστη και απεριοδική. Στην περίπτωση που τα ενδεχόμενα παραμονής στην ίδια κατάσταση έχουν θετική πιθανότητα, τότε η MA είναι απεριοδική. Ενώ, για να εξασφαλίσουμε ότι η MA που παράγεται από τον αλγόριθμο είναι αδιαχώριστη, αρκεί να επιλέξουμε έναν πίνακα πιθανοτήτων μετάβασης R μίας αδιαχώριστης MA.

Έτσι λοιπόν, για κάθε πίνακα πιθανοτήτων μετάβασης P_{T_m} υπάρχει η αντίστοιχη στάσιμη κατανομή π_{T_m} . Επιπλέον, όπως είδαμε και στην παράγραφο 2.2, η π_{T_m} συγκλίνει στο διάνυσμα των βέλτιστων λύσεων e^* , καθώς $T_m \rightarrow 0$, δηλαδή

$$e_x^* = \lim_{T_m \rightarrow 0} \pi_{T_m}(x) = \begin{cases} |S_{opt}|^{-1} & \text{αν } x \in S_{opt} \\ 0 & \text{αλλιώς} \end{cases},$$

όπου S_{opt} το σύνολο των βέλτιστων λύσεων. Τα υπόλοιπα αποτελέσματα βασίζονται στη θεωρία των μη ομογενών Μαρκοβιανών αλυσίδων. Πιο συγκεκριμένα,

χρησιμοποιείται το θεώρημα A.4 για να δειχθεί ότι η MA είναι ισχυρά εργοδική και, κατά συνέπεια, αποδεικνύεται η σύγκλιση του αλγορίθμου στη βέλτιστη λύση.

Επομένως, δεδομένου ότι για κάθε P_{T_m} υπάρχει η αντίστοιχη στάσιμη κατανομή π_{T_m} , τότε, σύμφωνα με το θεώρημα A.4, για να είναι η MA ισχυρά εργοδική πρέπει να ικανοποιούνται οι ακόλουθες συνθήκες:

1. Η MA πρέπει να είναι ασθενώς εργοδική.
2. Για κάθε π_{T_m} πρέπει να ισχύει:

$$\sum_{T_m=0}^{\infty} \|\pi_{T_m} - \pi_{T_{m+1}}\| < \infty \quad (2.5)$$

Ασθενής εργοδικότητα της MA

Η ασθενής εργοδικότητα της MA αποδεικνύεται μέσω του θεωρήματος A.3. Αρχικά, κατασκευάζεται ένα κάτω φράγμα για τις πιθανότητες μετάβασης του αλγορίθμου σε οποιαδήποτε λύση ξεκινώντας από οποιαδήποτε λύση δεν αποτελεί τοπικό μέγιστο. Ενώ, στη συνέχεια, κατασκευάζεται ένα άνω φράγμα για τον συντελεστή εργοδικότητας χρησιμοποιώντας το φράγμα των πιθανοτήτων μετάβασης. Οπότε είναι, πλέον, δυνατόν με βάση το θεώρημα A.3, να προσδιοριστούν τις απαραίτητες συνθήκες που εξασφαλίζουν την ασθενή εργοδικότητα της MA.

Σε αυτό το σημείο, δίνονται ορισμοί κάποιων ποσοτήτων που χρειάζονται για τη συνέχεια της ανάλυσης μας. Πιο συγκεκριμένα, ορίζονται τα w, r, S_{max} και L ως εξής:

Το S_{max} είναι το σύνολο των τοπικών μεγίστων της αντικειμενικής συνάρτησης

$$S_{max} = \{x \in S \mid f(y) \leq f(x) \forall y \in N(x)\}.$$

Για το r ισχύει ότι

$$r = \min_{x \in S \setminus S_{max}} \max_{y \in S} d(x, y), \quad (2.6)$$

δηλαδή το r είναι ο ελάχιστος αριθμός μεταβάσεων που απαιτούνται, για να μεταβούμε από μία λύση που δεν ανήκει στο σύνολο των τοπικών μεγίστων της αντικειμενικής συνάρτησης σε οποιαδήποτε άλλη λύση. Η λύση στην οποία επιτυγχάνεται το ελάχιστο στην (2.6) συμβολίζεται με \hat{x} .

Το L αποτελεί ένα άνω φράγμα της τοπικής κλίσης της αντικειμενικής συνάρτησης

$$L = \max_{x \in S} \max_{y \in N(x)} |f(y) - f(x)|, \quad (2.7)$$

ενώ για το w ισχύει ότι

$$w = \min_{x \in S} \min_{y \in N(x)} \{R(x, y): R(x, y) > 0\}. \quad (2.8)$$

Έτσι λοιπόν, σε αυτό το σημείο είναι δυνατόν με χρήση των παραπάνω ποσοτήτων να κατασκευασθεί ένα κάτω φράγμα για τις πιθανότητες μετάβασης του αλγορίθμου που ορίστηκαν στην (2.1).

Στην περίπτωση που $y \in N(x)$, τότε από τις 2.1, 2.7, 2.8 προκύπτει ότι,

$$P_{T_m}(x, y) \geq w \exp(-L/T_m) \quad (2.9)$$

Όσον αφορά, τα διαγώνια στοιχεία του πίνακα πιθανοτήτων μετάβασης $P_{T_m}(x, x)$, $x \in S - S_{max}$, πρέπει να παρατηρήσουμε ότι οι τιμές των συγκεκριμένων όρων αυξάνονται καθώς μειώνεται η θερμοκρασία. Έστω ότι βρισκόμαστε στην κατάσταση x , τότε

$$P_{T_m}(x, x) = 1 - \sum_{z \neq x} P_{T_m}(x, z).$$

Το παραπάνω άθροισμα χωρίζεται σε δύο μέρη: α) το άθροισμα των πιθανοτήτων μετάβασης από την κατάσταση x σε καταστάσεις που βελτιώνουν την αντικειμενική συνάρτηση β) το άθροισμα των πιθανοτήτων μετάβασης από την κατάσταση x σε καταστάσεις που δεν βελτιώνουν την αντικειμενική συνάρτηση. Από την δομή του αλγορίθμου είναι γνωστό ότι το πρώτο άθροισμα είναι σταθερό και ανεξάρτητο του T_m . Από την άλλη, το δεύτερο άθροισμα, μειώνεται καθώς αυξάνεται η θερμοκρασία. Αυτό συνεπάγεται ότι οι τιμές των όρων $P_{T_m}(x, x)$ αυξάνονται καθώς μειώνεται η θερμοκρασία.

Ως εκ τούτου, $\exists k_0, k_0 < \infty$ τέτοιο ώστε $\forall x \in S - S_{max}$,

$$P_{T_m}(x, x) \geq w \exp(-L/T_m) \quad m \geq k_0 r \quad (2.10)$$

αφού το αριστερό μέρος της ανίσωσης αυξάνεται, ενώ από την άλλη το δεξί μέρος μειώνεται καθώς αυξάνεται η θερμοκρασία.

Ενώ, συνδυάζοντας τις 2.9 και 2.10 προκύπτει ότι $\forall x \in S$ και $\forall m \geq k_0 r$

$$\begin{aligned} P^{(m-r, m)}(x, \hat{x}) &\geq \prod_{n=m-r}^{m-1} \{w \exp(-L/T_n)\} \\ &\geq w^r \exp(-rL/T_{m-1}). \end{aligned} \quad (2.11)$$

Οπότε, πλέον έχει κατασκευαστεί ένα κάτω φράγμα για τις πιθανότητες μετάβασης σε οποιαδήποτε λύση ξεκινώντας από οποιαδήποτε λύση δεν αποτελεί τοπικό μέγιστο.

Χρησιμοποιώντας την (2.11) βλέπουμε ότι για τον συντελεστή εργοδικότητας ισχύει η ακόλουθη σχέση:

$$\begin{aligned}\tau_1(\mathbf{P}^{(kr-r,r)}) &\leq 1 - \min_{x,y} \{\min(P(x, \hat{x}), P(y, \hat{x}))\} \\ &\leq 1 - w^r \exp(-rL/T_{kr-1}), k \geq k_0. \quad (2.12)\end{aligned}$$

Επομένως, σύμφωνα με το θεώρημα A.2 και την (2.7), για να είναι η MA ασθενώς εργοδική πρέπει να ισχύει

$$\sum_{k=k_0}^{\infty} \exp(-rL/T_{kr-1}) = \infty. \quad (2.13)$$

Στο ακόλουθο θεώρημα οι Mitra et al. προσδιορίζουν την ακριβή μορφή του χρονοδιαγράμματος ψύξης, ώστε να ικανοποιείται η (2.13).

Θεώρημα 2.1

Η MA που προκύπτει από τον αλγόριθμο της προσομοιωμένης ανόπτησης με το ακόλουθο χρονοδιάγραμμα ψύξης

$$T_m = \frac{\gamma}{\log(m + m_0 + 1)} \quad (2.14),$$

όπου $m = 0, 1, 2, \dots$ και $1 \leq m_0 < \infty$ είναι ασθενώς εργοδική, αν

$$\gamma \geq rL.$$

Απόδειξη: Αν αντικαταστήσουμε στην (2.12) το T_m με την (2.14), τότε προκύπτει ότι

$$\begin{aligned}\tau_1(\mathbf{P}^{(kr-r,r)}) &\leq 1 - w^r \exp\left(\frac{-rL}{\gamma} \log(kr + m_0)\right) \\ &= 1 - w^r (kr + m_0)^{\frac{-rL}{\gamma}} = 1 - \frac{a}{(k + \frac{m_0}{r})^\mu}\end{aligned}$$

όπου $\mu = \frac{rL}{\gamma}$, $a = \frac{w^r}{r^\mu}$.

Έτσι, λοιπόν, από το θεώρημα A.2 προκύπτει ότι η MA είναι ασθενώς εργοδική, δεδομένου ότι για $\mu \leq 1$ ισχύει:

$$\sum_{k=l}^{\infty} \frac{a}{(k + \frac{m_0}{r})^\mu} = \infty.$$

Οπότε, για να είναι πλέον η MA ισχυρά εργοδική, το μόνο που πρέπει να δειχθεί είναι ότι ισχύει η (2.5). Στην απόδειξή τους οι Mitra et al.(1986) χρησιμοποιούν την ακόλουθη πρόταση, σχετικά με τη μονοτονία της στάσιμης κατανομής (2.3).

Πρόταση 2.1

- i. Για κάθε $x \in S_{opt}$ ισχύει ότι $\pi_{T_{m+1}}(x) - \pi_{T_m}(x) > 0, \forall m > 0$.
- ii. Για κάθε $x \notin S_{opt}$ υπάρχει $\tilde{m}, \tilde{m} < \infty$ τέτοιο, ώστε για κάθε $x \notin S_{opt}$

$$\pi_{T_{m+1}}(x) - \pi_{T_m}(x) < 0, \quad \forall m > \tilde{m}$$

Ξεκινώντας από την $\|\pi_{T_{m+1}} - \pi_{T_m}\|$ έχουμε ότι,

$$\|\pi_{T_{m+1}} - \pi_{T_m}\| = \sum_{x \in S_{opt}} |\pi_{T_{m+1}}(x) - \pi_{T_m}(x)| + \sum_{x \notin S_{opt}} |\pi_{T_{m+1}}(x) - \pi_{T_m}(x)|,$$

Ενώ, χρησιμοποιώντας την πρόταση 2.1, έχουμε ότι $\forall m > \tilde{m}$

$$\|\pi_{T_{m+1}} - \pi_{T_m}\| = \sum_{x \in S_{opt}} \{\pi_{T_{m+1}}(x) - \pi_{T_m}(x)\} - \sum_{x \notin S_{opt}} \{\pi_{T_{m+1}}(x) - \pi_{T_m}(x)\}.$$

Επίσης, λόγω του ότι η π_{T_m} είναι κατανομή πιθανότητας, έχουμε

$$\sum_{x \in S_{opt}} \pi_{T_m}(x) + \sum_{x \notin S_{opt}} \pi_{T_m}(x) = 1 \quad \forall T_m > 0. \quad (2.15)$$

οπότε, χρησιμοποιώντας την (2.15), καταλήγουμε στην επόμενη σχέση

$$\|\pi_{T_{m+1}} - \pi_{T_m}\| = 2 \left\{ \sum_{x \in S_{opt}} \{\pi_{T_{m+1}}(x) - \pi_{T_m}(x)\} \right\}, \quad \forall m > \tilde{m} \quad (2.16)$$

άρα

$$\sum_{T_m = T_{\tilde{m}}} \|\pi_{T_{m+1}} - \pi_{T_m}\| = 2 \left\{ \sum_{T_m = T_{\tilde{m}}} \sum_{x \in S_{opt}} \pi_{T_{m+1}}(x) - \sum_{T_m = T_{\tilde{m}}} \sum_{x \in S_{opt}} \pi_{T_m}(x) \right\} \leq 2.$$

Έτσι, λοιπόν, συνοψίζοντας τα παραπάνω, η ΜΑ που παράγεται από τον αλγόριθμο της προσομοιωμένης ανόπτησης ικανοποιεί τις εξής συνθήκες:

- i. Για κάθε T_m υπάρχει η αντίστοιχη στάσιμη κατανομή π_{T_m}
- ii. Η Μ.α. είναι ασθενώς εργοδική υπό τις προϋποθέσεις του θεωρήματος 2.1.
- iii. $\sum_{T_m = T_{\tilde{m}}} \|\pi_{T_{m+1}} - \pi_{T_m}\| < \infty$

Επομένως, από το θεώρημα Α.3 προκύπτει ότι η ΜΑ είναι ισχυρά εργοδική. Επιπλέον, δεδομένου ότι

$$e^* = \lim_{T_m \rightarrow 0} \pi_{T_m}$$

τότε για κάθε T_m

$$\lim_{n \rightarrow \infty} \sup_{g(0)} \|g^{(m,n)} - e^*\| = 0,$$

όπου $g(0)$ μία αρχική κατανομή και $g^{(m,n)} = g(0) \prod_{i=0}^{n-1} P_{(m+i)}$. Συνεπώς, η ΜΑ που παράγεται από τον αλγόριθμο συγκλίνει στο σύνολο των βέλτιστων λύσεων.

2.6 Χρονοδιάγραμμα ψύξης

Στις ακόλουθες παραγράφους θα δώσουμε μία αναλυτική περιγραφή του χρονοδιαγράμματος ψύξης του αλγορίθμου της προσομοιωμένης ανόπτωσης και θα παρουσιάσουμε κάποιες από τις επιλογές των παραμέτρων του.

Το χρονοδιάγραμμα ψύξης καθορίζεται πλήρως από τις ακόλουθες παραμέτρους:

1. Αρχική τιμή της θερμοκρασίας T_0 .
2. Τελική τιμή της θερμοκρασίας.
3. Έναν κανόνα μετάβασης από την τρέχουσα θερμοκρασία T_m στην επόμενη, την T_{m+1} .
4. Αριθμός των επαναλήψεων του αλγορίθμου που εκτελούνται σε κάθε θερμοκρασία T_k .

Η αρχική θερμοκρασία επιλέγεται έτσι, ώστε $\forall x, y \in S$ να ισχύει $\exp\left(-\frac{-\Delta f(x,y)}{T_0}\right) \approx 1$. Επιθυμούμε, δηλαδή, να υπάρχει η δυνατότητα να επισκεφθούμε κάθε σημείο του χώρου λύσεων.

Μία από τις πιο απλές μεθόδους για τον καθορισμό της αρχικής θερμοκρασίας είναι η ακόλουθη εμπειρική μέθοδος, η οποία προτάθηκε από τους Kirpatrick et al: Αρχικά, επιλέγουμε μία υψηλή τιμή για τη θερμοκρασία και πραγματοποιούμε έναν αρκετά μεγάλο αριθμό μεταβάσεων. Στη συνέχεια, υπολογίζουμε τον λόγο του αριθμού των μεταβάσεων που έγιναν αποδεκτές προς τον αριθμό των μεταβάσεων που προτάθηκαν. Αν ο παραπάνω λόγος είναι μικρότερος από έναν προκαθορισμένο αριθμό x_0 κοντά στη μονάδα, π.χ. οι Kirpatrick et al.(1983) χρησιμοποιούν $x_0 = 0.8$, τότε διπλασιάζουμε τη θερμοκρασία. Η παραπάνω διαδικασία επαναλαμβάνεται μέχρι ο λόγος να υπερβεί το x_0 .

Συνήθως, ως κριτήριο τερματισμού επιλέγουμε κάποια θερμοκρασία πολύ κοντά στο μηδέν. Εναλλακτικά, ο αλγόριθμος τερματίζεται, εφόσον δεν παρατηρείται βελτίωση στη λύση για κάποιον αριθμό διαφορετικών θερμοκρασιών.

Δεδομένου ότι οι λογάριθμοι είναι συναρτήσεις που αυξάνονται με πολύ αργό ρυθμό, αν χρησιμοποιήσουμε το λογαριθμικό χρονοδιάγραμμα ψύξης, θα πρέπει να περιμένουμε πάρα πολύ μέχρι η θερμοκρασία να προσεγγίσει το μηδέν. Γι' αυτόν τον λόγο, συνήθως χρησιμοποιούμε εναλλακτικούς κανόνες μείωσης της θερμοκρασίας.

Ένας από πιο διαδεδομένους κανόνες ψύξης είναι ο γεωμετρικός, όπου η θερμοκρασία μειώνεται σύμφωνα με το ακόλουθο σχήμα:

$$T_m = \alpha T_{m-1}$$

Εδώ, το α παίρνει τιμές μεταξύ 0,8 και 0,99.

Όσον αφορά τον αριθμό των επαναλήψεων, μία προφανής επιλογή είναι ένας σταθερός αριθμός για κάθε T_k . Σχετικά δε με τον ακριβή προσδιορισμό αυτού του σταθερού αριθμού, ίσως χρειαστεί να πειραματιστούμε με διάφορες τιμές. Βεβαίως, θα πρέπει να σημειώσουμε ότι ο αριθμός των επαναλήψεων επηρεάζεται από τον κανόνα μείωσης της θερμοκρασίας που χρησιμοποιούμε. Αυτό συμβαίνει, διότι για κάθε θερμοκρασία παράγεται μία Μαρκοβιανή αλυσίδα, που επιθυμούμε να προσεγγίσει την κατάσταση ισορροπίας. Επομένως, αν το βήμα μείωσης της θερμοκρασίας είναι μικρό, αναμένουμε ότι δεν χρειάζεται μεγάλος αριθμός επαναλήψεων για την προσέγγιση της κατάστασης ισορροπίας στη νέα θερμοκρασία. Δηλαδή, ο αριθμός των επαναλήψεων είναι ανάλογος με τον ρυθμό μείωσης της θερμοκρασίας.

Εναλλακτικά, ο αριθμός των επαναλήψεων για κάθε θερμοκρασία μπορεί να μεταβάλλεται κατά τη διάρκεια εκτέλεσης του αλγορίθμου. Μία λογική προσέγγιση θα ήταν να αυξάνουμε τον αριθμό των επαναλήψεων σε χαμηλότερες θερμοκρασίες, ώστε, σε περίπτωση που βρεθούμε στην περιοχή κάποιου τοπικού ακροτάτου, να έχουμε στη διάθεσή μας τον απαραίτητο χρόνο για να εξερευνήσουμε την περιοχή γύρω του.

Τέλος, θα πρέπει να αναφερθεί ότι στις προηγούμενες παραγράφους παραθέσαμε κάποιες από τις απλούστερες επιλογές για τις παραμέτρους του χρονοδιαγράμματος ψύξης. Όποιος ενδιαφέρεται για πιο εξεζητημένα χρονοδιαγράμματα ψύξης μπορεί, ενδεικτικά, να ανατρέξει στο κεφάλαιο 5 του βιβλίου των Aarts και Laarhoven, *Simulated Annealing: Theory and Applications*.

2.7 Παράδειγμα

Σε αυτή την παράγραφο, θα υλοποιήσουμε τον αλγόριθμο σε ένα απλό παράδειγμα. Έστω ότι θέλουμε να μεγιστοποιήσουμε τη συνάρτηση f με

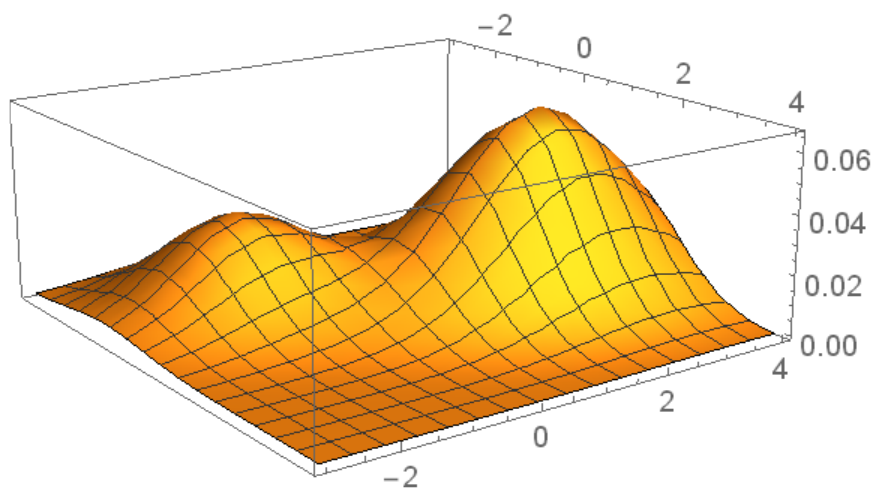
$$f(x_1, x_2) = pf_1(x_1, x_2) + (1 - p)f_2(x_1, x_2)$$

όπου

$$f_1(x_1, x_2) = \frac{1}{2\pi} \exp\left(-\frac{1}{2}[(x_1 + 1)^2 + (x_2 + 1)^2]\right)$$

$$f_2(x_1, x_2) = \frac{1}{\pi\sqrt{15}} \exp\left(-\frac{8}{15}\left[(x_1 - 1)^2 + \frac{(x_2 - 2.5)^2}{2} - \frac{1}{4}(x_1 - 1)(x_2 - 2.5)\right]\right)$$

και $p = 0.2$.



Σχήμα 2.1 Η τρισδιάστατη αναπαράσταση της $f(x)$, όπου διακρίνονται δύο κορυφές.

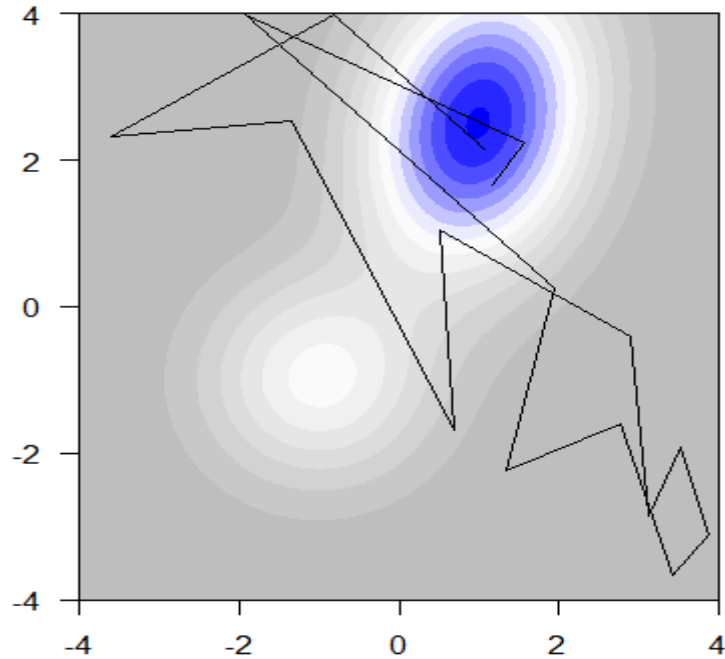
Η παραπάνω συνάρτηση μετά από αριθμητική βελτιστοποίηση, που πραγματοποιήθηκε χρησιμοποιώντας το Mathematica, φαίνεται να παρουσιάζει μέγιστο 0.065759 στο σημείο

(0.999599, 2.498955).

