



Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης
Πολυτεχνική Σχολή
Τμήμα Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών

Προσδιορισμός Αναλυτικής Έκφρασης συνεχούς
συνάρτησης f με χρήση Γκαουσιανών συναρτήσεων

Αμαραντίδου Ευθυμία
7ο Εξάμηνο
Α.Ε.Μ.: 9762

Τεχνικές Βελτιστοποίησης
Project

14 Φεβρουαρίου 2022

Περιεχόμενα

1	Εισαγωγή	2
1.1	Περιγραφή του προβλήματος	2
1.2	Γενετικός Αλγόριθμος	2
1.2.1	Μέθοδοι εξέλιξης του πληθυσμού	3
1.2.2	Επιλογή παραμέτρων γενετικού αλγορίθμου	4
2	Πειραματικά Αποτελέσματα	6
2.1	Βέλτιστη προσέγγιση της αναλυτικής μορφής της συνάρτησης f	7
2.2	Δοκιμές για μελέτη επιρροής παραμέτρων	7
2.2.1	Παράμετροι υπό μελέτη	9

Κεφάλαιο 1

Εισαγωγή

Στην εργασία αυτή θα προταθεί μια χαμηλής πολυπλοκότητας αναλυτική έκφραση συνεχούς συνάρτησης f , με χρήση **γενετικού αλγορίθμου**. Παράλληλα, θα γίνει και αξιολόγηση της συνάρτησης αυτής, καθώς και παρουσίαση των διαφόρων δοκιμών με στόχο να προκύψει η βέλτιστη προσέγγιση. Ο γενετικός αλγόριθμος που χρησιμοποιήθηκε έχει υλοποιηθεί με τη βοήθεια του `Matlab` και θα παρουσιαστεί στη συνέχεια.

1.1 Περιγραφή του προβλήματος

Θεωρούμε το στατικό σύστημα δύο εισόδων u_1, u_2 και μιας εξόδου y , που περιγράφεται από τη σχέση 1.1.

$$y = f(u_1, u_2) \quad (1.1)$$

Η αναλυτική έκφραση της f είναι άγνωστη, αλλά **συνεχής** συνάρτηση των u_1, u_2 και θα προσεγγιστεί με τη βοήθεια μετρήσεων εισόδου-εξόδου καθώς και κατάλληλα σχεδιασμένου **γραμμικού συνδυασμού** -το πολύ 15- **γκαουσιανών συναρτήσεων** της μορφής:

$$G(u_1, u_2) = Ae^{-\left(\frac{(u_1 - c_1)^2}{2\sigma_1^2} + \frac{(u_2 - c_2)^2}{2\sigma_2^2}\right)} \quad (1.2)$$

Για την παραγωγή δεδομένων εισόδου-εξόδου, καθώς και την αξιολόγηση των αποτελεσμάτων θα χρησιμοποιηθεί η **συνάρτηση** 1.3.

$$f(u_1, u_2) = \sin(u_1 + u_2)\sin(u_2^2), \quad u_1 \in [-1, 2] \quad u_2 \in [-2, 1] \quad (1.3)$$

1.2 Γενετικός Αλγόριθμος

Οι γενετικοί αλγόριθμοι, αποτελούν αλγορίθμους βελτιστοποίησης προβλημάτων εύρεσης ακρότατων και βασίζονται στις αρχές της Βιολογικής Εξέλιξης. Κύρια στοιχεία των αλγορίθμων αυτών αποτελούν ο **πληθυσμός**, τα **χρωμοσώματα** και τα **γονίδια**.

Συγκεκριμένα, στην περίπτωση της εργασίας αυτής, οι παράμετροι $A, c_1, c_2, \sigma_1, \sigma_2$ της

1.2 αποτελούν τα **γονίδια**, ενώ το χρωμόσωμα περιλαμβάνει **κάθε ένα από τα γονίδια** αυτά **για όλες τις γκαουσσιανές** συναρτήσεις που χρησιμοποιούνται.

Στο σημείο αυτό χρειάζεται να αναφέρουμε πως ένα από τα κύρια χαρακτηριστικά των γενετικών αλγορίθμων αποτελεί η αναπαράσταση των παραμέτρων με δυαδικούς αριθμούς. Η επιλογή του αριθμού των ψηφίων που χρησιμοποιούνται για την δυαδική αναπαράσταση, επηρεάζει και την ακρίβεια του αλγορίθμου. Για λόγους μείωσης της υπολογιστικής πολυπλοκότητας και του χρόνου εκτέλεσης του αλγορίθμου, στην συγκεκριμένη υλοποίηση οι τιμές των παραμέτρων δεν κωδικοποιούνται αλλά αναπαριστώνται ως πραγματικοί αριθμοί.