Όμως, όπως φαίνεται και από το σχήμα, η συνάρτηση παρουσιάζει δύο κορυφές. Έτσι, λοιπόν, θα εφαρμόσουμε τον αλγόριθμο 2.1, και θα περιμένουμε να δούμε αν ο αλγόριθμος θα συγκλίνει στην κύρια κορυφή της συνάρτησης ή αν θα καταλήξει στη μικρότερη κορυφή. Βεβαίως, πρέπει να θυμηθούμε ότι ο αλγόριθμος έχει διατυπωθεί για προβλήματα ελαχιστοποίησης, γι' αυτό, πριν τον εφαρμόσουμε, θα χρειαστεί κατάλληλη τροποποίηση. Πιο συγκεκριμένα, θα πρέπει να αντικαταστήσουμε το $-\Delta f$ με Δf και τον έλεγχο αρνητικότητας $\Delta f \leq 0$ με $\Delta f \geq 0$.

Στο ακόλουθο πλαίσιο, παραθέτουμε τον αλγόριθμο της προσομοιωμένης απόκτησης για τη μεγιστοποίηση της f στο χωρίο $[-4,4] \times [-4,4]$ μέσω της R .

```
f<-function(x1,x2){f<-0.2*dnorm(x1,-1,1)*dnorm(x2,-1,1)+0.8*
(exp((-8/15)*((x1-1)^2+((x2-2.5)^2)/2-((x1-1)*(x2-
2.5))/4)))/(pi*sqrt(15))
return(f)}
for(a in seq(0.85,0.99,0.01)){
X1=NULL; X2=NULL;Iterations=NULL;
for(i in 1:1000){
x1=runif(1); x2=runif(1);
T=10; T_end=0.0001; step=0;
while(T>T_end){
  F=f(x1,x2)
  new_x1=x1+rnorm(1,0,2)
  new_x2=x2+rnorm(1,0,2)
  if (abs(new_x1)<=4 & abs(new_x2)<4){
    step=step+1
    new_F=f(new_x1,new_x2)
    dF=new_F-F
    ar=min(1,exp(dF/T))
    u<-runif(1)
    if(u<ar){
      x1=new_x1; x2=new_x2;}}
T=T*a;}
X1[i]=x1; X2[i]=x2; Iterations[i]=step;}
X=cbind(X1,X2);k=0;Y1=NULL;Y2=NULL;
for(i in 1:1000){
  d1=sqrt((X[i,1]+0.955467)^2+(X[i,2]+0.956622)^2)
  d2=sqrt((X[i,1]-0.999599)^2+(X[i,2]-2.498955)^2)
  if(d2<d1){
    k=k+1;
Y1[k]=X[i,1];Y2[k]=X[i,2]}}
Y=cbind(Y1,Y2)
cat("a=",a,"\n");cat("k=",k,"\n");
cat("mean(Iterations)=",mean(Iterations),"\n");
cat("mean(Y[,1])=",mean(Y[,1]),"\n");
cat("sd(Y[,1])=",sd(Y[,1]),"\n")
cat("mean(Y[,2])=",mean(Y[,2]),"\n");cat("sd(Y[,2])=",sd(Y[,2]),"\n"
);
cat("MSE(x1)=",sd(Y[,1])^2+(0.999599-mean(Y[,1]))^2,"\n");
cat("MSE(x2)=",sd(Y[,2])^2+(2.498955-mean(Y[,2]))^2,"\n");
Function<-mapply(function(x1,x2){f<-0.2*dnorm(x1,-1,1)*dnorm(x2,-
1,1)+0.8*(exp((-8/15)*((x1-1)^2+((x2-2.5)^2)/2-((x1-1)*(x2-
2.5))/4)))/(pi*sqrt(15))},Y[,1],Y[,2])
cat("mean(Function)=",mean(Function),"\n");
cat("sd(Function)=",sd(Function),"\n");}
```



Σχήμα 2.2 Στο ακόλουθο διάγραμμα ισοϋψών καμπυλών βλέπουμε την ακολουθία των λύσεων μίας τυχαίας πραγματοποίησης του αλγορίθμου για $\alpha = 0.85$.

Έτσι, λοιπόν, τρέξαμε τον αλγόριθμο 2.1 1000 φορές έχοντας επιλέξει τις εξής παραμέτρους: α) Η θερμοκρασία ξεκινά από το 10 και καταλήγει στο 0.0001 και β) για τη μείωση της θερμοκρασίας χρησιμοποιούμε το γεωμετρικό χρονοδιάγραμμα ψύξης. Όσον αφορά την επιλογή των αρχικών τιμών, αυτή γίνεται με τυχαίο τρόπο. Ακόμη, για την παραγωγή των υποψήφιων λύσεων για τις μεταβλητές x_1, x_2 χρησιμοποιούμε την ακόλουθη σχέση:

$$x_{new} = x + rnorm(1,0,2).$$

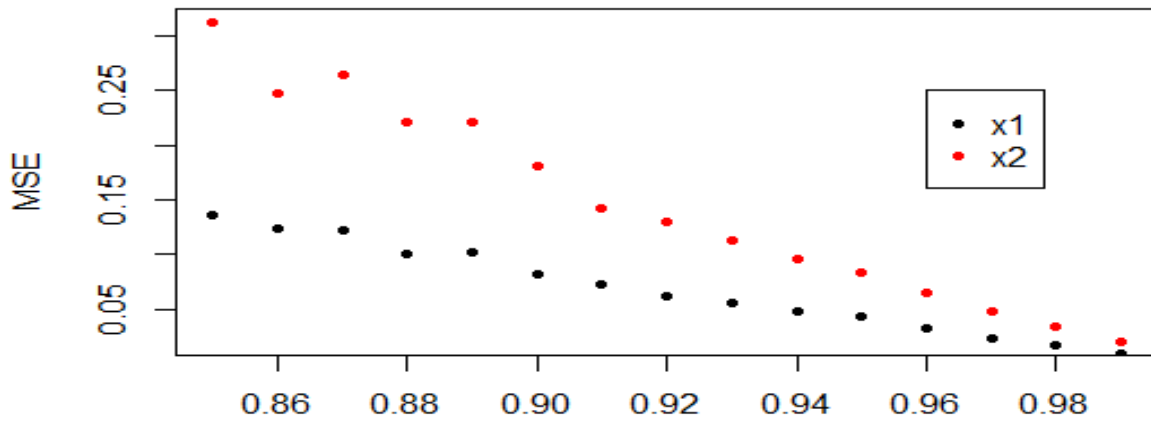
Δηλαδή, η υποψήφια λύση ακολουθεί την κανονική κατανομή με κέντρο την προηγούμενη παρατήρηση και τυπική απόκλιση 2.

Η συγκεκριμένη διαδικασία πραγματοποιήθηκε για μία ακολουθία διαφορετικών τιμών της παραμέτρου α , ξεκινώντας από $\alpha = 0.85$ και καταλήγοντας στο 0.99. Ο στόχος μας μέσω αυτής της διαδικασίας ήταν να μελετήσουμε εμπειρικά την επίδραση της παραμέτρου α στην ταχύτητα και την ακρίβεια σύγκλισης του αλγορίθμου. Στον Πίνακα 2.1 παραθέτουμε τις εκτιμήσεις για το MSE , τη μέση τιμή και τη διασπορά των παραμέτρων, που προκύπτουν στην περίπτωση σύγκλισης στην κύρια κορυφή της f . Υπενθυμίζεται ότι το μέσο τετραγωνικό σφάλμα (MSE) των εκτιμήσεων είναι το άθροισμα της διακύμανσης των εκτιμήσεων συν το τετράγωνο της μεροληψίας. Όπως

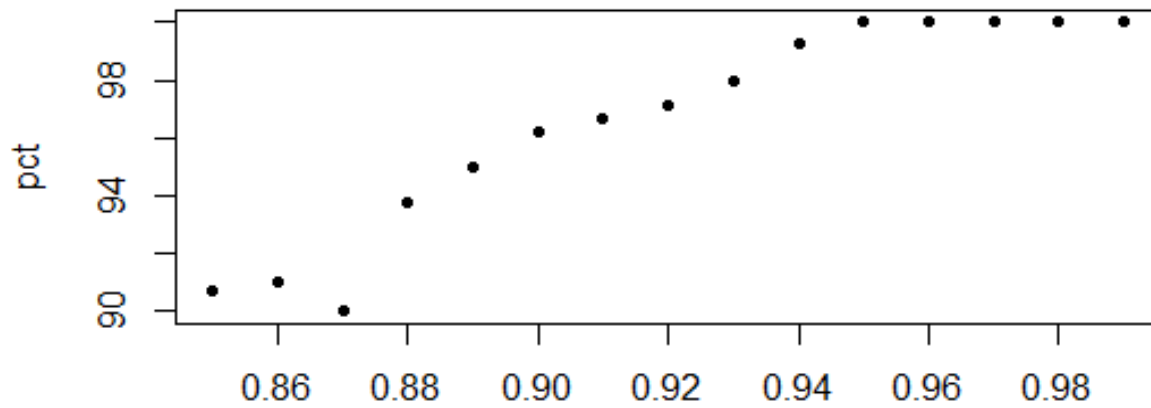
έχει ήδη αναφερθεί η f μετά από αριθμητική βελτιστοποίηση, που πραγματοποιήθηκε χρησιμοποιώντας το Mathematica, φαίνεται να παρουσιάζει στα σημεία $(-0.955467, -0.956622)$ τοπικό μέγιστο και στο σημείο $(0.999599, 2.498955)$ ολικό μέγιστο αντίστοιχα. Έτσι λοιπόν, για να διακρίνουμε σε ποια από τις δύο κορυφές συγκλίνει μία εκτίμηση (\hat{x}_1, \hat{x}_2) υπολογίζουμε την απόσταση της από τις δύο κορυφές και θεωρούμε ότι συγκλίνει στην κορυφή που είναι πιο κοντά.

a	Μέση τιμή \hat{x}_1	Τυπική απόκλιση \hat{x}_1	Μέση τιμή \hat{x}_2	Τυπική απόκλιση \hat{x}_2	Μέση τιμή $\hat{f}(x_1, x_2)$	Τυπική απόκλιση $\hat{f}(x_1, x_2)$	Μέσος αριθμός επαναλήψεων	Ποσοστό ορθών λύσεων
0.85	0.9645	0.3670	2.4245	0.5542	0.0572	0.0079	49.006	90.7%
0.86	0.9676	0.3493	2.4331	0.4927	0.0585	0.0072	53.181	91%
0.87	0.9755	0.3487	2.4217	0.5079	0.0583	0.0071	57.069	90%
0.88	0.9885	0.3173	2.4594	0.4678	0.0594	0.0061	62.577	93.8%
0.89	0.9895	0.3183	2.4583	0.4689	0.0594	0.0064	68.219	95%
0.90	0.9810	0.2860	2.4867	0.4253	0.0605	0.0054	75.159	96.2%
0.91	0.9873	0.2685	2.4756	0.3761	0.0571	0.0046	84.356	96.7%
0.92	0.9799	0.2488	2.4711	0.3602	0.0580	0.0043	95.507	97.1%
0.93	0.9829	0.2337	2.4852	0.3354	0.0594	0.0037	108.563	98%
0.94	0.9975	0.2201	2.4848	0.3085	0.0603	0.0033	127.698	99.3%
0.95	1.0153	0.2062	2.5100	0.2899	0.0610	0.0029	153.194	99.7%
0.96	0.9949	0.1779	2.4954	0.2539	0.0621	0.0023	191.938	100%
0.97	0.9899	0.1529	2.4972	0.2198	0.0629	0.0015	257.107	100%
0.98	0.9919	0.1277	2.4937	0.1828	0.0646	0.0011	388.052	100%
0.99	0.9987	0.0999	2.4984	0.1415	0.0651	0.0006	779.25	100%

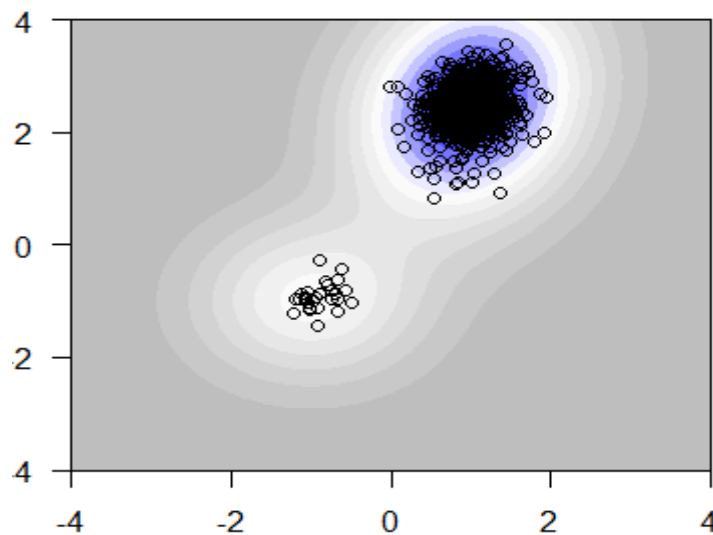
Πίνακας 2.1 Τα αποτελέσματα από την εκτέλεση του αλγορίθμου της προσομοιωμένης απόπτωσης 10^3 φορές, για διάφορες τιμές του a όπου οι υποψήφιες λύσεις παράγονται σύμφωνα με τη σχέση $x_{new} = x + rnorm(0,2)$



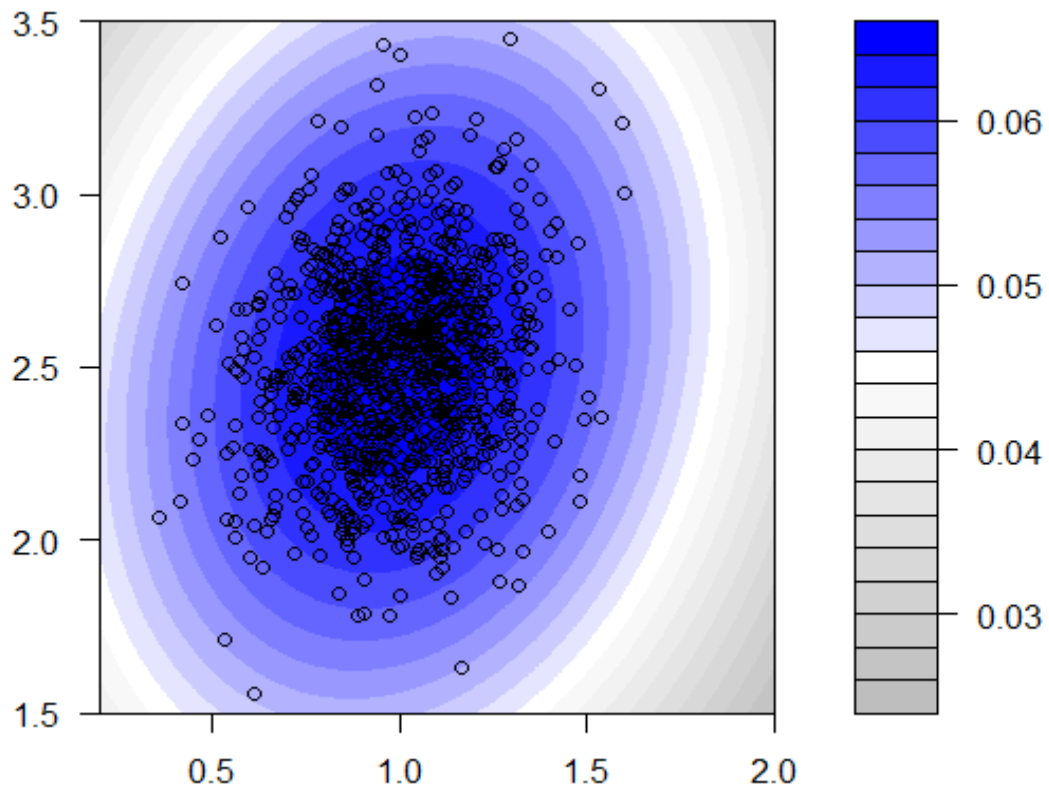
Σχήμα 2.3. Στο παραπάνω διάγραμμα απεικονίζεται το MSE για τα x_1, x_2 συναρτήσει του α .



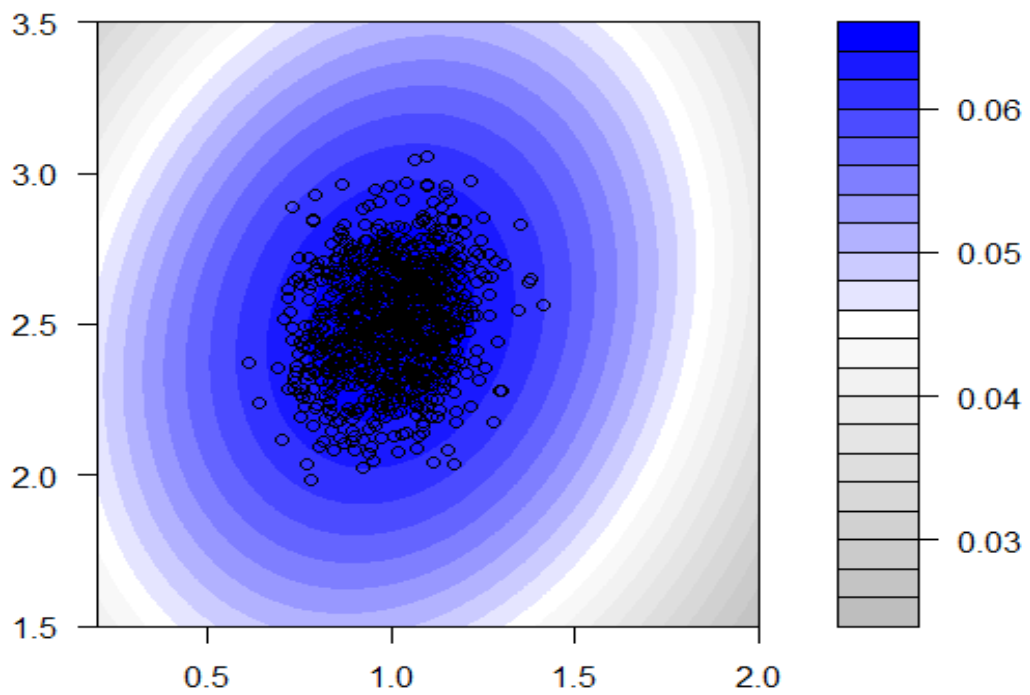
Σχήμα 2.4. Στο παραπάνω διάγραμμα απεικονίζεται το ποσοστό των ορθών λύσεων συναρτήσει του α .



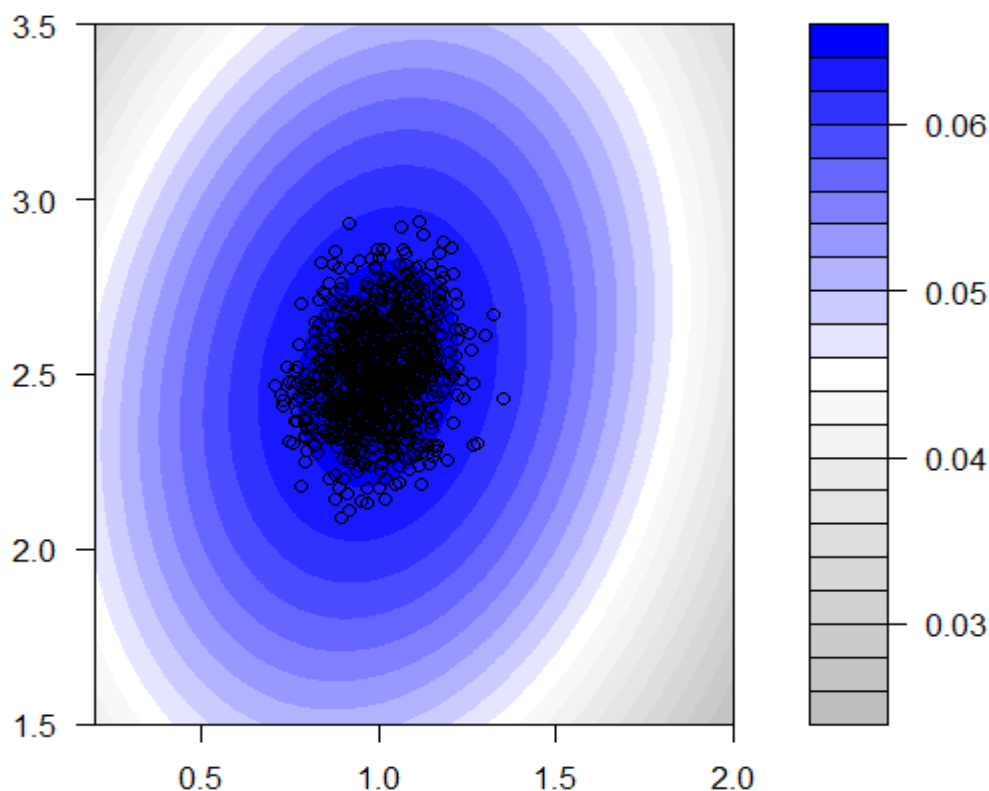
Σχήμα 2.5. Στο παραπάνω διάγραμμα ισοϋψών καμπυλών βλέπουμε τις εκτιμήσεις των λύσεων που προκύπτουν από 1000 εκτιμήσεις του αλγορίθμου για $\alpha = 0.92$.



Σχήμα 2.4 Στο παραπάνω διάγραμμα ισοψών καμπυλών βλέπουμε την διασπορά των εκτιμήσεων γύρω από την κύρια κορυφή που προκύπτει από 1000 επαναλήψεις του αλγορίθμου για $\alpha = 0.95$.



Σχήμα 2.5 Στο παραπάνω διάγραμμα ισοϋψών καμπυλών βλέπουμε την διασπορά των εκτιμήσεων γύρω από την κύρια κορυφή που προκύπτει από 1000 επαναλήψεις του αλγορίθμου για $\alpha = 0.98$.



Σχήμα 2.6 Στο παραπάνω διάγραμμα ισοϋψών καμπυλών βλέπουμε την διασπορά των εκτιμήσεων γύρω από την κύρια κορυφή που προκύπτει από 1000 επαναλήψεις του αλγορίθμου για $\alpha = 0.99$.

Από τα αποτελέσματα του Πίνακα 2.1 βλέπουμε ότι για ότι όλες τις τιμές του α , οι μέσες τιμές των εκτιμήσεων του αλγορίθμου στην περίπτωση σύγκλισης στην κύρια κορυφή είναι αρκετά κοντά στις πραγματικές τιμές (0.999599 και 2.498955 αντίστοιχα). Όσον αφορά, τις τυπικές αποκλίσεις των εκτιμήσεων αν και οι αρχικές τιμές τους μπορούν να θεωρηθούν αρκετά υψηλές π.χ. για $\alpha = 0.85$ η μέση εκτίμηση και η τυπική απόκλιση του \hat{x}_1 είναι 0.9645 και 0.3670 αντίστοιχα. Ωστόσο, όπως φαίνεται από τον Πίνακα 2.1 οι τυπικές αποκλίσεις των εκτιμήσεων μειώνονται καθώς αυξάνεται το α . Στο ίδιο συμπέρασμα καταλήγουμε παρατηρώντας και το Σχήμα 2.2, καθώς όσο αυξάνεται το α μειώνεται το MSE και, συνεπώς, οι εκτιμήσεις που προκύπτουν θεωρούνται πιο αξιόπιστες. Αν και στην περίπτωση που το α ισούται με 0.87 και 0.89 παρατηρείτε μία παροδική αύξηση του MSE . Ωστόσο αυτή η αύξηση

μπορεί να αποδοθεί στη στοχαστικότητα της μεθόδου αλλά και στις πολύ μικρές διαφορές του πλήθους των επαναλήψεων του αλγορίθμου που έχουμε για τις αρχικές τιμές του α . Ακόμη, από το ποσοστό των ορθών συγκλίσεων βλέπουμε ότι για όλες τις τιμές του α το ποσοστό τους είναι άνω του 90% ενώ για α από 0.96 έως 0.99 το ποσοστό είναι 100%. Επομένως, επιβεβαιώνεται και πειραματικά η δυνατότητα του αλγορίθμου να μην εγκλωβίζεται σε περιοχές τοπικών ελαχίστων. Ένα άλλο σημαντικό σημείο που πρέπει να λάβουμε υπόψιν είναι η σχέση ανάμεσα στην τιμή της παραμέτρου α και του υπολογιστικού κόστους του αλγορίθμου. Στο συγκεκριμένο παράδειγμα, το υπολογιστικό κόστος προσδιορίζεται μέσω του μέσου αριθμού επαναλήψεων που απαιτούνται για μία εκτέλεση του αλγορίθμου. Σύμφωνα, με τα αποτελέσματα του Πίνακα 2.1 είναι προφανές ότι ναι μεν οι εκτιμήσεις μας βελτιώνονται καθώς αυξάνεται το α αλλά ταυτόχρονα αυξάνεται και το υπολογιστικό κόστος του αλγορίθμου. Παραδείγματος χάριν, για $\alpha = 0.99$ οι εκτιμήσεις προκύπτουν είναι οι βέλτιστες ως προς τις τυπικές αποκλίσεις ωστόσο απαιτούνται κατά μέσο όρο 779.25 επαναλήψεις για κάθε εκτέλεση του αλγορίθμου. Γι' αυτό το λόγο, μία καλή επιλογή για την τιμή της παραμέτρου α θεωρείται αυτή που μας δίνει ικανοποιητικές εκτιμήσεις χωρίς πολύ μεγάλο αριθμό επαναλήψεων όπως π.χ. για $\alpha = 0.96$.

Κεφάλαιο 3

Χρονολογικές Σειρές

Το παρόν κεφάλαιο αποτελείται από δύο μέρη. Αρχικά γίνεται μια συνοπτική παρουσίαση των αυτοπαλίνδρομων υποδειγμάτων κινητού μέσου (*ARMA models*) ενώ, στη συνέχεια, το φαινόμενο της δεσμευμένης ετεροσκεδαστικότητας, καθώς και τα βασικά χαρακτηριστικά των μοντέλων *ARCH* και *GARCH* που έχουν αναπτυχθεί για να το περιγράψουν.

3.1 Γραμμικές διαδικασίες και αυτοπαλίνδρομα υποδείγματα κινητού μέσου

Μία χρονοσειρά $\{X_t\}$ καλείται γραμμική διαδικασία, αν έχει την ακόλουθη μορφή:

$$X_t = \sum_{k=-\infty}^{\infty} \psi_k W_{t-k} \quad (3.1)$$

για κάθε t . Η $\{W_t\}$ είναι μία διαδικασία λευκού θορύβου (white noise process), δηλαδή είναι μία ακολουθία ανεξάρτητων και ισόνομων τ.μ., με μέση τιμή μηδέν και σταθερή διακύμανση.

Επιπλέον, όταν το άθροισμα των συντελεστών ψ_k είναι πεπερασμένο, δηλαδή $\sum_{k=-\infty}^{\infty} |\psi_k| < \infty$, τότε η χρονοσειρά είναι ασθενώς στάσιμη. Με τον όρο «ασθενώς στάσιμη» εννοούμε ότι η μέση τιμή της χρονοσειράς και οι αυτοδιακυμάνσεις είναι ανεξάρτητες από τη χρονική στιγμή t . Δηλαδή ισχύει ότι:

$$\begin{aligned} E(X_t) &= \mu \quad \forall t \\ \text{Var}(X_t) &= \sigma^2 \quad \forall t \\ \text{Cov}(X_t, X_{t+k}) &= \gamma_k \quad \forall t \end{aligned}$$

3.2 Το Αυτοπαλίνδρομο Υπόδειγμα

Μία άλλη σημαντική κλάση υποδειγμάτων είναι αυτή των αυτοπαλίνδρομων υποδειγμάτων. Σύμφωνα με αυτό το υπόδειγμα, η $\{X_t\}$ γράφεται ως γραμμικός

συνδυασμός των πιο πρόσφατων παρατηρήσεων p , διαταραγμένος από μία διαδικασία λευκού θορύβου. Η μορφή του υποδείγματος είναι η εξής:

$$X_t = W_t + \varphi_1 X_{t-1} + \dots + \varphi_p X_{t-p} \quad (3.2)$$

και το συμβολίζουμε ως $AR(p)$. Οι συντελεστές φ_i είναι πραγματικοί αριθμοί.

Η εξίσωση 3.2 μπορεί να γραφεί σε μία πιο συμπαγή μορφή με χρήση των τελεστών υστέρησης. Πιο συγκεκριμένα, εφαρμόζουμε τον εξής μετασχηματισμό:

$$X_{t-m} = L^m X_t$$

Έτσι, λοιπόν, η εξίσωση 3.2 χρησιμοποιώντας τον τελεστή υστέρησης γράφεται:

$$(1 - \varphi_1 L - \varphi_2 L^2 - \dots - \varphi_p L^p) X_t = W_t$$

$$\varphi(L) X_t = W_t \quad (3.3),$$

όπου $\varphi(L) = 1 - \sum_{j=1}^p \varphi_j L^j$.

Η ικανή και αναγκαία συνθήκη για να είναι η διαδικασία $AR(p)$ στάσιμη είναι οι ρίζες του

$$L^p - \varphi_1 L^{p-1} - \dots - \varphi_{p-1} L - \varphi_p = 0$$

να βρίσκονται εντός του μοναδιαίου κύκλου, στο μιγαδικό επίπεδο.

Για την απόδειξη της παραπάνω συνθήκης παραπέμπουμε τον αναγνώστη στο βιβλίο των George E.P. Box, Gwilym M. Jenkins, Gregory C. Reinsel και Greta M. Ljung, *Time Series Analysis: Forecasting and Control (2015)*.

3.3 Το υπόδειγμα κινητού μέσου

Στην περίπτωση που από την εξίσωση (3.1) το πλήθος από τους συντελεστές ψ_k (μόνο πεπερασμένοι) είναι μη αρνητικοί, τότε έχουμε μία διαδικασία κινητού μέσου. Πιο συγκεκριμένα, η X_t γράφεται ως εξής:

$$X_t = W_t + \theta_1 W_{t-1} + \dots + \theta_q W_{t-q} \quad (3.4)$$

και την συμβολίζουμε $MA(q)$. Οι συντελεστές θ_i είναι πραγματικοί αριθμοί.

Η εξίσωση 3.4 μπορεί να γραφεί σε μία πιο συμπαγή μορφή με χρήση των τελεστών υστέρησης. Πιο συγκεκριμένα, εφαρμόζουμε τον εξής μετασχηματισμό:

$$W_{t-m} = L^m W_t,$$

οπότε η 3.4 γράφεται ως εξής:

$$X_t = (1 + \theta_1 L + \dots + \theta_q L^q) W_t$$

$$X_t = \theta(L) W_t \quad (3.5).$$

Η $MA(q)$ είναι μία στάσιμη χρονοσειρά, αφού γράφεται ως πεπερασμένο άθροισμα όρων λευκού θορύβου. Επιπλέον, όταν οι ρίζες του πολυωνύμου $\theta(L)$ βρίσκονται εκτός του μοναδιαίου κύκλου, τότε η X_t είναι αντιστρέψιμη. Αυτό σημαίνει ότι η X_t μπορεί να γραφεί ως ένα αυτοπαλίνδρομο υπόδειγμα απείρων όρων.

3.4 Το Αυτοπαλίνδρομο Υπόδειγμα Κινητού Μέσου ARMA(p,q)

Το αυτοπαλίνδρομο υπόδειγμα κινητού μέσου τάξης (p,q) ορίζεται ως εξής:

$$X_t = \varphi_1 X_{t-1} + \dots + \varphi_p X_{t-p} + W_t + \theta_1 W_{t-1} + \dots + \theta_q W_{t-q} \quad (3.6)$$

ή, εναλλακτικά, με χρήση των τελεστών υστέρησης

$$(1 - \varphi_1 L - \varphi_2 L^2 - \dots - \varphi_p L^p) X_t = (1 + \theta_1 L + \dots + \theta_q L^q) W_t$$

$$\varphi(L) X_t = \theta(L) W_t.$$

Είναι προφανές, ότι το συγκεκριμένο υπόδειγμα μπορεί να θεωρηθεί ως ένας συνδυασμός ενός αυτοπαλίνδρομου και ενός υποδείγματος κινητού μέσου.

Η στασιμότητα του υποδείγματος εξαρτάται από το αυτοπαλίνδρομο μέρος του, δηλαδή το υπόδειγμα μπορεί να θεωρηθεί στάσιμο, όταν οι ρίζες του πολυωνύμου $\varphi(L)$ βρίσκονται εκτός του μοναδιαίου κύκλου. Από την άλλη, η αντιστρεψιμότητα καθορίζεται από το MA μέρος του υποδείγματος. Έτσι, λοιπόν, το υπόδειγμα είναι αντιστρέψιμο, όταν οι ρίζες του πολυωνύμου $\theta(L)$ βρίσκονται εκτός του μοναδιαίου κύκλου.

3.5 Υποδείγματα Δεσμευμένης Ετεροσκεδαστικότητας

Ένα από τα βασικά χαρακτηριστικά των χρονοσειρών που αφορούν αποδόσεις αξιογράφων είναι η ομαδοποίηση της αστάθειας (volatility clustering). Με τον συγκεκριμένο όρο εννοούμε ότι υπάρχουν χρονικές περιόδους, κατά τις οποίες οι αποδόσεις των αξιογράφων παρουσιάζουν παρόμοια υψηλή διακύμανση, και άλλες (χρονικές περιόδους), κατά τις οποίες παρουσιάζουν παρόμοια χαμηλή διακύμανση. Αυτό σημαίνει ότι η δεσμευμένη διακύμανση των αποδόσεων

$$Var(Y_t | F_{t-1})$$

δεν είναι σταθερή συναρτήσει του χρόνου.

Επίσης, εμπειρικά, από τη μελέτη χρονοσειρών αποδόσεων γνωρίζουμε ότι οι χρονοσειρές αποδόσεων δεν εμφανίζουν ισχυρές αυτοσυσχετίσεις. Εντούτοις, δεν

συμβαίνει το ίδιο και με τα τετράγωνα των αποδόσεων, για τα οποία συνήθως προκύπτει ότι εμφανίζουν ισχυρές αυτοσυσχετίσεις. Επομένως, οι όροι των χρονοσειρών αποδόσεων δεν είναι ανεξάρτητοι.

Έτσι, λοιπόν, στη βιβλιογραφία έχουμε προτάσεις διαφόρων μοντέλων που ενσωματώνουν τα ανωτέρω χαρακτηριστικά των χρονοσειρών αποδόσεων.

Έστω ότι η χρονοσειρά των αποδόσεων ενός αξιογράφου έχει την ακόλουθη μορφή:

$$Y_t = E(Y_t|F_{t-1}) + \varepsilon_t \quad (3.7)$$

όπου το F_{t-1} αντιπροσωπεύει το σύνολο της πληροφορίας μέχρι τη χρονική στιγμή $t - 1$, η $E(Y_t|F_{t-1})$ είναι η δεσμευμένη μέση τιμή και ε_t είναι ο διαταρακτικός όρος.

Δεδομένου ότι οι όροι της χρονοσειράς των αποδόσεων δεν παρουσιάζουν ισχυρές αυτοσυσχετίσεις, συμπεραίνουμε ότι η εξίσωση του μέσου πρέπει να έχει απλή μορφή. Γι' αυτόν τον λόγο, συνήθως υποθέτουμε ότι ακολουθεί μία στάσιμη ARMA(p,q) διαδικασία:

$$E(Y_t|F_{t-1}) = \varphi_0 + \sum_{i=1}^p \varphi_i Y_{t-i} + \sum_{i=1}^q \theta_i \varepsilon_{t-i} \quad (3.8)$$

Επομένως, από την εξίσωση (3.7) προκύπτει

$$\text{Var}(Y_t|F_{t-1}) = \text{Var}(\varepsilon_t|F_{t-1}),$$

δηλαδή η δεσμευμένη διακύμανση των αποδόσεων της χρονοσειράς ισούται με τη δεσμευμένη διακύμανση του διαταρακτικού όρου.

3.5.1 Το υπόδειγμα ARCH

Ο Engle (1982) πρότεινε το υπόδειγμα ARCH (Autoregressive Conditional Heteroskedastic) για να περιγράψει το φαινόμενο της δεσμευμένης ετεροσκεδαστικότητας. Σύμφωνα με αυτό το υπόδειγμα, η δεσμευμένη διακύμανση, τη χρονική στιγμή t , εξαρτάται από τις παρελθοντικές τιμές του διαταρακτικού όρου.

Ορισμός 3.1 Υπόδειγμα ARCH(p)

Ο διαταρακτικός όρος ε_t σε ένα ARCH(p) υπόδειγμα είναι της μορφής

$$\varepsilon_t = \sigma_t z_t, \\ \sigma_t^2 = \omega + \sum_{k=1}^p a_k \varepsilon_{t-k}^2 \quad (3.9)$$

όπου το $\{z_t\}$ είναι μία ακολουθία ανεξάρτητων και ισόνομων τ.μ., με μέση τιμή 0 και διασπορά 1, $\omega > 0$ και $a_k \geq 0$ για κάθε k .

Από τη δομή του μοντέλου *ARCH* παρατηρούμε ότι μεγάλες τιμές για τα τετράγωνα των παρελθοντικών διαταρακτικών $\{\varepsilon_{t-i}^2\}$ όρων έχουν ως αποτέλεσμα μεγάλη τιμή και για τη δεσμευμένη διακύμανση σ_t^2 του διαταρακτικού όρου ε_t . Επομένως, είναι πιθανόν να έχουμε μία μεγάλη τιμή και για τον διαταρακτικό όρο ε_t . Σε ανάλογα συμπεράσματα οδηγούμαστε στην περίπτωση που οι τιμές των τετραγώνων των παρελθοντικών διαταρακτικών όρων είναι μικρές. Ως εκ τούτου, μπορούμε να πούμε ότι το μοντέλο *ARCH*(p) είναι κατάλληλο για να ερμηνεύσει την ομαδοποίηση της αστάθειας.

Ωστόσο, εμπειρικά αποτελέσματα από την εφαρμογή του μοντέλου *ARCH*(p) δείχνουν ότι συνήθως χρειαζόμαστε έναν μεγάλο αριθμό υστερήσεων p , προκειμένου να ερμηνεύσουμε ικανοποιητικά το φαινόμενο της δεσμευμένης ετεροσκεδαστικότητας. Πρακτικά, όμως, ένα μοντέλο με πολλές παραμέτρους είναι δύσχρηστο. Εξαιτίας αυτού, ο Bollerslev (1986) πρότεινε ένα πιο ευέλικτο υπόδειγμα, το *GARCH*, το οποίο χρησιμοποιεί λιγότερες παραμέτρους.

3.5.2 Το υπόδειγμα *GARCH*

Ορισμός 3.2 Υπόδειγμα *GARCH*(p, q)

Ο διαταρακτικός όρος ε_t σε ένα *GARCH*(p, q) υπόδειγμα είναι της μορφής

$$\begin{aligned} \varepsilon_t &= \sigma_t z_t \\ \sigma_t^2 &= \omega + \sum_{k=1}^p a_k \varepsilon_{t-k}^2 + \sum_{j=1}^q \beta_j \sigma_{t-j}^2 \end{aligned} \quad (3.11)$$

όπου $\omega > 0, a_i > 0, \beta_j \geq 0$.

Το $\{z_t\}$ είναι μία ακολουθία ανεξάρτητων και ισόνομων τ.μ., με μέση τιμή 0 και διασπορά 1. Συνήθως, θεωρούμε ότι $z_t \sim N(0,1)$. Ωστόσο, εμπειρικά γνωρίζουμε ότι οι χρονοσειρές αποδόσεων εμφανίζουν ασυμμετρία και πιο βαριές ουρές από την κανονική κατανομή. Γι' αυτόν τον λόγο, πολλοί θεωρούν ότι διαφορετικές κατανομές όπως η Student ή η Γάμμα είναι πιο αντιπροσωπευτικές για τις χρονοσειρές αποδόσεων.

Αν θέσουμε $W_t = \varepsilon_t^2 - \sigma_t^2, k = 0, 1, \dots, q$ και αντικαταστήσουμε στην (3.11), τότε προκύπτει η ακόλουθη εξίσωση

$$\varepsilon_t^2 = \omega + \sum_{k=1}^r (a_k + \beta_k) \varepsilon_{t-k}^2 - \sum_{k=1}^q \beta_k W_{t-k} + W_t \quad (3.12)$$

όπου $r = \max(p, q)$, $a_j \equiv 0$ για $j > p$ και $\beta_j \equiv 0$ για $j > q$.

Η εξίσωση (3.12) έχει πλέον τη μορφή ενός ARMA(p,q) για το ε_t^2 . Σ' αυτήν την περίπτωση, η διαδικασία $\{\varepsilon_t^2\}$ είναι ασθενώς στάσιμη, αν οι ρίζες του πολυωνύμου

$$1 - (a_1 + \beta_1)z - (a_2 + \beta_2)z^2 - \dots - (a_r + \beta_r)z^r = 0 \quad (3.13)$$

βρίσκονται εκτός του μοναδιαίου κύκλου. Όμως, δεδομένου ότι τα a_j και β_j είναι μη αρνητικοί αριθμοί, τότε για να ικανοποιείται η (3.13) αρκεί να ισχύει

$$(a_1 + \beta_1) + (a_2 + \beta_2) + \dots + (a_r + \beta_r) < 1 \quad (3.14)$$

ενώ, όταν ισχύει η (3.14), προκύπτει ότι η μη δεσμευμένη διακύμανση είναι ίση με

$$E(\varepsilon_t^2) = \frac{\omega}{1 - \sum_{k=1}^r (a_k + \beta_k)}.$$

Αν $p = q = 1$, τότε έχουμε το υπόδειγμα $GARCH(1,1)$ που, παρά την απλή μορφή του, αποδεικνύεται, πολλές φορές, επαρκές για να περιγράψει το φαινόμενο της δεσμευμένης διακύμανσης.

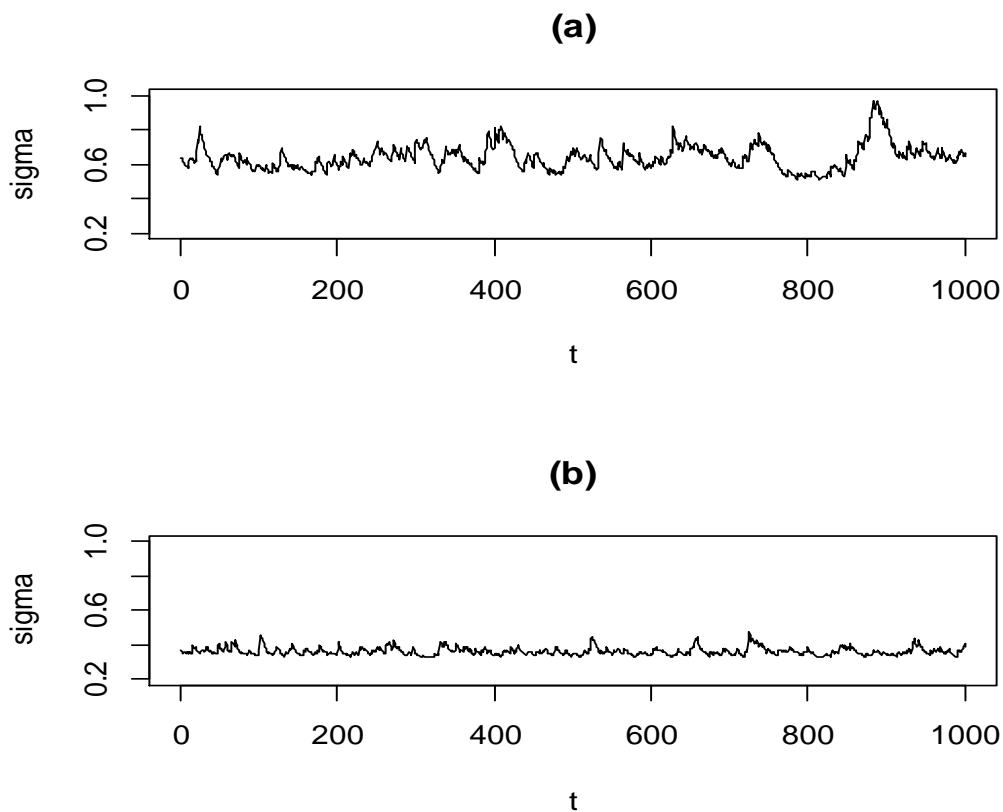
Ορισμός 3.3 Υπόδειγμα $GARCH(1,1)$

Η ακολουθία $\{\varepsilon_t\}$ σε ένα $GARCH(1,1)$ υπόδειγμα είναι της μορφής

$$\begin{aligned} \varepsilon_t &= \sigma_t z_t \\ \sigma_t^2 &= \omega + a_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \end{aligned} \quad (3.11)$$

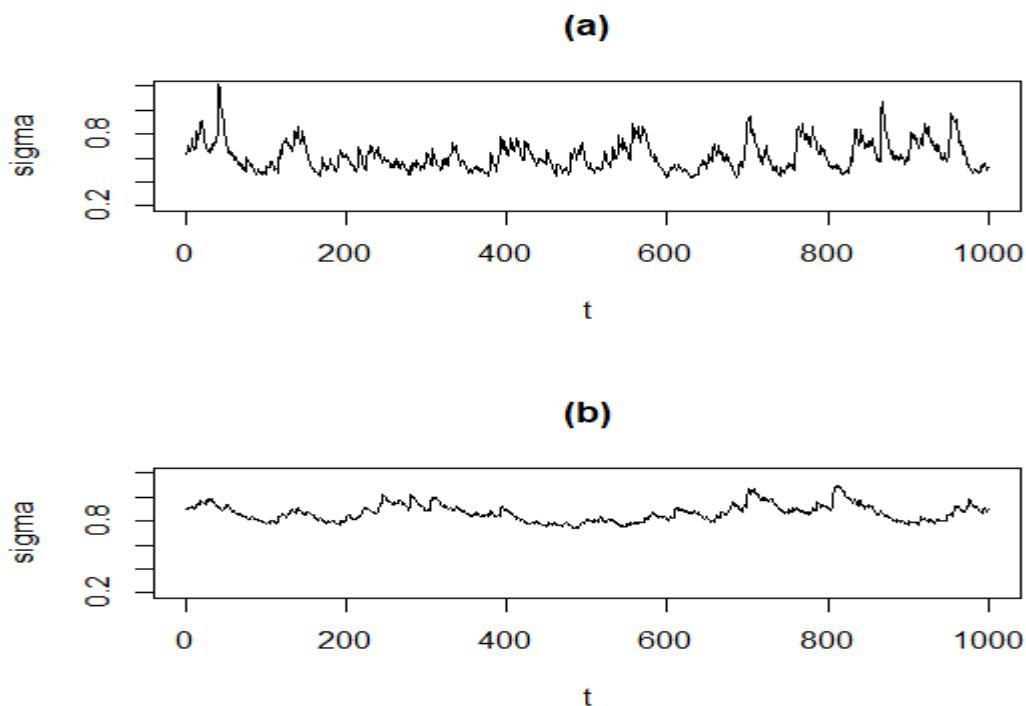
όπου $\omega > 0$, $a_1 \geq 0$, $\beta_1 \geq 0$ και η z_t είναι μία ακολουθία ανεξάρτητων και ισόνομων τ.μ., με μέση τιμή 0 και διασπορά 1.

Στη συνέχεια, ακολουθούν γραφήματα της δεσμευμένης διακύμανσης για διαφορετικές τιμές των παραμέτρων ω , a_1 και β_1 . Με αυτόν τον τρόπο, προσπαθούμε να συμπεράνουμε γραφικά πώς επηρεάζουν οι μεταβολές των τιμών των παραμέτρων τη δεσμευμένη διακύμανση.



Σχήμα 3.1 Γράφημα της δεσμευμένης διακύμανσης δύο υποδειγμάτων $GARCH(1,1)$ με παραμέτρους (a) ($\omega=0.02, \alpha_1=0.05, \beta_1=0.9$), (b) ($\omega=0.02, \alpha_1=0.05, \beta_1=0.8$).

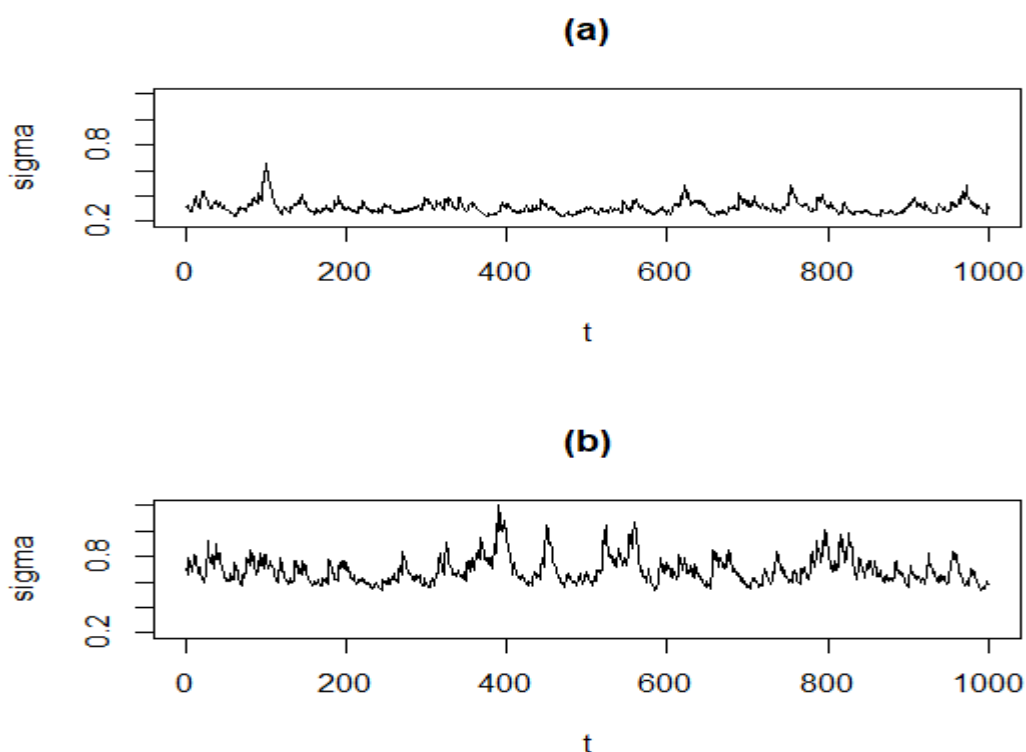
Στο Σχήμα 3.1 απεικονίζονται οι προσομοιωμένες τιμές της δεσμευμένης διακύμανσης δύο υποδειγμάτων $GARCH(1,1)$ με ίδιες τιμές για τις παραμέτρους ω, α_1 , αλλά με διαφορετικό β_1 . Παρατηρούμε ότι με $\beta_1 = 0.8$ (βλέπε Σχήμα 3.1b) έχουμε πολύ μικρές μεταβολές της δεσμευμένης διακύμανσης. Από την άλλη πλευρά, για μεγαλύτερες τιμές της παραμέτρου, όπως $\beta_1 = 0.9$ (βλέπε Σχήμα 3.1a), έχουμε μεγαλύτερες τιμές για τη δεσμευμένη διακύμανση. Επιπλέον, η χρονοσειρά για $\beta_1 = 0.9$ παρουσιάζει μεγαλύτερες διαταραχές και, ως εκ τούτου, έχουμε πιο αργή σύγκλιση στο μακροχρόνιο μέσο.



Σχήμα 3.2 Γράφημα της δεσμευμένης διακύμανσης δύο υποδειγμάτων GARCH(1,1) με παραμέτρους (a) ($\omega=0.02$, $\alpha_1=0.1$, $\beta_1=0.85$), (b) ($\omega=0.02$, $\alpha_1=0.025$, $\beta_1=0.95$).

Στο Σχήμα 3.2 απεικονίζονται οι προσομοιωμένες τιμές της δεσμευμένης διακύμανσης δύο υποδειγμάτων GARCH(1,1) με ίδια τιμή για την παράμετρο ω , αλλά με διαφορετικά α_1 , β_1 . Παρατηρούμε ότι για $\alpha_1 = 0.1$ και $\beta_1 = 0.85$ (βλέπε σχήμα 3.2a) το γράφημα της δεσμευμένης διακύμανσης παρουσιάζει σχετικά μεγάλες διαταραχές, σε σχέση με το γράφημα για το GARCH(1,1) με παραμέτρους $\alpha_1 = 0.025$ και $\beta_1 = 0.95$. Δηλαδή, μπορούμε να συμπεράνουμε ότι για μεγαλύτερες τιμές για το α_1 και μικρότερες τιμές για το β_1 η δεσμευμένη διακύμανση παρουσιάζει μεγάλη μεταβλητότητα, εν αντιθέσει με τη δεύτερη περίπτωση, στην οποία έχουμε μικρότερη τιμή για το α_1 και μεγαλύτερη τιμή για το β_1 , όπου η δεσμευμένη διακύμανση βρίσκεται σε ισορροπία.

Στο επόμενο γράφημα, βλέπουμε πώς μεταβάλλεται η δεσμευμένη διακύμανση, όταν μεταβάλλουμε το ω , ενώ οι παράμετροι α_1 , β_1 παραμένουν σταθεροί.



Σχήμα 3.3 Γράφημα της δεσμευμένης διακύμανσης δύο υποδειγμάτων $GARCH(1,1)$ με παραμέτρους (a) ($\omega=0.01, \alpha_1=0.1, \beta_1=0.8$), (b) ($\omega=0.05, \alpha_1=0.1, \beta_1=0.8$).

Από το Σχήμα 3.3 βλέπουμε ότι μικρότερες τιμές του ω οδηγούν σε χαμηλότερες τιμές για τη δεσμευμένη διακύμανση. Επίσης, για $\omega = 0.05$ παρατηρούμε σχετικά μεγαλύτερη μεταβλητότητα στο γράφημα της δεσμευμένης διακύμανσης.

3.5.3 Εκτίμηση παραμέτρων του υποδείγματος $GARCH(1, 1)$

Η εκτίμηση των παραμέτρων του υποδείγματος $GARCH(1,1)$ θα γίνει με τη μέθοδο της μέγιστης πιθανοφάνειας. Αν υποθέσουμε ότι η z_t είναι μία ακολουθία ανεξάρτητων και ισόνομων τ.μ. από την $N(0,1)$, τότε η $\varepsilon_t \sim N(0, \sigma_t^2)$. Επομένως, η συνάρτηση πιθανοφάνειας του υποδείγματος είναι

$$\begin{aligned} L(\boldsymbol{\theta}; \varepsilon_1, \dots, \varepsilon_T) &= f(\varepsilon_T | F_{T-1}; \boldsymbol{\theta}) f(\varepsilon_{T-1} | F_{T-2}; \boldsymbol{\theta}) \dots f(\varepsilon_2 | F_1; \boldsymbol{\theta}) f(\varepsilon_1; \boldsymbol{\theta}) \\ &= f(\varepsilon_1; \boldsymbol{\theta}) \prod_{t=2}^T f(\varepsilon_t | F_{t-1}; \boldsymbol{\theta}) = f(\varepsilon_1; \boldsymbol{\theta}) \prod_{t=2}^T \frac{1}{\sqrt{2\pi\sigma_t^2}} \exp\left(-\frac{\varepsilon_t^2}{2\sigma_t^2}\right). \end{aligned}$$

Δεδομένου ότι δεν γνωρίζουμε την ακριβή έκφραση της σ.π.π. $f(\varepsilon_1; \boldsymbol{\theta})$ και ότι διαθέτουμε ένα επαρκώς μεγάλο δείγμα, συνήθως επιλέγουμε να την παραλείψουμε

από τη συνάρτηση πιθανοφάνειας. Άρα, προκύπτει η ακόλουθη δεσμευμένη πιθανοφάνεια:

$$L(\boldsymbol{\theta}; \varepsilon_2, \dots, \varepsilon_T | \varepsilon_1) = \prod_{t=2}^T \frac{1}{\sqrt{2\pi\sigma_t^2}} \exp\left(-\frac{\varepsilon_t^2}{2\sigma_t^2}\right)$$

Ως γνωστόν, η μεγιστοποίηση της πιθανοφάνειας είναι ισοδύναμη με τη μεγιστοποίηση του λογαρίθμου της

$$\log(L(\boldsymbol{\theta}; \varepsilon_2, \dots, \varepsilon_T | \varepsilon_1)) = \sum_{t=2}^T \left[-\frac{1}{2} \ln(2\pi) - \frac{1}{2} \ln(\sigma_t^2) - \frac{\varepsilon_t^2}{2\sigma_t^2} \right]$$

και παραλείποντας τις προσθετικές σταθερές καταλήγουμε ότι η λογαριθμική συνάρτηση πιθανοφάνειας που πρέπει να μεγιστοποιήσουμε είναι:

$$\log(L(\boldsymbol{\theta}; \varepsilon_2, \dots, \varepsilon_T | \varepsilon_1)) = \sum_{t=2}^T \left[-\frac{1}{2} \ln(\sigma_t^2) - \frac{\varepsilon_t^2}{2\sigma_t^2} \right]$$

όπου $\boldsymbol{\theta} = (\omega, a_1, \beta_1)$.

Για περισσότερες πληροφορίες σχετικά με τα υποδείγματα δεσμευμένης ετεροσκεδαστικότητας ο αναγνώστης παραπέμπεται στα συγγράμματα των A. Arratia (2014), Ruey S. Tsay (2012).

Κεφάλαιο 4

Εφαρμογή της προσομοιωμένης ανόπτωσης με την βοήθεια της R

4.1 Εισαγωγή

Σε αυτό το κεφάλαιο πρόκειται να εφαρμόσουμε τον αλγόριθμο της προσομοιωμένης ανόπτωσης για τον προσδιορισμό των παραμέτρων ενός υποδείγματος $GARCH(1,1)$. Όπως είδαμε στην παράγραφο 3.5.3, η εκτίμηση των παραμέτρων του υποδείγματος γίνεται με τη μέθοδο της μέγιστης πιθανοφάνειας, απ' όπου προκύπτει και η προς μεγιστοποίηση συνάρτηση.

Πιο συγκεκριμένα, το πρόβλημα μεγιστοποίησης μπορεί να διατυπωθεί ως εξής: Το ζητούμενο είναι η εύρεση των $\theta^* = (\omega, a_1, \beta_1)$ που μεγιστοποιούν τη συνάρτηση του λογαρίθμου της πιθανοφάνειας

$$F(\theta) = \sum_{t=2}^T \left[-\frac{1}{2} \ln(\sigma_t^2) - \frac{\varepsilon_t^2}{2\sigma_t^2} \right].$$

Επιπροσθέτως, πρέπει να ικανοποιούνται τα εξής:

$$\varepsilon_t = \sigma_t z_t, \quad z_t \sim N(0,1)$$

$$\sigma_t^2 = \omega + a_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2$$

όπου: $\omega > 0$, $a_1 \geq 0$, $\beta_1 \geq 0$ και $a_1 + \beta_1 < 1$.

Αρχικά, μέσω του ακόλουθου κώδικα R, προσομοιώνουμε την πραγματοποίηση ενός υποδείγματος $GARCH(1,1)$ με παραμέτρους $(\omega, a_1, \beta_1) = (0.3, 0.2, 0.5)$.

```
omega=0.3; alpha=0.2;beta=0.5;epsilon<-NULL; n=10^3;
sigma<-rep(sqrt(omega/(1-alpha-beta)),n);
for(t in 1:n)
{
  epsilon[t]=sigma[t]*rnorm(1,0,1)
  sigma[t+1]=sqrt(omega+(beta*sigma[t]^2)+(alpha*epsilon[t]^2))
}
```

Στη συνέχεια, δίνεται η συνάρτηση του λογαρίθμου της πιθανοφάνειας μέσω της R. Ο υπολογισμός της συνάρτησης αυτής πραγματοποιείται, αφού δοθούν οι τιμές των

παραμέτρων $(\omega, \alpha_1, \beta_1)$, καθώς και οι τιμές των $\varepsilon_t, t = 1, 2, \dots, n$. Οι τελευταίες (τιμές) συνήθως προκύπτουν από μία σχέση, της μορφής $\varepsilon_t = \ln(S_t/S_{t-1})$ για κάποια παρατηρούμενη χρονοσειρά $S_t, t = 0, 1, 2, \dots, n$, που εκφράζει τιμές ενός αξιογράφου. Από τις τιμές των ε_t υπολογίζονται αναδρομικά και οι τιμές $\sigma_t, t = 2, 3, \dots, n$ (χρησιμοποιώντας τη σχέση $\sigma_t^2 = \omega + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2$), οι οποίες είναι απαραίτητες για τον υπολογισμό της πιθανοφάνειας. Ως αρχική τιμή του σ_t^2 λαμβάνεται η στάσιμη τιμή του.

```
loglikelihood=function(epsilon,omega,alpha,beta)
{
  sumlog=0;sigma=sqrt(omega/(1-alpha-beta))
  for(t in 2:n)
  {
    sigma=sqrt(omega+alpha*epsilon[t-1]^2+beta*sigma^2);
    sumlog=sumlog+(log(sigma^2)/(-2))-epsilon[t]^2/(2*sigma^2)
  }
  return(sumlog)
}
```

4.2 Εκτίμηση των παραμέτρων μέσω της προσομοιωμένης ανόπτησης

Χρησιμοποιώντας την ακολουθία ε_t και τη συνάρτηση της πιθανοφάνειας μπορούμε να εφαρμόσουμε τον αλγόριθμο της προσομοιωμένης ανόπτησης μέσω της συνάρτησης SA, η δομή της οποίας περιγράφεται ως εξής:

1. Αρχικοποίηση
 - a. Δίνουμε τιμές στα ορίσματα της συνάρτησης, τα οποία είναι: α , αρχική θερμοκρασία, τελική θερμοκρασία, τυπική απόκλιση βημάτων, ε_t , ω , α_1 , και β_1 .
 - b. Ορίζουμε τέσσερις κενούς πίνακες OMEGA, B, A και Lik, στους οποίους, μετά από κάθε εκτέλεση, θα αποθηκεύονται οι τιμές των παραμέτρων και η τιμή της συνάρτησης πιθανοφάνειας.
 - c. Θέτουμε τη μεταβλητή *step* (μέσω της οποίας καταγράφουμε τον αριθμό των επαναλήψεων του αλγορίθμου) ίση με μηδέν.
2. Εφόσον η συνθήκη $T > T_end$ είναι αληθής επαναλαμβάνουμε τα εξής:

- a. Υπολογίζουμε την συνάρτηση της πιθανοφάνειας L για τις τρέχουσες τιμές των παραμέτρων.
- b. Προσομοιώνουμε νέες υποψήφιες τιμές για τα ω , α_1 , και β_1 σύμφωνα με τη σχέση $x_{new} = x + rnorm(1,0,s)$.
- c. Αν $\omega_{new} > 0$, $\alpha_{1_{new}} > 0$, $\beta_{1_{new}} > 0$ και $\alpha_{1_{new}} + \beta_{1_{new}} < 1$, τότε:
 Θέτουμε $step = step + 1$
 Υπολογίζουμε την συνάρτηση της πιθανοφάνειας L_{new} για τις παραμέτρους ω_{new} , $\alpha_{1_{new}}$, $\beta_{1_{new}}$.
 Υπολογίζουμε την πιθανότητα αποδοχής $\min\left(1, \exp\left(\frac{L_{new}-L}{T}\right)\right)$.
 Προσομοιώνουμε $u \sim U(0,1)$.
- d. Αν $u \leq ar$ τότε θέτουμε: $\omega = \omega_{new}$, $\alpha_1 = \alpha_{1_{new}}$ και $\beta_1 = \beta_{1_{new}}$.
- e. Ενημερώνουμε τους πίνακες OMEGA, B, A και Lik με τις τρέχουσες τιμές των παραμέτρων.
- f. Θέτουμε $T = T * a$.

3. Κατασκευάζουμε τα διαγράμματα των παραμέτρων και της συνάρτησης της πιθανοφάνειας συναρτήσει του αριθμού των επαναλήψεων.
 Επιστρέφουμε ένα διάνυσμα που περιέχει τις τρέχουσες τιμές των παραμέτρων ω , α_1 , και β_1 και τα διαγράμματα.

Ο κώδικας της R μέσω του οποίου υλοποιείται η συνάρτηση SA είναι ο ακόλουθος:

```
SA<-function(a,T,T_end,s,epsilon,omega,alpha,beta)
{
step=0;OMEGA=NULL;B=NULL;A=NULL; Lik=NULL;
while(T>T_end)
{
L=loglikelihood(epsilon,omega,alpha,beta)
omega_new=omega+rnorm(1,0,s)
beta_new=beta+rnorm(1,0,s)
alpha_new=alpha+rnorm(1,0,s)
if((omega_new>0) && (alpha_new>=0) && (beta_new>=0)
&& (alpha_new+beta_new<1))
{L_new=loglikelihood(epsilon,omega_new,alpha_new,beta_new)
step=step+1; DL=L_new-L
ar=min(1,exp(DL/T))
u=runif(1)
if(u<=ar)
{omega=omega_new; alpha=alpha_new; beta=beta_new;}
}
}
OMEGA[step]=omega; A[step]=alpha; B[step]=beta;
```

```

Lik[step]=loglikelihood(epsilon,omega,alpha,beta)
T=T*a
}
maxLik=loglikelihood(epsilon,omega,alpha,beta)
iterations=step; par(mfrow=c(2,2))
omega.plot<-plot(OMEGA,type="l",xlab="iterations"
,ylab="omega")
alpha.plot<-plot(A,type="l",xlab="iterations" ,ylab="alpha")
beta.plot<-plot(B,type="l",xlab="iterations" ,ylab="beta")
F.plot<-plot(Lik,type="l",xlab="iterations" ,ylab="F")

return(list(omega,alpha,beta,maxLik,iterations,omega.plot,
alpha.plot, beta.plot, F.plot))
}

```

4.3 Παραδείγματα

4.3.1 Εκτίμηση των παραμέτρων ενός *GARCH(1, 1)* υποδείγματος για διάφορες τιμές της παραμέτρου α

Σε αυτό το σημείο, έχουμε τη δυνατότητα να εξερευνήσουμε τη συμπεριφορά του αλγορίθμου. Δεδομένου ότι η μέθοδος είναι στοχαστική, δεν είναι δυνατόν να αξιολογήσουμε την αποτελεσματικότητά της από μία μόνο πραγματοποίηση. Γι' αυτόν τον λόγο, θα πραγματοποιήσουμε 1000 επαναλήψεις του αλγορίθμου για μία σειρά διαφορετικών τιμών των παραμέτρων.

Κατ' αρχάς, θα εκτελέσουμε 1000 επαναλήψεις του αλγορίθμου για μία σειρά διαφορετικών τιμών της παραμέτρου α (0.95, 0.98, 0.99, 0.999), ενώ οι υπόλοιπες παράμετροι θα παραμένουν ίδιες σε όλες τις εκτελέσεις του αλγορίθμου. Πιο συγκεκριμένα, έχουμε επιλέξει ως αρχικές τιμές για το διάνυσμα των παραμέτρων $(\omega_0, \alpha_{1,0}, \beta_{1,0}) = (0.1, 0.1, 0.1)$ και για το s την τιμή 0.005. Ως αρχική θερμοκρασία επιλέγουμε την $T = 100$, κι αυτό διότι, ύστερα από δοκιμές, παρατηρήσαμε ότι με τη συγκεκριμένη θερμοκρασία επιτυγχάνουμε ποσοστό αποδοχής άνω του 80% στα 50 πρώτα βήματα του αλγορίθμου. Συνεπώς, η συγκεκριμένη θερμοκρασία επιτρέπει στον αλγόριθμο να κινηθεί αρχικά σε όλο τον χώρο λύσεων, ενώ ως τελική θερμοκρασία επιλέγουμε την $T_{end} = 0.001$. Υπενθυμίζεται ότι οι πραγματικές τιμές των παραμέτρων είναι $\omega = 0.3, \alpha_1 = 0.2, \beta_1 = 0.5$.

Στους παρακάτω πίνακες παρουσιάζονται οι εκτιμήσεις της μέσης τιμής, της τυπικής απόκλισης και του μέσου τετραγωνικού σφάλματος που προκύπτουν για τα $\omega, \alpha_1, \beta_1$

από την παραπάνω διαδικασία. Υπενθυμίζεται ότι το μέσο τετραγωνικό σφάλμα (MSE) είναι το άθροισμα της διακύμανσης των εκτιμήσεων συν το τετράγωνο της μεροληψίας.

α	μέση τιμή $\hat{\omega}$	τυπική απόκλιση $\hat{\omega}$	μέση τιμή $\hat{\alpha}_1$	τυπική απόκλιση $\hat{\alpha}_1$
0.95	0.4236	0.03927582	0.2570	0.0407391
0.98	0.3527	0.07845325	0.2082	0.04008411
0.99	0.3218	0.09188184	0.20031	0.04253976
0.999	0.3151	0.09752066	0.19837	0.04460864

Πίνακας 4.1 Οι εκτιμήσεις των παραμέτρων ω, α_1 , που προκύπτουν από την εκτέλεση 1000 επαναλήψεων του αλγορίθμου της προσομοιωμένης ανόπτησης για διάφορες τιμές του α .

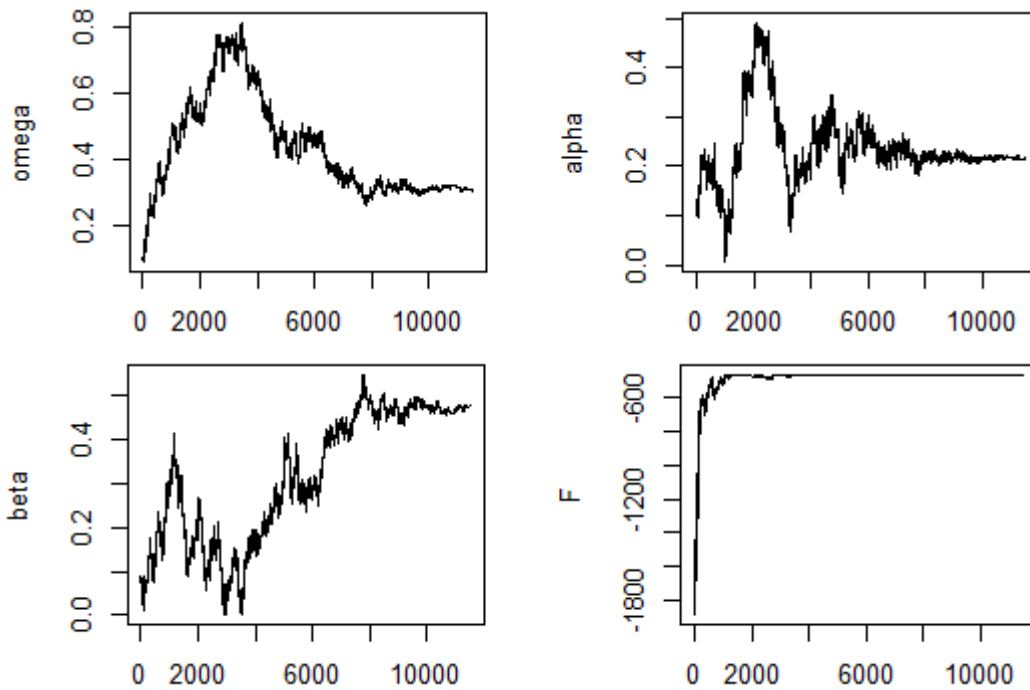
α	μέση τιμή $\hat{\beta}_1$	τυπική απόκλιση $\hat{\beta}_1$	Μέσος αριθμός επαναλήψεων
0.95	0.28199	0.0529786	225
0.98	0.43640	0.08641039	570
0.99	0.475305	0.1098878	1146
0.999	0.4841582	0.1195758	11496

Πίνακας 4.2 Οι εκτιμήσεις της παραμέτρου β_1 και ο μέσος αριθμός επαναλήψεων, που προκύπτουν από την εκτέλεση 1000 επαναλήψεων του αλγορίθμου της προσομοιωμένης ανόπτησης για διάφορες τιμές του α .

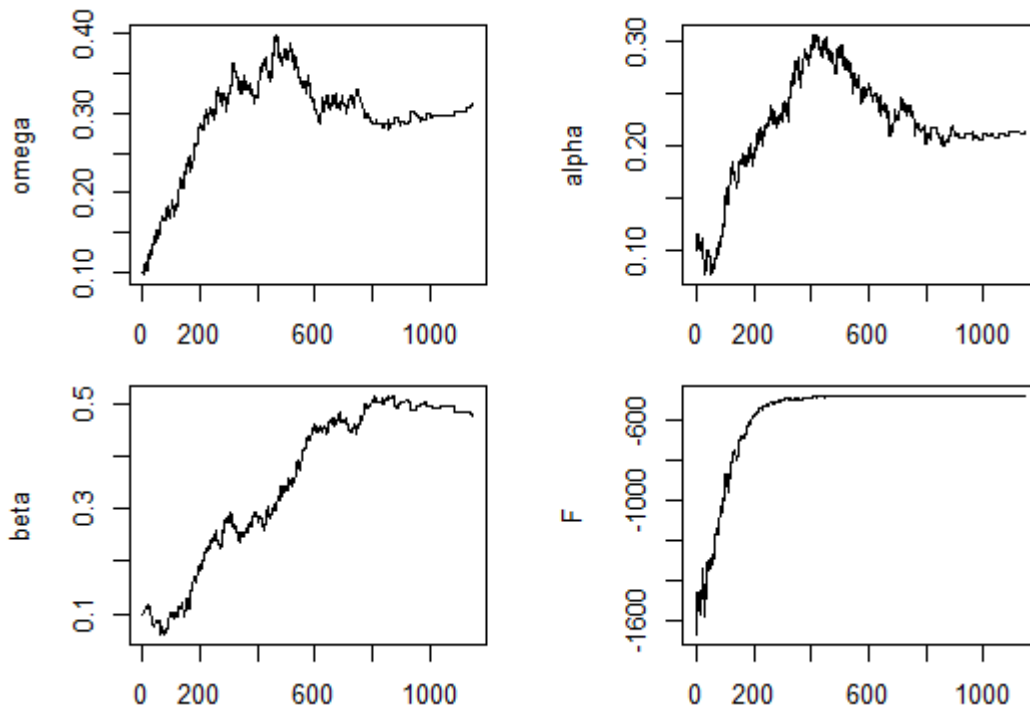
α	$MSE \hat{\omega}$	$MSE \hat{\alpha}_1$	$MSE \hat{\beta}_1$	Χρόνος εκτέλεσης
0.95	0.01681955	0.004908674	0.05033509	6.48 min
0.98	0.008932202	0.001673976	0.01151172	15.77 min
0.99	0.008917513	0.001809727	0.01268517	32.41 min
0.999	0.009738289	0.001992588	0.01454933	5.74 h

Πίνακας 4.3 Το μέσο τετραγωνικό σφάλμα των εκτιμήσεων συναρτήσεως του α και ο συνολικός τρόπος εκτέλεσης του αλγορίθμου.

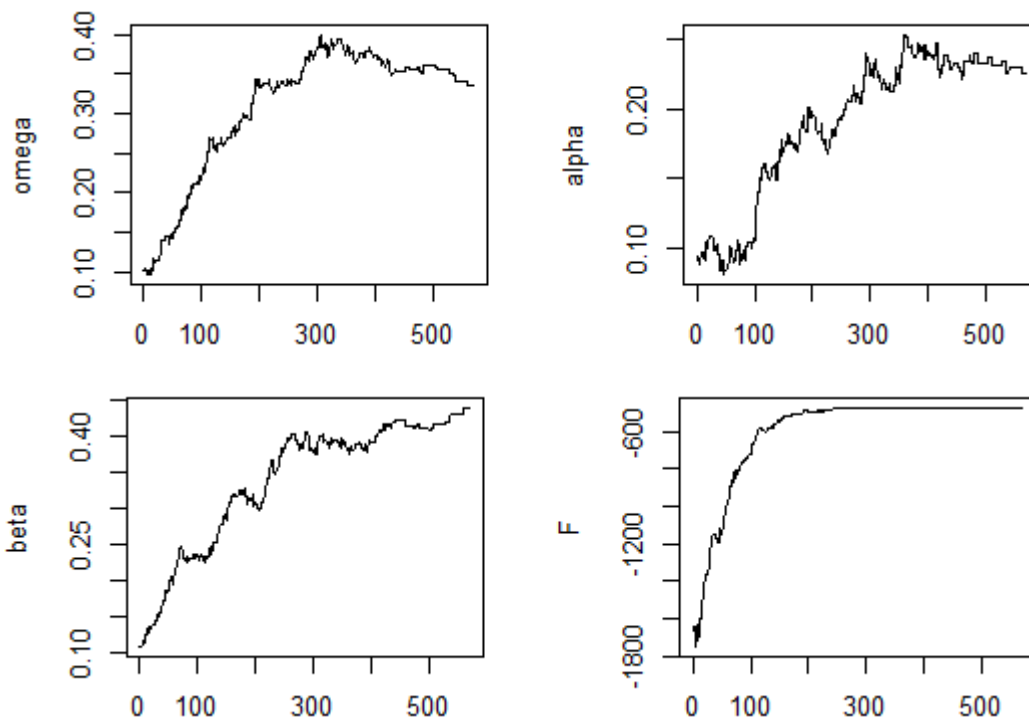
Στη συνέχεια, παραθέτουμε τα γραφήματα των εκτιμήσεων των παραμέτρων ω , a_1 , β_1 και F συναρτήσει του αριθμού των επαναλήψεων που προκύπτουν από μία πραγματοποίηση του αλγορίθμου για τις διάφορες τιμές του α .



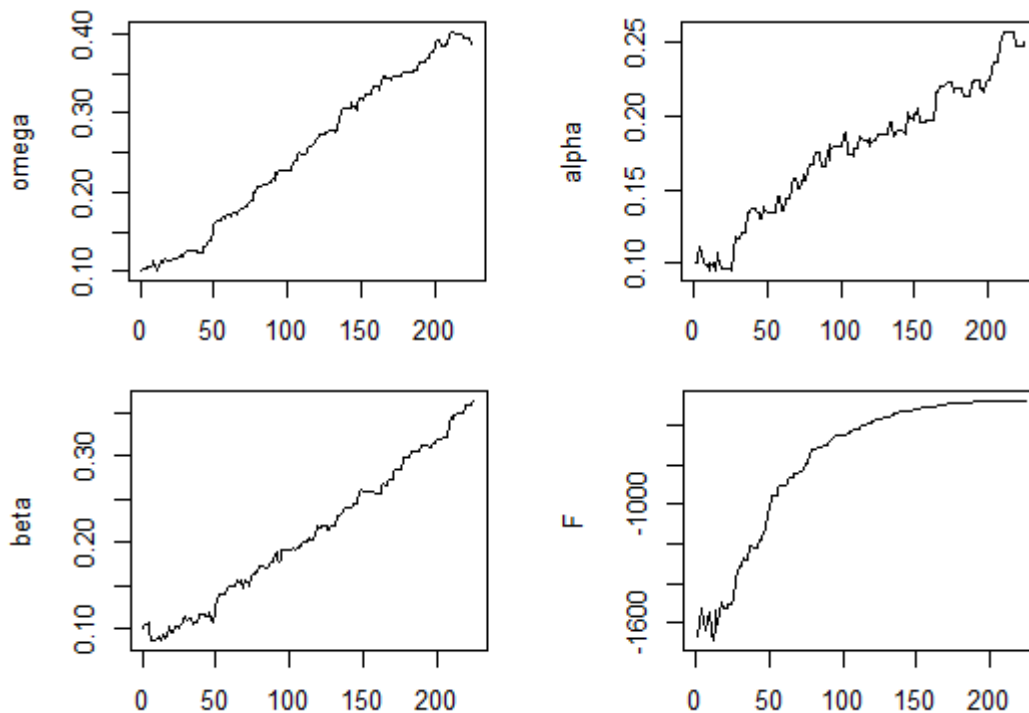
Σχήμα 4.1 Γράφημα της εκτίμησης των παραμέτρων ω , a_1 , β_1 και της F συναρτήσει του αριθμού των επαναλήψεων για $\alpha = 0.999$.



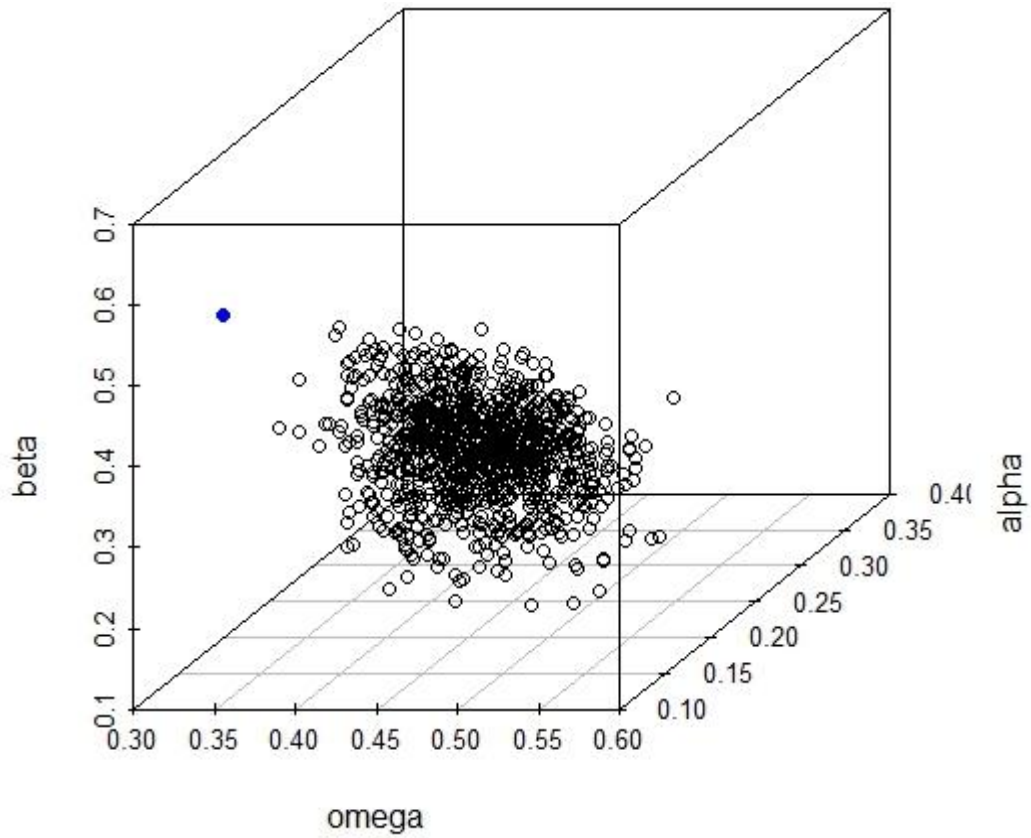
Σχήμα 4.2 Γράφημα της εκτίμησης των παραμέτρων ω , a_1 , β_1 και της F συναρτήσει του αριθμού των επαναλήψεων για $\alpha = 0.99$.



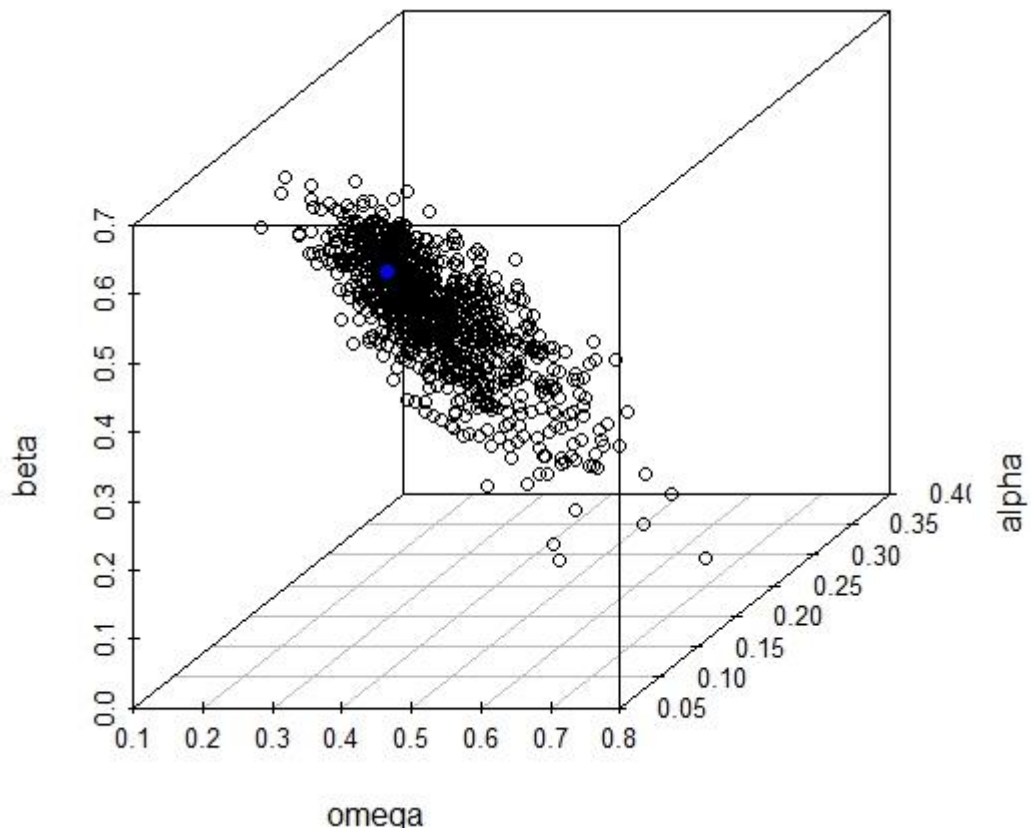
Σχήμα 4.3 Γράφημα της εκτίμησης των παραμέτρων ω, a_1, β_1 και της F συναρτήσεως του αριθμού των επαναλήψεων για $\alpha = 0.98$.



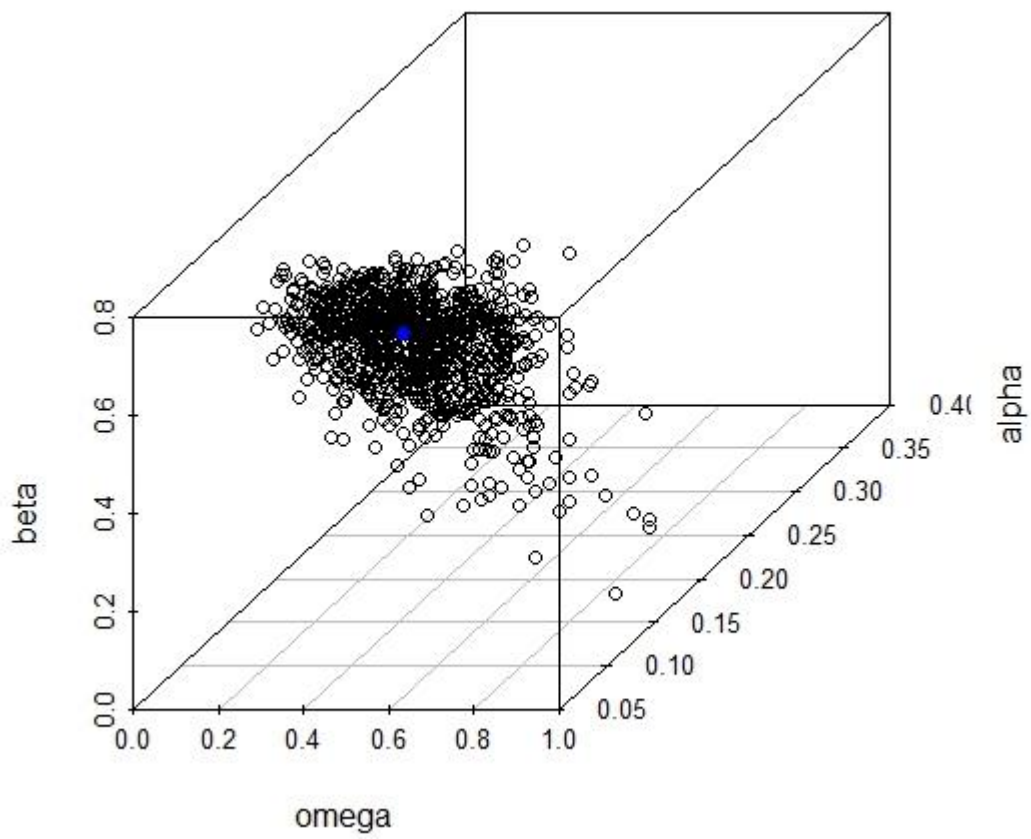
Σχήμα 4.4 Γράφημα της εκτίμησης των παραμέτρων ω, a_1, β_1 και της F συναρτήσεως του αριθμού των επαναλήψεων για $\alpha = 0.95$.



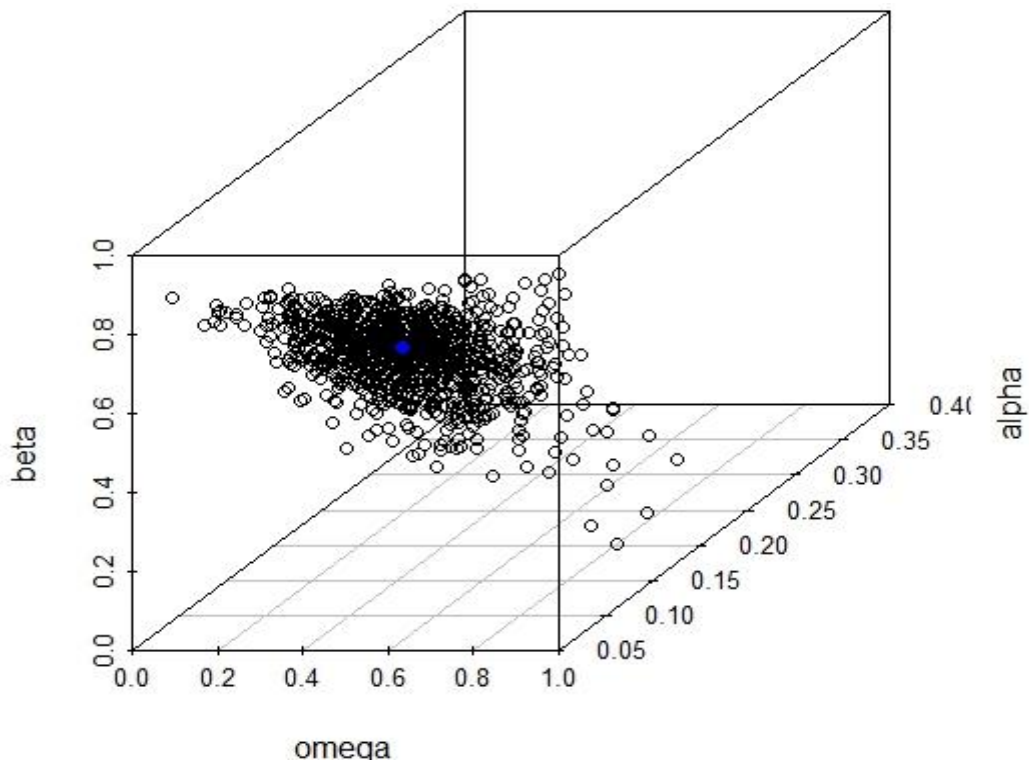
Σχήμα 4.5 Το διάγραμμα διασποράς των εκτιμήσεων που προκύπτουν για $\alpha = 0.95$.



Σχήμα 4.6 Το διάγραμμα διασποράς των εκτιμήσεων που προκύπτουν για $\alpha = 0.98$.



Σχήμα 4.7 Το διάγραμμα διασποράς των εκτιμήσεων που προκύπτουν για $\alpha = 0.99$.



Σχήμα 4.8 Το διάγραμμα διασποράς των εκτιμήσεων που προκύπτουν για $\alpha = 0.999$.

Παρατηρώντας τα αποτελέσματα από την εκτέλεση του αλγορίθμου προκύπτει ότι: για $\alpha = 0.95$ η θερμοκρασία μειώνεται πολύ γρήγορα και ο αλγόριθμος δεν καταφέρνει να συγκλίνει στις βέλτιστες λύσεις για τις παραμέτρους $\omega, \alpha_1, \beta_1$. Αυτό φαίνεται και από το Σχήμα 4.5 στο οποίο βλέπουμε ότι καμία από τις εκτιμήσεις που προέκυψαν από τον αλγόριθμο δεν έχει προσεγγίσει το σημείο που απεικονίζεται με μπλε χρώμα, το οποίο αντιστοιχεί στις πραγματικές τιμές των παραμέτρων. Σε όλες τις άλλες περιπτώσεις, ο αλγόριθμος μάς δίνει καλές εκτιμήσεις των παραμέτρων. Ειδικότερα, παρατηρούμε ότι, καθώς αυξάνεται το α , οι εκτιμήσεις μας βελτιώνονται, αφού προσεγγίζουν με μεγαλύτερη ακρίβεια τις πραγματικές τιμές. Από τα αποτελέσματα για το *MSE* βλέπουμε ότι για $\alpha = 0.98$ οι τιμές μειώνονται σημαντικά σε σχέση με τις τιμές που προκύπτουν για $\alpha = 0.95$. Ενώ για $\alpha = 0.99$ και $\alpha = 0.999$, παρατηρείται μία μικρή αυξητική τάση των τιμών του *MSE*.

Επιπλέον, παρατηρώντας τα αποτελέσματα βλέπουμε ότι οι διαφορετικές τιμές του α μας οδηγούν σε εκτιμήσεις με πολύ διαφορετικό υπολογιστικό κόστος. Για $\alpha = 0.98$ ο αλγόριθμος πραγματοποιεί, κατά μέσο όρο, 570 επαναλήψεις και χρειάστηκαν συνολικά 15.77 λεπτά για την εκτέλεση του. Ακόμη, για $\alpha = 0.99$ ο αλγόριθμος πραγματοποιεί, κατά μέσο όρο 1146 επαναλήψεις, και ο συνολικός χρόνος εκτέλεσης ανέρχεται σε 32.41 λεπτά. Για $\alpha = 0.999$ ο αλγόριθμος χρειάστηκε 5.74 ώρες για να εκτελεστεί, πραγματοποιώντας κατά μέσο όρο 11496 επαναλήψεις. Ενώ θα πρέπει να σημειώσουμε ότι παρά το τεράστιο υπολογιστικό κόστος που έχουμε για $\alpha = 0.999$ δεν βελτιώνεται ανάλογα η ποιότητα της λύσης μας. Παραδείγματος χάριν, για $\alpha = 0.99$ η μέση τιμή του $\hat{\beta}_1$ εκτιμάται ότι είναι 0.475305, ενώ για $\alpha = 0.999$ εκτιμάται στο 0.4841582. Κλείνοντας, αυτήν την παράγραφο να αναφέρουμε ότι τα χαρακτηριστικά του Η/Υ που χρησιμοποιήθηκε είναι: (CPU) Intel Core i3-2370M 2.40Ghz, (RAM) 4GB.

Συνοψίζοντας τα αποτελέσματα της παραπάνω διαδικασίας, μπορούμε να πούμε ότι επιβεβαιώνεται και πειραματικά η καθοριστική σημασία της επιλογής της παραμέτρου α στην αποτελεσματικότητα του αλγορίθμου. Πιο συγκεκριμένα, δύο είναι τα βασικά συμπεράσματα σχετικά με την επιλογή τιμής της παραμέτρου α . Πρώτον, αν μειώνεται η θερμοκρασία πολύ γρήγορα (π.χ. η περίπτωση του αλγορίθμου με τις παραμέτρους που δηλώθηκαν παραπάνω για $\alpha = 0.95$), ο αλγόριθμος δεν θα προσεγγίσει τη βέλτιστη λύση. Δεύτερον, μετά από ένα «κατώφλι», η επιλογή ενός πιο

αργού χρονοδιαγράμματος ψύξης έχει υπερβολικά μεγάλο κόστος χωρίς σημαντικές διαφοροποιήσεις ως προς την ακρίβεια των εκτιμήσεων.

4.3.2 Εκτίμηση των παραμέτρων ενός $GARCH(1, 1)$ υποδείγματος για διάφορες τιμές της παραμέτρου s

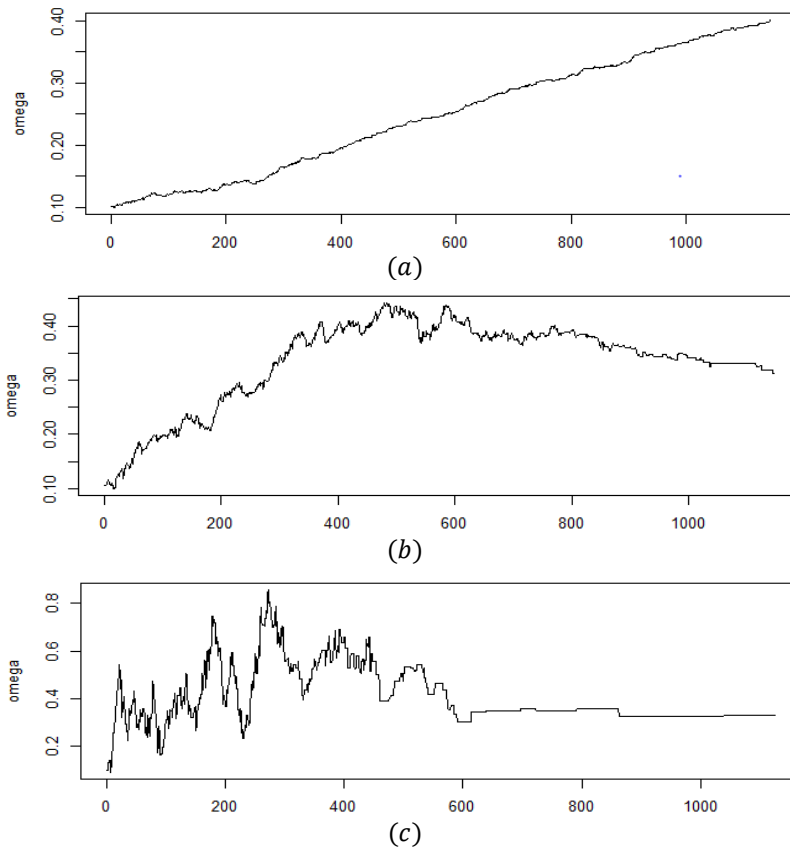
Σε αυτή την παράγραφο ο στόχος μας είναι να δούμε πώς επηρεάζεται η αποτελεσματικότητα του αλγορίθμου, αν μεταβάλλουμε την τυπική απόκλιση της κανονικής κατανομής, μέσω της οποίας παράγουμε τις υποψήφιες λύσεις. Γι' αυτόν τον λόγο, θα εκτελέσουμε 1000 επαναλήψεις του αλγορίθμου για $s = \{0.001, 0.005, 0.05\}$, ενώ οι υπόλοιπες παράμετροι θα παραμένουν ίδιες σε όλες τις εκτελέσεις του αλγορίθμου. Οι τιμές των υπόλοιπων παραμέτρων είναι $T = 100$, $T_{end} = 0.001$, $\alpha = 0.99$ και το αρχικό διάνυσμα παραμέτρων $(\omega_0, \alpha_{1,0}, \beta_{1,0}) = (0.1, 0.1, 0.1)$. Στους ακόλουθους πίνακες βλέπουμε τα αποτελέσματα αυτής της διαδικασίας.

s	μέση τιμή $\hat{\omega}$	τυπική απόκλιση $\hat{\omega}$	μέση τιμή $\hat{\alpha}_1$	τυπική απόκλιση $\hat{\alpha}_1$
0.001	0.4044	0.02076663	0.2544	0.0219588
0.005	0.3205	0.09070014	0.20216	0.04337671
0.05	0.3167	0.09533164	0.20113	0.04488332

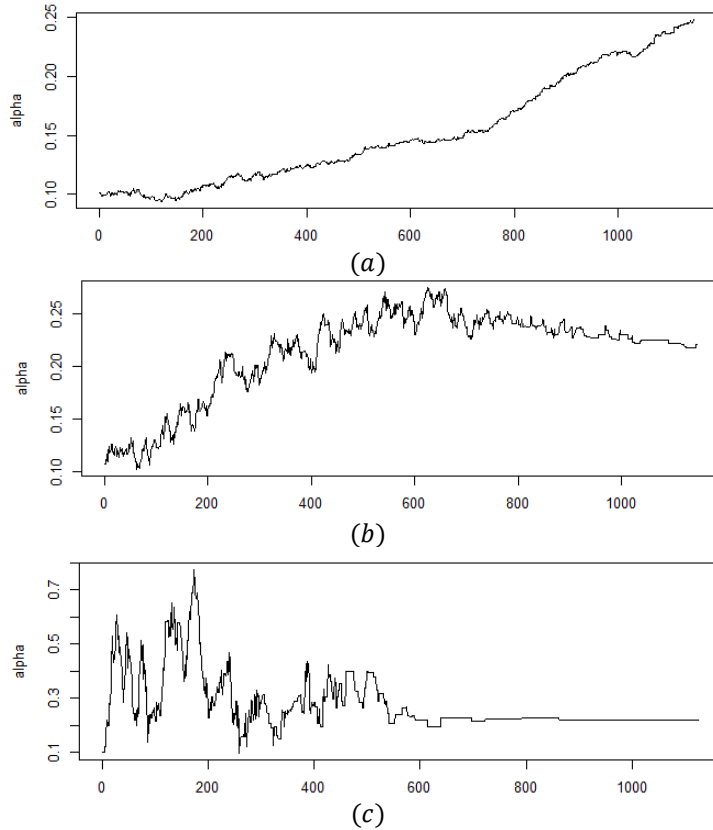
Πίνακας 4.4 Οι εκτιμήσεις των παραμέτρων ω, α_1 , που προκύπτουν από την εκτέλεση 1000 επαναλήψεων του αλγορίθμου της προσομοιωμένης ανόπτησης για $s = \{0.001, 0.005, 0.05\}$.

s	μέση τιμή $\hat{\beta}_1$	τυπική απόκλιση $\hat{\beta}_1$	$MSE\hat{\omega}$	$MSE\hat{\alpha}_1$	$MSE\hat{\beta}_1$
0.001	0.2587	0.02890433	0.0113306	0.0034415	0.2421355
0.005	0.47665	0.1078436	0.0086468	0.0018862	0.0121755
0.05	0.481654	0.1149527	0.0093670	0.0020158	0.0135507

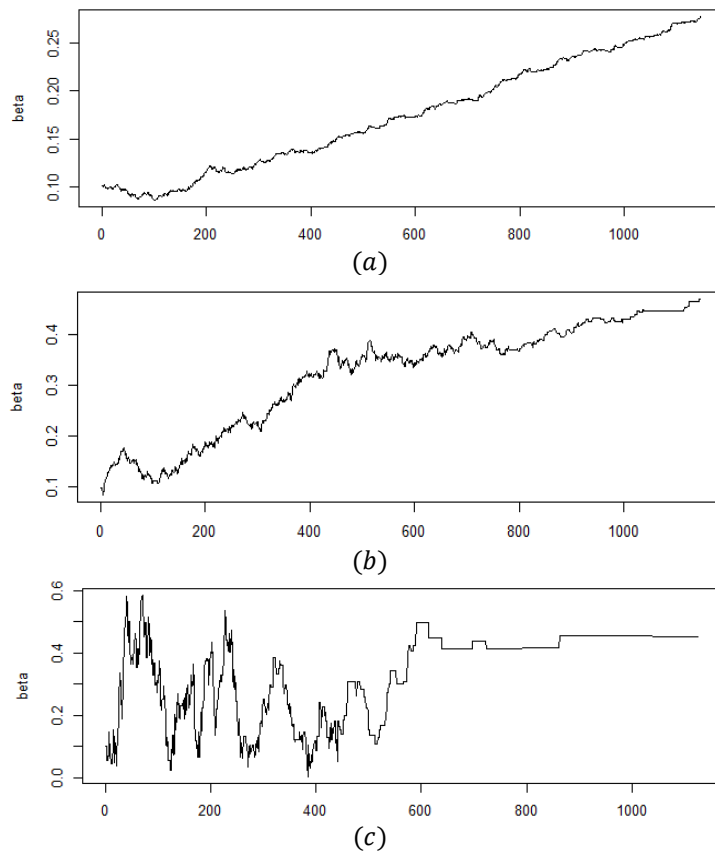
Πίνακας 4.5 Οι εκτιμήσεις της παραμέτρου β_1 και του MSE , που προκύπτουν από την εκτέλεση 1000 επαναλήψεων του αλγορίθμου της προσομοιωμένης ανόπτησης για $s = \{0.001, 0.005, 0.05\}$.



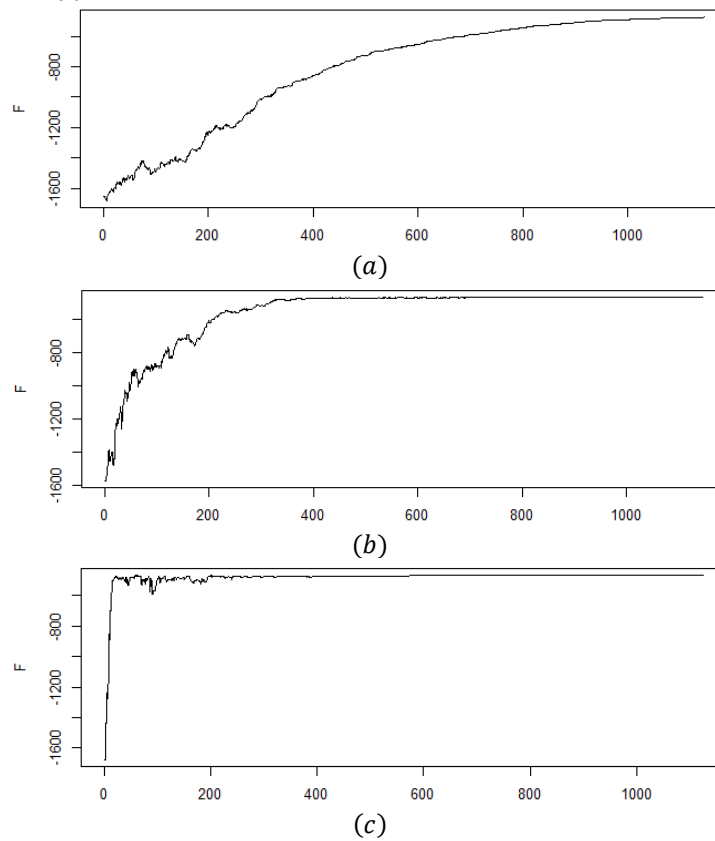
Σχήμα 4.9 Γράφημα της εκτίμησης της παραμέτρου ω συναρτήσει του αριθμού των επαναλήψεων: (a) $s = 0.001$, (b) $s = 0.005$, (c) $s = 0.05$.



Σχήμα 4.10 Γράφημα της εκτίμησης της παραμέτρου α_1 συναρτήσει του αριθμού των επαναλήψεων: (a) $s = 0.001$, (b) $s = 0.005$, (c) $s = 0.05$.



Σχήμα 4.11 Γράφημα της εκτίμησης της παραμέτρου β_1 συναρτήσει του αριθμού των επαναλήψεων: (a) $s = 0.001$, (b) $s = 0.005$, (c) $s = 0.05$.



Σχήμα 4.12 Γράφημα της συνάρτησης F συναρτήσει του αριθμού των επαναλήψεων: (a) $s = 0.001$, (b) $s = 0.005$, (c) $s = 0.05$.

Παρατηρώντας τα αποτελέσματα που παρατίθενται στους Πίνακες 4.4 και 4.5 αλλά και τα γραφήματα των εκτιμήσεων των παραμέτρων $\omega, \alpha_1, \beta_1$ και F συναρτήσει του αριθμού των επαναλήψεων, που προκύπτουν από μία πραγματοποίηση του αλγορίθμου για $s = \{0.001, 0.005, 0.05\}$, καταλήγουμε στα ακόλουθα συμπεράσματα. Για $s = 0.001$ ο αλγόριθμος δεν μας δίνει καλές εκτιμήσεις για τις παραμέτρους. Αυτό συμβαίνει, διότι, εξαιτίας της μικρής τιμής του s , έχουμε μικρές μετακινήσεις στον χώρο καταστάσεων και ο αλγόριθμος τερματίζει πριν συγκλίνει στη βέλτιστη λύση. Έτσι, λοιπόν, αν θέλουμε να χρησιμοποιήσουμε τη συγκεκριμένη τιμή του s , θα πρέπει να αυξήσουμε είτε το T είτε το α , προκειμένου να δώσουμε στον αλγόριθμο αρκετό χρόνο, ούτως ώστε να συγκλίνει σε κάποια λύση. Στις άλλες δύο περιπτώσεις, οι εκτιμήσεις των παραμέτρων είναι πολύ κοντά στις πραγματικές τιμές.

Ακόμη, παρατηρώντας τα γραφήματα βλέπουμε ότι ο τρόπος που ο αλγόριθμος εξερευνά τον χώρο των λύσεων διαφέρει σε κάθε περίπτωση. Ειδικότερα, για $s = 0.05$ παρατηρούμε ότι ο αλγόριθμος πραγματοποιεί αρχικά μεγάλες μετακινήσεις στον χώρο καταστάσεων. Ωστόσο, μετά τις 600 επαναλήψεις ο αλγόριθμος απορρίπτει τις περισσότερες από τις υποψήφιες λύσεις, με αποτέλεσμα να παραμένει στην ίδια λύση. Από την άλλη, για $s = 0.005$ ο αλγόριθμος πραγματοποιεί μικρές μετακινήσεις στον χώρο καταστάσεων

4.3.3 Σύγκριση μεταξύ των μεθόδων Quasi-Newton και προσομοιωμένης ανόπτησης

Είναι γνωστό ότι η εκτίμηση των παραμέτρων ενός υποδείγματος $GARCH(1,1)$ μπορεί να πραγματοποιηθεί και με άλλες μεθόδους εκτός της προσομοιωμένης ανόπτησης. Γι' αυτόν τον λόγο, σε αυτή την ενότητα, θα εκτιμήσουμε τις παραμέτρους του υποδείγματος χρησιμοποιώντας τη μέθοδο Quasi-Newton αλλά και τη μέθοδο της προσομοιωμένης ανόπτησης. Έτσι, θα μπορέσουμε να συγκρίνουμε την αποτελεσματικότητα των δύο μεθόδων.

Πιο συγκεκριμένα, θα προσομοιώσουμε την πραγματοποίηση 1000 υποδειγμάτων $GARCH(1,1)$ με παραμέτρους $(\omega, \alpha_1, \beta_1) = (0.3, 0.2, 0.5)$ και, στη συνέχεια, για κάθε ένα από αυτά θα εκτιμήσουμε τις παραμέτρους και με τις δύο

μεθόδους. Η συγκεκριμένη διαδικασία θα επαναληφθεί τρεις φορές, με διαφορετικές επιλογές αρχικών σημείων σε κάθε περίπτωση.

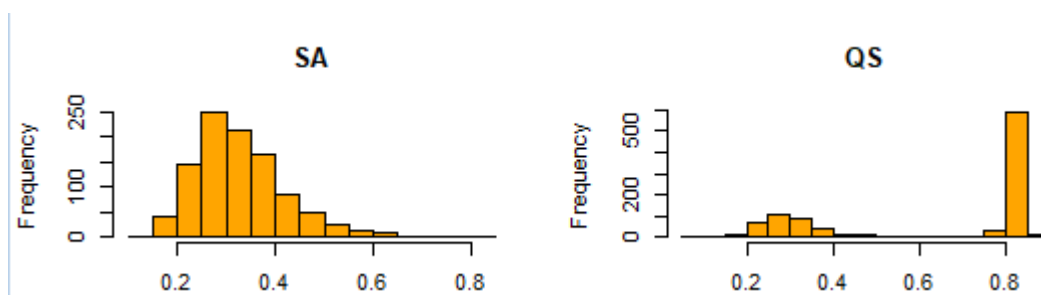
1^η Περίπτωση

Στην πρώτη περίπτωση, θα χρησιμοποιήσουμε ως αρχικό διάνυσμα παραμέτρων το $(\omega_0, \alpha_{1,0}, \beta_{1,0}) = (0.1, 0.1, 0.1)$ και για τις δύο μεθόδους, ενώ οι υπόλοιπες παράμετροι του αλγορίθμου της προσομοιωμένης ανόπτωσης είναι $T = 100$, $T_{end} = 0.001$, $s = 0.005$ και $\alpha = 0.99$. Η εφαρμογή της μεθόδου Quasi-Newton θα γίνει μέσω της συνάρτησης *GARCH*, που περιλαμβάνεται στο πακέτο *tseries* της R.

Στους ακόλουθους πίνακες παραθέτουμε τα περιγραφικά χαρακτηριστικά της ακολουθίας των εκτιμήσεων, που προέκυψαν και από τις δύο μεθόδους:

	$\hat{\omega}$	
	<i>SA</i>	<i>QS</i>
<i>min</i>	0.1310	0.1418
<i>mean</i>	0.3275	0.6145
<i>max</i>	0.7426	0.8519
<i>sd</i>	0.09359158	0.2606245

Πίνακας 4.6 Τα περιγραφικά χαρακτηριστικά για την ακολουθία των εκτιμήσεων της παραμέτρου $\hat{\omega}$, που παράγονται από 1000 επαναλήψεις των αλγορίθμων *SA* και *QS* με αρχικό διάνυσμα παραμέτρων το $(\omega_0, \alpha_{1,0}, \beta_{1,0}) = (0.1, 0.1, 0.1)$.



Σχήμα 4.13 Τα ιστογράμματα της κατανομής των εκτιμήσεων του ω , που παρήχθησαν με τη μέθοδο *SA* και τη μέθοδο *QS* και αρχικό διάνυσμα παραμέτρων το $(\omega_0, \alpha_{1,0}, \beta_{1,0}) = (0.1, 0.1, 0.1)$.

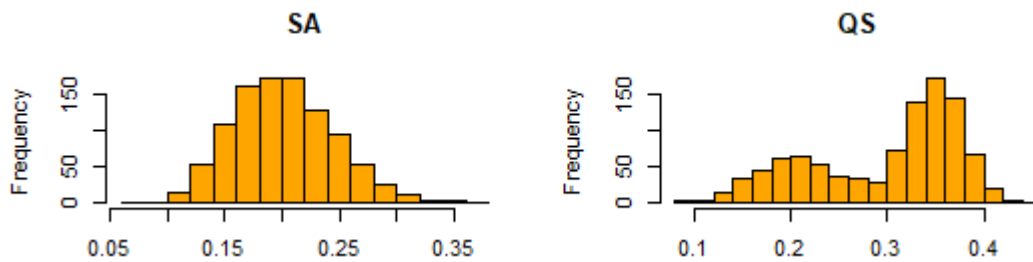
Για την παράμετρο ω παρατηρούμε ότι, χρησιμοποιώντας τη μέθοδο *SA*, η μέση τιμή του $\hat{\omega}$ εκτιμάται 0.3275, η τυπική απόκλιση εκτιμάται 0.09359158 και το εύρος των τιμών της εκτίμησης του $\hat{\omega}$ κυμαίνεται μεταξύ των 0.1310 και 0.7426. Από την άλλη, σύμφωνα με τη μέθοδο *QS*, η μέση τιμή του $\hat{\omega}$ εκτιμάται 0.6145, η τυπική

απόκλιση εκτιμάται 0.2606245 και το εύρος των τιμών της εκτίμησης του $\hat{\omega}$ κυμαίνεται μεταξύ 0.1418 και 0.8519.

Αυτό που επιθυμούμε είναι οι μέσες τιμές να βρίσκονται κοντά στην πραγματική τιμή της παραμέτρου, δηλαδή 0.3, και να έχουν όσο το δυνατόν μικρότερες τυπικές αποκλίσεις. Επομένως, μπορούμε να πούμε ότι η μέθοδος SA μας δίνει καλά αποτελέσματα για το ω , αφού ικανοποιεί τις παραπάνω προϋποθέσεις. Αντιθέτως, οι εκτιμήσεις που προκύπτουν από τη μέθοδο QS δεν προσεγγίζουν την πραγματική τιμή της παραμέτρου. Αυτό γίνεται, διότι, όπως φαίνεται και στο Σχήμα 4.13, σε ένα μεγάλο ποσοστό των επαναλήψεων γίνεται λανθασμένη εκτίμηση της παραμέτρου (κοντά στο 0.8).

	\hat{a}_1	
	SA	QS
<i>min</i>	0.02364	0.07681
<i>mean</i>	0.20135	0.29334
<i>max</i>	0.36483	0.43852
<i>sd</i>	0.04504164	0.07679961

Πίνακας 4.7 Τα περιγραφικά χαρακτηριστικά για την ακολουθία των εκτιμήσεων της παραμέτρου a_1 , που παράγονται από 1000 επαναλήψεις των αλγορίθμων SA και QS με αρχικό διάνυσμα παραμέτρων το $(\omega, a_1, \beta_1) = (\omega_0, a_{1,0}, \beta_{1,0})$.



Σχήμα 4.14 Τα ιστογράμματα της κατανομής των εκτιμήσεων του a_1 , που παρήχθησαν με τη μέθοδο SA και τη μέθοδο QS και αρχικό διάνυσμα παραμέτρων το $(\omega_0, a_{1,0}, \beta_{1,0}) = (0.1, 0.1, 0.1)$.

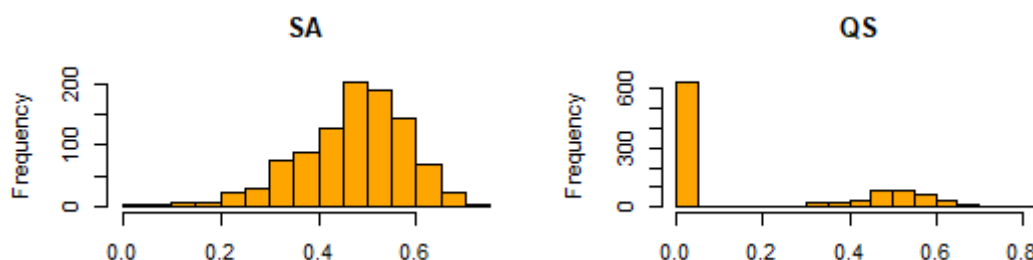
Για την παράμετρο a_1 χρησιμοποιώντας τη μέθοδο SA παρατηρούμε ότι η μέση τιμή του \hat{a}_1 εκτιμάται 0.20135, η τυπική απόκλιση εκτιμάται 0.04504164 και το εύρος των εκτιμήσεων του \hat{a}_1 κυμαίνεται μεταξύ 0.02364 και 0.36483. Από την άλλη, σύμφωνα με τη μέθοδο QS, η μέση τιμή του \hat{a}_1 εκτιμάται 0.29334, η τυπική απόκλιση

εκτιμάται 0.07679961 και το εύρος των εκτιμήσεων του \hat{a}_1 κυμαίνεται μεταξύ 0.07681 και 0.43852.

Έτσι, λοιπόν, και σε αυτή την περίπτωση, μπορούμε να χαρακτηρίσουμε ικανοποιητικά τα αποτελέσματα του αλγορίθμου SA. Ωστόσο, δεν συμβαίνει το ίδιο με τα αποτελέσματα του QS. Όπως βλέπουμε και από το ιστόγραμμα της κατανομής των εκτιμητών του α_1 (Σχήμα 4.14), έχουμε δύο κορυφές, και μάλιστα η μικρότερη από αυτές αντιστοιχεί στο 0.2, που είναι η πραγματική τιμή της παραμέτρου. Επίσης, φαίνεται ότι σε ένα μεγάλο ποσοστό των επαναλήψεων γίνεται λανθασμένη εκτίμηση της παραμέτρου (κοντά στο 0.35).

	$\hat{\beta}_1$	
	SA	QS
<i>min</i>	0.0005338	0.0000
<i>mean</i>	0.4693466	0.2027
<i>max</i>	0.7160110	0.7368
<i>sd</i>	0.1122564	0.2552294

Πίνακας 4.8 Τα περιγραφικά χαρακτηριστικά για την ακολουθία των εκτιμήσεων της παραμέτρου β_1 , που παράγονται από 1000 επαναλήψεις των αλγορίθμων SA και QS με αρχικό διάνυσμα παραμέτρων το $(\omega, a_1, \beta_1) = (0.1, 0.1, 0.1)$.



Σχήμα 4.15 Τα ιστογράμματα της κατανομής των εκτιμήσεων του β_1 , που παρήχθησαν με τη μέθοδο SA και τη μέθοδο QS και αρχικό διάνυσμα παραμέτρων το $(\omega, a_1, \beta_1) = (0.1, 0.1, 0.1)$.

Για την παράμετρο β_1 η μέση τιμή της, που προκύπτει από τον αλγόριθμο SA, εκτιμάται 0.4693466, η τυπική απόκλιση εκτιμάται 0.1122564 και το εύρος των εκτιμήσεων του $\hat{\beta}_1$ κυμαίνεται μεταξύ 0.0005338 και 0.7160110. Ενώ, σύμφωνα με τη μέθοδο QS, η μέση τιμή της παραμέτρου $\hat{\beta}_1$ εκτιμάται 0.2027, η τυπική απόκλιση εκτιμάται 0.2552294 και το εύρος των εκτιμήσεων του $\hat{\beta}_1$ κυμαίνεται μεταξύ 0 και

0.7368. Συνεπώς, παρατηρούμε ότι, και σε αυτήν την παράμετρο, μόνο ο αλγόριθμος SA προσεγγίζει την πραγματική τιμή της παραμέτρου β_1 . Αντιθέτως, για τις εκτιμήσεις του β_1 που παράγονται με τη μέθοδο QS βλέπουμε ότι σε ένα μεγάλο ποσοστό των επαναλήψεων γίνεται λανθασμένη εκτίμηση της παραμέτρου, κοντά στο 0.

Συνοψίζοντας τα αποτελέσματα που προκύπτουν από την εκτέλεση των δύο μεθόδων, βλέπουμε ότι, για όλες τις παραμέτρους, ο αλγόριθμος SA καταφέρνει σε όλες τις επαναλήψεις να προσεγγίσει τις πραγματικές τιμές, σε αντίθεση με τη μέθοδο QS. Ωστόσο, αυτό πιθανόν να οφείλεται στην επιλογή του αρχικού διάνυσματος, καθώς είναι γνωστό ότι η συγκεκριμένη επιλογή επηρεάζει την απόδοση της μεθόδου QS.

2^η Περίπτωση

Εδώ πρόκειται να επαναλάβουμε την παραπάνω διαδικασία, χρησιμοποιώντας όμως διαφορετικό αρχικό διάνυσμα παραμέτρων. Οι παράμετροι των μεθόδων που θα χρησιμοποιήσουμε είναι: $T = 100$, $T_{end} = 0.001$, $s = 0.005$, $\alpha = 0.99$ και αρχικό διάνυσμα παραμέτρων το $(\omega, a_1, \beta_1) = (0.5, 0.1, 0.8)$.

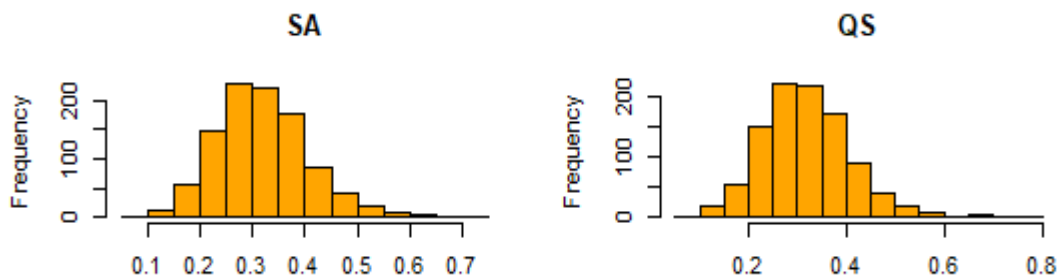
Στη συνέχεια, παρουσιάζουμε τα αποτελέσματα που προκύπτουν από αυτή τη διαδικασία:

	$\hat{\omega}$		$\hat{\alpha}_1$	
	SA	QS	SA	QS
<i>min</i>	0.1097	0.08682	0.07112	0.06724
<i>mean</i>	0.3167	0.31887	0.19960	0.19977
<i>max</i>	0.7503	0.78068	0.34191	0.34112
<i>sd</i>	0.08814322	0.09703225	0.04473739	0.04508281

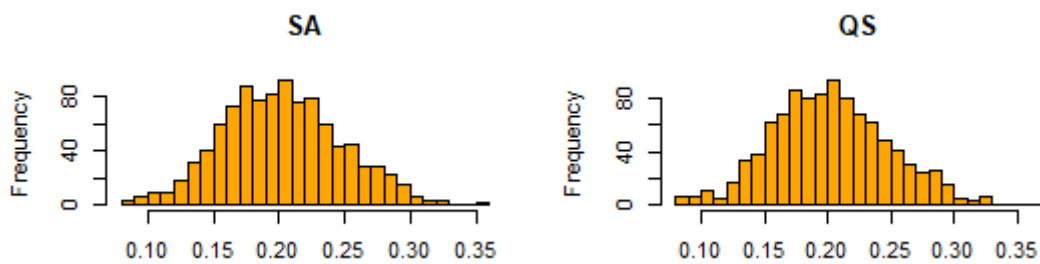
Πίνακας 4.9 Τα περιγραφικά χαρακτηριστικά για την ακολουθία των εκτιμήσεων της παραμέτρου ω, α_1 που παράγονται από 1000 επαναλήψεις των αλγορίθμων SA και QS με αρχικό διάνυσμα παραμέτρων το $(\omega, a_1, \beta_1) = (0.5, 0.1, 0.8)$.

	$\hat{\beta}_1$	
	SA	QS
<i>min</i>	0.002119	0.0000
<i>mean</i>	0.482905	0.4805
<i>max</i>	0.769637	0.8449
<i>sd</i>	0.111071	0.1207606

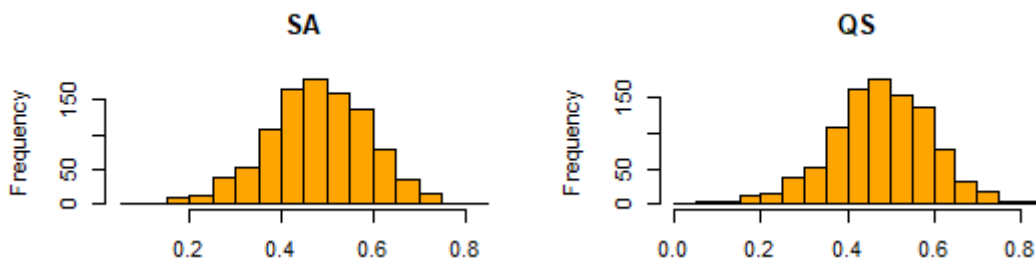
Πίνακας 4.10 Τα περιγραφικά χαρακτηριστικά για την ακολουθία των εκτιμήσεων της παραμέτρου β_1 , που παράγονται από 1000 επαναλήψεις των αλγορίθμων SA και QS με αρχικό διάνυσμα παραμέτρων το $(\omega, a_1, \beta_1) = (0.5, 0.1, 0.8)$.



Σχήμα 4.16 Τα ιστογράμματα της κατανομής των εκτιμήσεων του ω , που παρήχθησαν με τις μεθόδους SA και QS με αρχικό διάνυσμα παραμέτρων το $(\omega, a_1, \beta_1) = (0.5, 0.1, 0.8)$



Σχήμα 4.17 Τα ιστογράμματα της κατανομής των εκτιμήσεων του a_1 , που παρήχθησαν με τις μεθόδους SA και QS και αρχικό διάνυσμα παραμέτρων το $(\omega, a_1, \beta_1) = (0.5, 0.1, 0.8)$.



Σχήμα 4.18 Τα ιστογράμματα της κατανομής των εκτιμήσεων του β_1 , που παρήχθησαν με τις μεθόδους SA και QS και αρχικό διάνυσμα παραμέτρων το $(\omega, a_1, \beta_1) = (0.5, 0.1, 0.8)$.

Παρατηρώντας τα αποτελέσματα που προκύπτουν από την εφαρμογή των μεθόδων SA και QS, βλέπουμε ότι, και για τις τρεις παραμέτρους, οι μέσες τιμές των εκτιμήσεών τους αποκλίνουν ελάχιστα από τις πραγματικές τιμές των παραμέτρων. Επιπλέον, οι εκτιμητές που προκύπτουν μέσω της μεθόδου SA έχουν μικρότερες τυπικές

αποκλίσεις συγκρινόμενοι με τους αντίστοιχους εκτιμητές που προκύπτουν μέσω της μεθόδου *QS*.

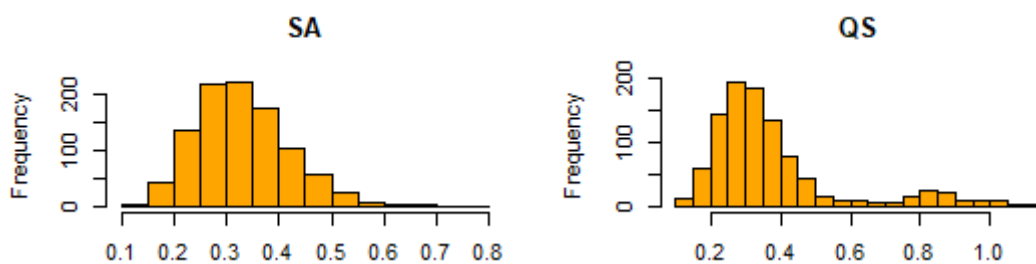
Έτσι, λοιπόν, μετά την εφαρμογή των μεθόδων για δύο διαφορετικά αρχικά σημεία, είναι ξεκάθαρο ότι η μέθοδος *SA* λειτουργεί καλά, ανεξαρτήτως της επιλογής αρχικού σημείου, ενώ η απόδοση της μεθόδου *QS* εξαρτάται από αυτή την επιλογή. Γι' αυτόν τον λόγο, θα επαναλάβουμε τη διαδικασία με τυχαία επιλογή αρχικού σημείου σε κάθε επανάληψη. Με αυτόν τον τρόπο, επιχειρούμε να εξαλείψουμε την επίδραση της συγκεκριμένης επιλογής στην απόδοση των αλγορίθμων.

3^η Περίπτωση

Σε αυτή την περίπτωση θα εφαρμόσουμε τις δύο μεθόδους χρησιμοποιώντας ως αρχικό διάνυσμα παραμέτρων τυχαίους αριθμούς από τη κατανομή $U(0,1)$. Από την άλλη πλευρά, οι υπόλοιπες παράμετροι του αλγορίθμου της προσομοιωμένης ανόπτησης είναι $T = 100$, $T_{end} = 0.001$, $s = 0.005$ και $\alpha = 0.99$, όπως και τις προηγούμενες φορές. Στη συνέχεια, ακολουθούν τα αποτελέσματα που προέκυψαν από αυτή τη διαδικασία.

	$\hat{\omega}$	
	<i>SA</i>	<i>QS</i>
<i>min</i>	0.1326	0.1032
<i>mean</i>	0.3327	0.3806
<i>max</i>	0.7645	1.1184
<i>sd</i>	0.09367538	0.1982291

Πίνακας 4.11 Τα περιγραφικά χαρακτηριστικά για την ακολουθία των εκτιμήσεων της παραμέτρου ω , που παράγονται από 1000 επαναλήψεις των αλγορίθμων *SA* και *QS* με αρχικό διάνυσμα παραμέτρων τυχαίους αριθμούς από την κατανομή $U(0,1)$.



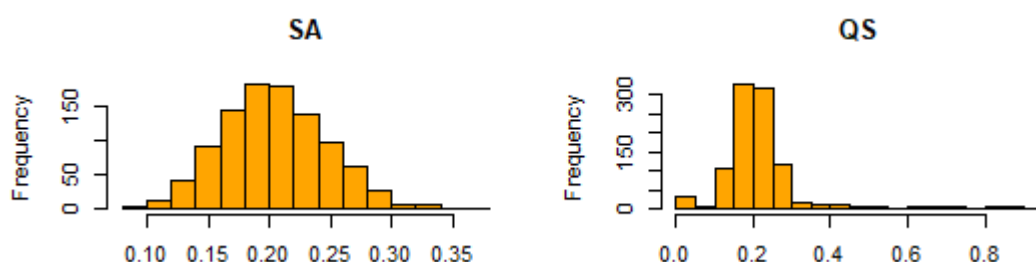
Σχήμα 4.15 Τα ιστογράμματα της κατανομής των εκτιμήσεων του ω , που παρήχθησαν με τη μέθοδο SA και τη μέθοδο QS και αρχικό διάνυσμα παραμέτρων τυχαίους αριθμούς από την κατανομή $U(0,1)$.

Σύμφωνα με τη μέθοδο SA, η μέση τιμή του $\hat{\omega}$ εκτιμάται 0.3327, η τυπική απόκλιση εκτιμάται 0.09367538 και το εύρος των τιμών της εκτίμησης του ω κυμαίνεται μεταξύ 0.1326 και 0.7645. Από την άλλη, σύμφωνα με τη μέθοδο QS, η μέση τιμή του $\hat{\omega}$ εκτιμάται 0.3806, η τυπική απόκλιση εκτιμάται 0.1982291 και το εύρος των τιμών της εκτίμησης του ω κυμαίνεται μεταξύ 0.1032 και 1.1184.

Συνεπώς, μπορούμε να πούμε ότι η μέθοδος SA μας δίνει καλύτερες εκτιμήσεις από τη μέθοδο QS. Αυτό φαίνεται και από τα ιστογράμματα των εκτιμήσεων, όπου βλέπουμε ότι για την κατανομή που παράγεται με τη μέθοδο QS υπάρχει ένα μικρό ποσοστό λανθασμένων εκτιμήσεων, κοντά στο 0.8. Από την άλλη, αν και η κατανομή των εκτιμήσεων που παράγεται με τη μέθοδο SA είναι λοξή προς τα δεξιά, προσεγγίζει καλύτερα την πραγματική τιμή της παραμέτρου.

	\hat{a}_1	
	SA	QS
<i>min</i>	0.0806	0.0000
<i>mean</i>	0.2046	0.2256
<i>max</i>	0.3751	0.9302
<i>sd</i>	0.04334419	0.1266237

Πίνακας 4.12 Τα περιγραφικά χαρακτηριστικά για την ακολουθία των εκτιμήσεων της παραμέτρου a_1 , που παράγονται από 1000 επαναλήψεις των αλγορίθμων SA και QS με αρχικό διάνυσμα παραμέτρων τυχαίους αριθμούς από την κατανομή $U(0,1)$.



Σχήμα 4.16 Τα ιστογράμματα της κατανομής των εκτιμήσεων του a_1 , που παρήχθησαν με τις μεθόδους SA και QS και αρχικό διάνυσμα παραμέτρων τυχαίους αριθμούς από την κατανομή $U(0,1)$.

Χρησιμοποιώντας τη μέθοδο SA η μέση τιμή του \hat{a}_1 εκτιμάται 0.2046, η τυπική απόκλιση εκτιμάται 0.04334419 και το εύρος των εκτιμήσεων του a_1 κυμαίνεται

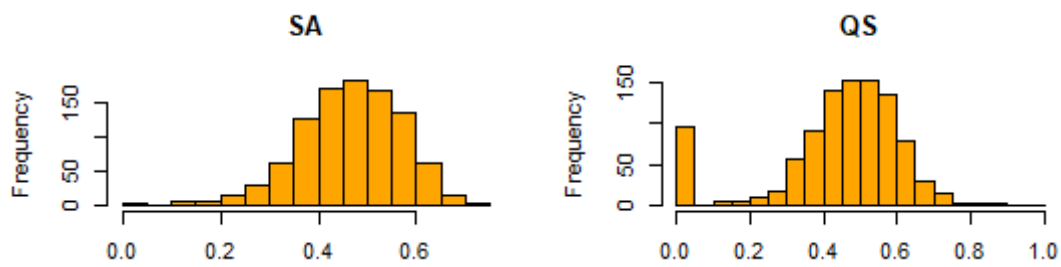
μεταξύ 0.0806 και 0.3751. Από την άλλη, σύμφωνα με τη μέθοδο *QS*, η μέση τιμή του \hat{a}_1 εκτιμάται 0.2256, η τυπική απόκλιση εκτιμάται 0.1266237 και το εύρος των εκτιμήσεων του a_1 κυμαίνεται μεταξύ 0.0000 και 0.9302.

Για την παράμετρο a_1 και οι δύο μέθοδοι προσεγγίζουν την πραγματική τιμή της παραμέτρου. Βέβαια, οι εκτιμήσεις της μεθόδου *SA* είναι καλύτερες και ως προς τις μέσες τιμές αλλά και ως προς τις τυπικές αποκλίσεις. Αυτό είναι λογικό, αφού στο ιστογράμμα των εκτιμήσεων της μεθόδου *QS* βλέπουμε ότι υπάρχει ένας μικρός αριθμός λανθασμένων εκτιμήσεων σε διάφορα σημεία. Ενώ, από την άλλη πλευρά, οι εκτιμήσεις της μεθόδου *SA* κατανέμονται συμμετρικά γύρω από το 0.2.

	$\hat{\beta}_1$	
	<i>SA</i>	<i>QS</i>
<i>min</i>	0.0007132	0.0000
<i>mean</i>	0.4628226	0.4414
<i>max</i>	0.7246728	0.9601
<i>sd</i>	0.1087074	0.1813755

Πίνακας 4.13 Τα περιγραφικά χαρακτηριστικά για την ακολουθία των εκτιμήσεων της παραμέτρου β_1 , που παράγονται από 1000 επαναλήψεις των αλγορίθμων *SA* και *QS* με αρχικό διάνυσμα παραμέτρων τυχαίους αριθμούς από την κατανομή $U(0,1)$.

Η μέση τιμή του $\hat{\beta}_1$ εκτιμάται 0.4628226, σύμφωνα με την μέθοδο *SA*, και η εκτίμηση της τυπικής απόκλισης είναι 0.1087074. Οι αντίστοιχες εκτιμήσεις που προκύπτουν από τη μέθοδο *QS* για τη μέση τιμή και την τυπική απόκλιση είναι 0.4414 και 0.1813755. Επομένως, και για αυτή την παράμετρο, μπορούμε να πούμε ότι οι εκτιμήσεις της μεθόδου *SA* είναι καλύτερες από αυτές της μεθόδου *QS*, διότι προσεγγίζουν καλύτερα την πραγματική τιμή της παραμέτρου και έχουν μικρότερη τυπική απόκλιση. Όπως φαίνεται και από τα ιστογράμματα των εκτιμήσεων που ακολουθούν, υπάρχει ένας σημαντικός αριθμός λανθασμένων εκτιμήσεων της παραμέτρου β_1 από τη μέθοδο *QS*, γύρω στο 0. Από την άλλη, το ποσοστό των εκτιμήσεων της παραμέτρου β_1 γύρω από το 0 μπορεί να χαρακτηριστεί αμελητέο, καθώς σχεδόν όλες οι εκτιμήσεις κατανέμονται συμμετρικά γύρω από το 0.5.



Σχήμα 4.17 Τα ιστογράμματα της κατανομής των εκτιμήσεων του β_1 , που παρήχθησαν με τις μεθόδους SA, QS και αρχικό διάνυσμα παραμέτρων τυχαίους αριθμούς από την κατανομή $U(0,1)$.

Οπότε, σύμφωνα με τα αποτελέσματα και των τριών περιπτώσεων που παρουσιάσαμε μπορούμε να πούμε ότι ο αλγόριθμος SA δίνει καλύτερες εκτιμήσεις από τη μέθοδο QS.

Ανακεφαλαιώνοντας, ο στόχος μας σε αυτό το κεφάλαιο ήταν η εφαρμογή του αλγορίθμου της προσομοιωμένης ανόπτησης για τον προσδιορισμό των παραμέτρων ενός $GARCH(1,1)$ υποδείγματος και η αξιολόγηση της αποτελεσματικότητας του. Γι' αυτόν το λόγο προσομοιώσαμε την πραγματοποίηση ενός $GARCH(1,1)$ υποδείγματος με παραμέτρους $(\omega, a_1, \beta_1) = (0.3, 0.2, 0.5)$ και στη συνέχεια, χρησιμοποιώντας την ακολουθία ε_t προχωρήσαμε στην εφαρμογή του αλγορίθμου. Σύμφωνα, με τα αποτελέσματα που παρουσιάσαμε οι τιμές των παραμέτρων του αλγορίθμου είναι καθοριστικής σημασίας για την αποτελεσματικότητα του.

Από τη θεωρία είναι γνωστό ότι αν το α μειώνεται με γρήγορο ρυθμό τότε ο αλγόριθμος δεν προσεγγίζει τη βέλτιστη λύση. Ωστόσο, είναι προφανές ότι αυτός είναι ένας γενικός κανόνας και η σημασία της φράσης με γρήγορο ρυθμό διαφέρει από εφαρμογή σε εφαρμογή. Ως εκ τούτου, κάναμε διάφορες δοκιμές ώστε να καταλήξουμε σε τιμές του α που οδηγούν σε σύγκλιση του αλγορίθμου στην βέλτιστη λύση για το $GARCH(1,1)$. Όπως είδαμε για τιμές του α από 0.98 και άνω, κρατώντας τις τιμές της θερμοκρασίας και του βήματος της τυπικής απόκλισης σταθερές, οι εκτιμήσεις μας είναι κατά μέσο όρο αρκετά κοντά στις πραγματικές τιμές των παραμέτρων. Με την ακρίβεια τους να βελτιώνεται καθώς αυξάνεται το α . Όσον αφορά, την τυπική απόκλιση της κανονικής κατανομής, που χρησιμοποιούμε για την παραγωγή των υποψήφιων λύσεων, θα πρέπει να είμαστε σίγουροι ότι είναι κατάλληλη για το α που χρησιμοποιούμε. Παραδείγματος χάριν, είδαμε ότι για $\alpha = 0.99$ με $s = 0.001$ ο

αλγόριθμος δεν προσέγγισε τη βέλτιστη λύση. Ακόμη σχετικά με την επιλογή των αρχικών σημείων μπορούμε να πούμε ότι φαίνεται ο αλγόριθμος να προσεγγίζει τη βέλτιστη λύση ανεξαρτήτως αρχικού σημείου. Αν και μία καλή επιλογή επιταχύνει τη σύγκλιση του αλγορίθμου.

Εντούτοις, από την ανάλυση που προηγήθηκε έγινε φανερό ότι τα αποτελέσματα που προκύπτουν από μεμονωμένες εκτελέσεις του αλγορίθμου δεν είναι αξιόπιστα. Παρατηρώντας, τα ιστογράμματα της κατανομής των εκτιμήσεων που προκύπτουν ύστερα από 1000 επαναλήψεις του αλγορίθμου βλέπουμε ότι σε όλες τις περιπτώσεις η τυπική απόκλιση της κατανομής είναι αρκετά μεγάλη.

Κλείνοντας, αυτήν την ενότητα μπορούμε να πούμε ότι ο αλγόριθμος της προσομοιωμένης ανόπτησης μας δίνει κατά μέσο όρο καλές εκτιμήσεις για τις παραμέτρους του υποδείγματος $GARCH(1,1)$, υπό την προϋπόθεση κατάλληλης επιλογής για τα α, s, T .

Παράρτημα

Μαρκοβιανές αλυσίδες

Στοχαστική ανέλιξη είναι μία συλλογή τυχαίων μεταβλητών $\{X(t), t \in T\}$ ορισμένη σ' έναν χώρο πιθανότητας (Ω, F, P) . Συνήθως, το σύνολο T αναφέρεται στον χρόνο και μπορεί να είναι συνεχές ή διακριτό. Το σύνολο των δυνατών τιμών των τ.μ. καλείται χώρος καταστάσεων και συμβολίζεται με S . Το S μπορεί να είναι επίσης συνεχές ή διακριτό.

Ορισμός A.1: Μία στοχαστική ανέλιξη $\{X(t), t \in T\}$ καλείται *Μαρκοβιανή*, αν ισχύει η σχέση

$$\begin{aligned} P(X(t_n) = i_n | X(t_{n-1}) = i_{n-1}, \dots, X(t_1) = i_1) \\ = P(X(t_n) = i_n | X(t_{n-1}) = i_{n-1}) \end{aligned}$$

για κάθε $i_1, i_2, \dots, i_n \in S$, και κάθε $t_1 < t_2 < \dots < t_n \in T$.

Μια Μαρκοβιανή ανέλιξη $\{X(t), t \in T\}$ με διακριτό χώρο καταστάσεων καλείται Μαρκοβιανή αλυσίδα.

Για μία Μ.α.δ.χ. η πιθανότητα $P^{(n-1,n)}(i, j) = P(X_n = j | X_{n-1} = i)$ καλείται πιθανότητα μετάβασης πρώτης τάξης από την κατάσταση i στην κατάσταση j στο n -οστό βήμα. Αν, επιπλέον, $P^{(n-1,n)}(i, j) = P(i, j)$ για κάθε n , η αλυσίδα καλείται ομογενής. Εφόσον ο χώρος καταστάσεων S είναι αριθμήσιμο σύνολο, μπορούμε να συνοψίσουμε τις πιθανότητες μετάβασης πρώτης τάξης σ' έναν πίνακα:

$$\mathbf{P} = \begin{pmatrix} p_{00} & p_{01} & \dots \\ p_{10} & p_{11} & \dots \\ \vdots & \vdots & \ddots \end{pmatrix}$$

Ο πίνακας \mathbf{P} λέγεται πίνακας πιθανοτήτων μετάβασης πρώτης τάξης. Μπορούμε να παρατηρήσουμε ότι τα στοιχεία κάθε γραμμής του πίνακα \mathbf{P} είναι μη αρνητικά και αθροίζουν στη μονάδα και, εξαιτίας αυτού, ο πίνακας \mathbf{P} λέγεται στοχαστικός.

Με $f_i^{(0)} = P\{X_0 = i\}$, $i \in S$ δηλώνουμε τη συνάρτηση πιθανότητας της αρχικής τυχαίας μεταβλητής X_0 , ενώ η πιθανότητα μετάβασης από μία κατάσταση i σε μία κατάσταση j σε n βήματα λέγεται πιθανότητα μετάβασης n τάξης και συμβολίζεται ως εξής :

$$P^{(n)}(i, j) = P(X_n = j | X_0 = i) \quad i, j \in S.$$

Η $P^{(n)}(i, j)$ είναι ίση με το άθροισμα των πιθανοτήτων όλων των πεπερασμένων ακολουθιών μεταβάσεων, οι οποίες οδηγούν από την κατάσταση i στην κατάσταση j , σε n ακριβώς βήματα. Πράγματι, από το θεώρημα ολικής πιθανότητας έχουμε ότι :

$$\begin{aligned} P^{(n)}(i, j) &= P(X_n = j | X_0 = i) \\ &= \sum_k P(X_n = j | X_{n-1} = k) P(X_{n-1} = k | X_0 = i) \\ &= \sum_k P^{(n)}(i, j) P(k, j) \quad i, j \in S \end{aligned}$$

Ισοδύναμα, υπό μορφή πινάκων, έχουμε ότι:

$$\mathbf{P}^{(n)} = \mathbf{P}^{(n-1)} \mathbf{P}.$$

Αναδρομικά, από την παραπάνω σχέση αποδεικνύεται ότι:

$$\mathbf{P}^{(n)} = \mathbf{P}^n.$$

Η κατανομή της τ.μ. X_n , δεδομένης της αρχικής κατανομής $f_i^{(0)}$, είναι η εξής:

$$\begin{aligned} f_j^{(n)} &= P(X_n = j) \\ &= \sum_i P(X_0 = i) P(X_n = j | X_0 = i) \\ &= \sum_i f_i^{(0)} P^{(n)}(i, j) \quad (j \in S). \end{aligned}$$

Ισοδύναμα, υπό μορφή πινάκων, έχουμε ότι

$$\mathbf{f}^{(n)} = \mathbf{f}^{(0)} \mathbf{P}^n.$$

Ταξινόμηση καταστάσεων

Ορισμός A.2: Θα λέμε ότι η κατάσταση $i \in S$ είναι προσιτή από την κατάσταση $j \in S$, και συμβολίζουμε $j \rightarrow i$, αν $P^{(n)}(i, j) > 0$ για κάποιο $n \geq 0$. Ακόμη, όταν η i είναι προσιτή από την j και η j είναι προσιτή από την i , λέμε ότι οι καταστάσεις επικοινωνούν ($j \leftrightarrow i$).

Η σχέση επικοινωνίας είναι σχέση ισοδυναμίας στον χώρο καταστάσεων S και, επομένως, τον διαμερίζει σε κλάσεις, τις οποίες αποκαλούμε κλάσεις επικοινωνίας. Είναι προφανές ότι οτιδήποτε ισχύει για μία κατάσταση της κλάσης θα ισχύει και για τις υπόλοιπες καταστάσεις της κλάσης.

Ορισμός A.3: Μία κλάση επικοινωνίας C καλείται κλειστή, αν και μόνο αν για κάθε $i \in C$ και $j \in S$, ώστε $i \rightarrow j$ είναι $j \in C$. Αν μία κλάση δεν είναι κλειστή, θα λέγεται ανοικτή.

Ορισμός A.4: Όταν ο χώρος καταστάσεων S είναι μία κλειστή κλάση επικοινωνίας, τότε η ΜΑ λέγεται αδιαχώριστη. Συνεπώς, όλες οι καταστάσεις επικοινωνούν μεταξύ τους.

Ορισμός A.5: Μία κατάσταση $i \in S$ λέγεται περιοδική με περίοδο d , αν ο μ.κ.δ. όλων των $n \geq 1$, για τα οποία $P^{(n)}(i, i) > 0$, είναι ο d . Μία κατάσταση με περίοδο 1 λέγεται απεριοδική.

Στάσιμες κατανομές

Ορισμός A.6: Έστω $\{X_n\}$ Μαρκοβιανή αλυσίδα με χώρο καταστάσεων S και πίνακα πιθανοτήτων μετάβασης P . Ένα διάνυσμα $\pi = (\pi(1), \dots, \pi(k))$ καλείται στάσιμη κατανομή της αλυσίδας, αν ισχύουν τα ακόλουθα:

- $\pi(i) \geq 0$ για $i = \{1, \dots, k\}$ και $\sum_{i=1}^k \pi(i) = 1$
- $\pi P = \pi$

Θεώρημα A.1: Εάν μία Μαρκοβιανή αλυσίδα είναι αδιαχώριστη και απεριοδική, τότε έχει μοναδική στάσιμη κατανομή π και η κατανομή $f^{(n)}$ της αλυσίδας, τη χρονική στιγμή n , συγκλίνει στην π , καθώς $n \rightarrow \infty$, ανεξάρτητα από την αρχική κατανομή $f^{(0)}$.

Αντιστρεψιμότητα

Ορισμός A.7: Έστω $\{X_n\}$ Μαρκοβιανή αλυσίδα με χώρο καταστάσεων S και πίνακα πιθανοτήτων μετάβασης P . Η Μ.α. καλείται αντιστρέψιμη ως προς την κατανομή π , αν για κάθε $i, j \in S$ ισχύει

$$\pi(i)P(i, j) = \pi(j)P(j, i)$$

Θεώρημα A.2: Έστω $\{X_n\}$ Μαρκοβιανή αλυσίδα με χώρο καταστάσεων S και πίνακα

πιθανοτήτων μετάβασης \mathbf{P} . Αν η Μ.α. είναι αντιστρέψιμη ως προς την κατανομή $\boldsymbol{\pi}$, τότε αυτή είναι, επίσης, και η στάσιμη κατανομή της.

Μη ομογενείς Μαρκοβιανές Αλυσίδες

Όπως αναφέραμε και παραπάνω, στις μη ομογενείς ΜΑ ο πίνακας μεταβάσεων που δίνει τις πιθανότητες μετάβασης από την κατάσταση i τη χρονική στιγμή $n - 1$, στην κατάσταση j τη χρονική στιγμή n , εξαρτάται από το n . Εξαιτίας αυτού, ο συγκεκριμένος πίνακας μεταβάσεων συμβολίζεται με \mathbf{P}_n .

Έτσι, λοιπόν, μία μη ομογενής ΜΑ καθορίζεται πλήρως από την αρχική κατανομή $\mathbf{f}^{(0)}$ και από μία ακολουθία $\{\mathbf{P}\}_{n=1}^{\infty}$ από πίνακες μεταβάσεων, δηλαδή

$$\mathbf{f}^{(n)} = \mathbf{f}^{(0)} \mathbf{P}_1 \mathbf{P}_2 \cdots \mathbf{P}_n \text{ και } \mathbf{f}^{(m,n)} = \mathbf{f}^{(0)} \mathbf{P}_{m+1} \mathbf{P}_{m+2} \cdots \mathbf{P}_n.$$

Αυτό που μας ενδιαφέρει είναι η μελέτη της μακροχρόνιας συμπεριφοράς της ΜΑ. Μιλώντας πιο επεξηγηματικά, θέλουμε να προσδιορίσουμε τη συμπεριφορά της κατανομής $\mathbf{f}^{(n)}$ καθώς $n \rightarrow \infty$, και το κατά πόσο η αρχική κατάσταση επηρεάζει ακόμη την εξέλιξη της Μ.α. Γι' αυτόν τον λόγο, θα χρησιμοποιήσουμε τις έννοιες της ασθενούς και ισχυρής εργοδικότητας.

Ορισμός Α.4: Μία μη ομογενής ΜΑ καλείται ασθενώς εργοδική, αν για κάθε m

$$\lim_{k \rightarrow \infty} \sup_{\mathbf{f}^{(0)}, \mathbf{g}^{(0)}} \|\mathbf{f}^{(m,k)} - \mathbf{g}^{(m,k)}\| = 0,$$

όπου $\mathbf{f}^{(0)}, \mathbf{g}^{(0)}$ είναι αρχικές κατανομές της Μ.α.

Με τη χρήση του παραπάνω ορισμού, όμως, είναι πρακτικά δύσκολο να δειχθεί ότι μία ΜΑ είναι ασθενώς εργοδική. Εξαιτίας αυτού, χρησιμοποιούμε τον συντελεστή εργοδικότητας που μας βοηθάει να διατυπώσουμε πιο εύχρηστα θεωρήματα για την εργοδικότητα.

Ορισμός A.5: Έστω \mathbf{P} ένας στοχαστικός πίνακας. συντελεστής εργοδικότητας του \mathbf{P} ορίζεται ως εξής:

$$\tau_1(\mathbf{P}) = \mathbf{1} - \min_{i,j} \sum_{k=1}^{\infty} \min\{P(i,k), P(j,k)\}$$

Θεώρημα A.3: Μία μη ομογενής MA είναι ασθενώς εργοδική, αν και μόνο αν υπάρχει μία γνησίως αύξουσα ακολουθία θετικών ακεραίων $\{k_i\}, i = 0,1,2, \dots$ τέτοιων, ώστε:

$$\sum_{i=0}^{\infty} [1 - \tau_1(\mathbf{P}^{(k_i, k_{i+1})})] = \infty.$$

Ορισμός A.6: Μία μη ομογενής MA είναι ισχυρά εργοδική, αν υπάρχει διάνυσμα \mathbf{q} με $q_i \geq 0$ και $\|\mathbf{q}\| = 1$ τέτοιο, ώστε για κάθε m :

$$\lim_{k \rightarrow \infty} \sup_{f(0)} \|f^{(m,k)} - \mathbf{q}\| = 0.$$

Θεώρημα A.4: Αν για κάθε m υπάρχει διάνυσμα $\boldsymbol{\pi}_m$, με $\boldsymbol{\pi}_m = \boldsymbol{\pi}_m \mathbf{P}, \|\boldsymbol{\pi}_m\| = 1$ και

$$\sum_{m=0}^{\infty} \|\boldsymbol{\pi}_m - \boldsymbol{\pi}_{m+1}\| < \infty$$

τότε, αν η MA είναι ασθενώς εργοδική, θα είναι και ισχυρά εργοδική.

Επιπλέον, αν

$$e^* = \lim_{m \rightarrow \infty} \boldsymbol{\pi}_m$$

τότε για κάθε m

$$\lim_{n \rightarrow \infty} \sup_{f(0)} \|f^{(m,n)} - e^*\| = 0.$$

Για περισσότερες πληροφορίες σχετικά με τις μη ομογενείς MA ο αναγνώστης παραπέμπεται στο βιβλίο των Isaacson, D.L. AND Madsen, R.W. (1976).

Σύνοψη-Συζήτηση Αποτελεσμάτων

Όπως έχει ήδη αναφερθεί, ο αλγόριθμος της προσομοιωμένης ανόπτησης είναι μία από τις πιο διαδεδομένες μεθόδους στοχαστικής βελτιστοποίησης. Στην παρούσα εργασία, αρχικά παρουσιάσαμε το θεωρητικό πλαίσιο του αλγορίθμου και, στη συνέχεια, τον εφαρμόσαμε σε δύο διαφορετικές περιπτώσεις: Η πρώτη εφαρμογή ήταν η μεγιστοποίηση της σ.π.π. της μείξης δύο κανονικών κατανομών και η δεύτερη η εκτίμηση των παραμέτρων ενός $GARCH(1,1)$ υποδείγματος. Για καθεμία από τις περιπτώσεις εκτελέσαμε 1000 επαναλήψεις του αλγορίθμου. Στις ακόλουθες παραγράφους συνοψίζονται τα συμπεράσματα που προέκυψαν από την παραπάνω διαδικασία.

Το πρώτο σημείο που είναι σημαντικό και θέλουμε να τονίσουμε είναι η ευελιξία της μεθόδου, καθώς αυτή δύναται να εφαρμοσθεί σε οποιοδήποτε υπολογιστικά δύσκολο πρόβλημα, χωρίς να επηρεάζεται από τυχόν περιορισμούς ή αναλυτικές ιδιότητες της αντικειμενικής συνάρτησης. Ωστόσο, προκειμένου να προχωρήσουμε στην εφαρμογή του αλγορίθμου, πρέπει να ορίσουμε τιμές για τις παραμέτρους του, οι οποίες, όπως είδαμε και στα παραδείγματα, επηρεάζουν σημαντικά την απόδοση. Παραδείγματος χάριν, αν η αρχική θερμοκρασία δεν είναι αρκετά υψηλή, ο αλγόριθμος ενδέχεται να τερματίσει, προτού προσεγγίσει τη βέλτιστη λύση.

Ένα άλλο ουσιαστικό πλεονέκτημα του αλγορίθμου είναι η ικανότητά του να μην εγκλωβίζεται σε περιοχές τοπικών ελαχίστων. Όπως έγινε φανερό και από τις εφαρμογές που παρουσιάστηκαν, στις περιπτώσεις που έχει γίνει κατάλληλη επιλογή στις τιμές των παραμέτρων του αλγορίθμου, το ποσοστό των λανθασμένων συγκλίσεων είναι αμελητέο. Ενώ παρατηρούμε ότι η ικανότητά του να προσεγγίζει το ολικό ακρότατο της αντικειμενικής συνάρτησης δεν επηρεάζεται από την επιλογή αρχικών σημείων, όπως συμβαίνει σε άλλες μεθόδους βελτιστοποίησης.

Ακόμη, θα πρέπει να αναφέρουμε ότι οι εκτιμήσεις στις οποίες βασίστηκε η ανάλυσή μας προέκυψαν κατόπιν 1000 επαναλήψεων του αλγορίθμου. Αυτό επελέγη διότι, λόγω της στοχαστικότητας της μεθόδου, υπήρχε μεγάλη πιθανότητα για τις εκτιμήσεις που προέρχονται από μεμονωμένες εκτελέσεις του αλγορίθμου, ακόμη και για α, T με υψηλές τιμές, να απέχουν σημαντικά από την πραγματική λύση του προβλήματος.

Μία ακόμη σημαντική παρατήρηση σχετικά με τον αλγόριθμο είναι ότι η ακρίβεια της λύσης που επιστρέφει (ο αλγόριθμος) εξαρτάται από το πλήθος των βημάτων του. Δηλαδή, όσο αυξάνεται η ακρίβεια της λύσης μας αυξάνεται και το υπολογιστικό κόστος. Γι' αυτόν τον λόγο, το να επιλέξουμε για τις παραμέτρους α, T κάποιες τιμές που μεγιστοποιούν την ακρίβεια δεν είναι η καλύτερη στρατηγική. Αυτό που πρέπει να κάνουμε είναι να επιλέξουμε τέτοιες τιμές, ούτως ώστε να υπάρχει ισορροπία μεταξύ ακρίβειας και υπολογιστικού κόστους. Μία άλλη τακτική που θα μπορούσαμε να ακολουθήσουμε, δεδομένης της ικανότητας του αλγορίθμου να έλκεται στην περιοχή του ολικού μεγίστου της αντικειμενικής συνάρτησης, είναι η εξής: Αρχικά εκτελούμε τον αλγόριθμο της προσομοιωμένης ανόπτησης για τιμές των α, T χωρίς πολύ μεγάλο υπολογιστικό κόστος και, στη συνέχεια, αφού εντοπίσουμε την περιοχή του ολικού μεγίστου, εφαρμόζουμε σε έναν πιο περιορισμένο χώρο λύσεων μία μέθοδο μεγαλύτερης ακρίβειας. Τέτοιοι συνδυασμοί μεθόδων είναι γνωστοί στη βιβλιογραφία με την ονομασία «υβριδικές μέθοδοι».

Για όλους τους παραπάνω λόγους, ο αλγόριθμος της προσομοιωμένης ανόπτησης αλλά και, πολύ περισσότερο πια, νέες τεχνικές που βασίζονται σε αυτόν, όπως οι υβριδικές μέθοδοι, χρησιμοποιούνται ευρέως για την επίλυση υπολογιστικά δύσκολων προβλημάτων βελτιστοποίησης.

Βιβλιογραφία

1. A. Arratia (2014) *Computational Finance: An Introductory Course with R*, Atlantis Press.
2. P. Brandimarte (2014) *Handbook in Monte Carlo Simulation*, Wiley, Hoboken, NJ.
3. Bollerslev, T. (1986), Generalized autoregressive conditional heteroskedasticity, *Journal of Econometrics*, **31**:307-27.
4. Cerny, V. A Thermodynamical Approach to the Travelling Salesman Problem: An Efficient Simulation Algorithm. *J. Optim. Theory Appl.* **45** : 41 – 51.
5. Engle, R.F., (1982), Autoregressive conditional heteroskedasticity, with estimates of the variance of the United Kingdom inflation. *Econometrica*, **50**: 987-1007.
6. Geman, S. & Geman, D. (1984) Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images, *IEEE Transactions on Pattern analysis and Machine Intelligence* **6**, 721 – 741.
7. M. Gendreau, J-Y Potvin (2010) *Handbook of Metaheuristics*, Springer, New York
8. O. Haggstrom (2002) *Finite Markov Chains and Algorithmic Applications*, Cambridge University Press.
9. Hajek, B. (1988). Cooling schedules for optimal annealing. *Mathematics of Operations Research*, **13**:375-395.
10. Isaacson, D.L. AND Madsen, R.W. (1976) *Markov Chains: Theory and Applications*. Wiley, New York.
11. Kirkpatrick, S., Gelatt, Jr., C.D., Vecchi, M.P.(1983): Optimization by Simulated Annealing. *Science*, **220**, 671 – 680.
12. D.P. Kroese, R.Y. Rubinstein (2007) *Simulation and the Monte Carlo Method* (2nd ed.) Wiley, Hoboken, NJ.
13. Jun S. Liu (2004) *Monte Carlo Strategies in Scientific Computing*, Springer, New York.
14. Metropolis N., Rosenbluth A.W., Rosenbluth M.N., Teller A.H. and Teller E., (1953), Equations of State Calculations by Fast Computing Machines, *Journal of Chemical Physics*, vol. **21**, 1087 – 1092.
15. Mitra, D., Romeo, F., Sangiovanni-Vincentelli, A.L.(1986): Convergence and finite time behavior of simulated annealing. *Adv. Appl. Probab.* **18**, 747 – 771.
16. C.P. Robert and G. Casella (2004), *Monte Carlo Statistical Methods*, Springer Verlag, New York. (2nd ed.)

17. C.P. Robert and G. Casella (2010), *Introducing Monte Carlo Methods with R*, Springer Verlag, New York.
18. S. Ross (1997) *Simulation* (2nd ed.), Academic Press, San Diego, CA.
19. Van Laarhoven, P.J.M., Aarts, E.H.L. (1987) *Simulated Annealing : Theory and applications*, Springer, New York.
20. Ruey S. Tsay (2012) *An Introduction to Analysis of Financial Data With R* ,Wiley, Hoboken NJ.
21. Μ. Μπούτσικας, (2004) *Μέθοδοι Προσομοίωσης και Στατιστικές Υπολογιστικές Τεχνικές*, Πανεπιστημιακές σημειώσεις, Πανεπιστήμιο Πειραιώς.