Η ευρύτερη λογική των γενετικών αλγορίθμων μπορεί να χωριστεί στα παρακάτω στάδια και περιγράφεται και γραφικά στο Σχήμα 1.1:

1. Δημιουργία αρχικού **τυχαίου** πληθυσμού λαμβάνοντας υπόψιν τους υπάρχοντες περιορισμούς κάθε γονιδίου.
2. Υπολογισμός της ικανότητας κάθε χρωμοσώματος-λύσης του εκάτοστε πληθυσμού.
3. Εξέλιξη του πληθυσμού και δημιουργία της επόμενης γενιάς μέσω 4 διαδικασιών:
 - Επιλογή βέλτιστων χρωμοσωμάτων
 - Επιλογή τυχαίων χρωμοσωμάτων
 - Διασταύρωση χρωμοσωμάτων
 - Μετάλλαξη χρωμοσωμάτων
4. Αξιολόγηση νέου πληθυσμού
5. Επανάληψη της διαδικασίας ώσπου να προκύψει μια ικανοποιητική λύση¹ ή να ολοκληρωθεί ο μέγιστος αριθμός γενεών που έχουμε ορίσει.

1.2.1 Μέθοδοι εξέλιξης του πληθυσμού

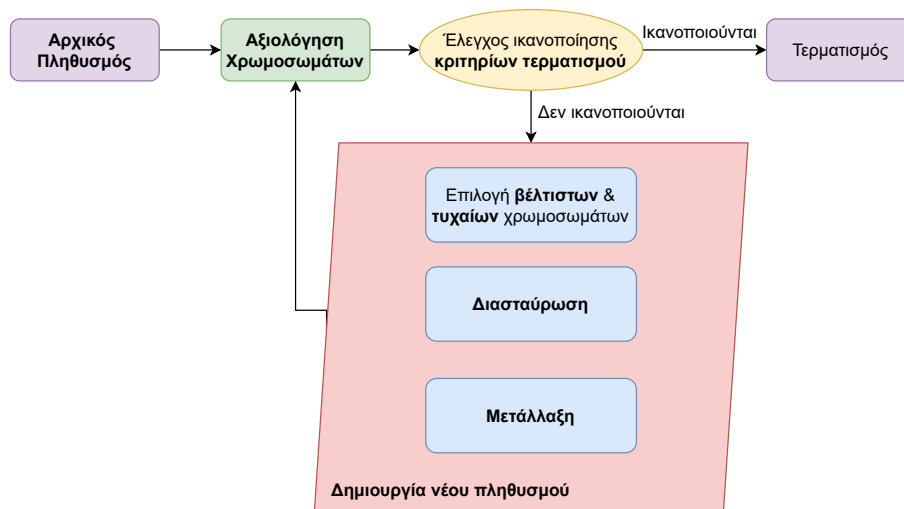
Επιλογή βέλτιστων χρωμοσωμάτων Η επιλογή των βέλτιστων χρωμοσωμάτων γίνεται με βάση τα αποτελέσματα που προκύπτουν από την αντικειμενική συνάρτηση. Η αντικειμενική συνάρτηση είναι αυτή που αξιολογεί πόσο ικανοποιητική είναι μια λύση (ένα χρωμόσωμα). Έτσι, ένα συγκεκριμένο ποσοστό των καλύτερων λύσεων επιλέγεται και μεταφέρεται και στην επόμενη γενιά.

Επιλογή τυχαίων χρωμοσωμάτων Ένα ποσοστό από τα υπόλοιπα (εκτός των βέλτιστων) χρωμοσώματα επιλέγεται τυχαία να μεταδωθεί και στην επόμενη γενιά.

Διασταύρωση χρωμοσωμάτων Η διασταύρωση αποτελεί τον κυριότερο τελεστή στους γενετικούς αλγορίθμους. Κατά τη διαδικασία αυτή, επιλέγονται τυχαία 2 γονίδια του πληθυσμού και διασταυρώνονται έτσι ώστε να προκύψουν αντίστοιχοι απόγονοι. Αναλόγως με την μέθοδο διασταύρωσης, ο αριθμός των απογόνων που θα προκύψουν είναι 1 ή 2. Οι μέθοδοι διασταύρωσης παρουσιάζονται στο Σχήμα 1.2 και είναι οι εξής:

1. Κάθε γονίδιο χωρίζεται σε δύο τμήματα και διασταυρώνεται με το άλλο με αποτέλεσμα να προκύψουν **2 απόγονοι**.

¹Ως ικανοποιητική λύση ορίζεται αυτή η οποία οδηγεί σε σφάλμα μικρότερο από το επιθυμητό.



Σχήμα 1.1: Γενετικός Αλγόριθμος

2. Υπολογίζεται το ημίθροισμα των εκάστοτε γονιδίων και δημιουργείται **ένας α-πόγονος**.²

Μετάλλαξη χρωμοσωμάτων Κατά τη μετάλλαξη των χρωμοσωμάτων επιλέγεται ένα γονίδιο του οποίου κάποιο ψηφίο αντιστρέφεται. Στην περίπτωση του δυαδικού συστήματος, η επιλογή του ψηφίου που θα αντιστραφεί είναι λίγο πιο περίπλοκη, καθώς το μέγεθος της μεταβολής, στην τιμή του γονιδίου, που προκαλείται εξαρτάται από τη σημαντικότητα του ψηφίου. Ωστόσο, στην εργασία αυτή, η μετάλλαξη πραγματοποιείται με αλλαγή της τιμής του επιλεγμένου γονιδίου.³ Η μετάλλαξη, αν και πραγματοποιείται με σχετικά μικρή πιθανότητα, είναι πολύ σημαντική για τη σύγκλιση του γενετικού αλγορίθμου, καθώς αποτρέπει τον εγκλοβισμό σε μη βέλτιστες λύσεις.

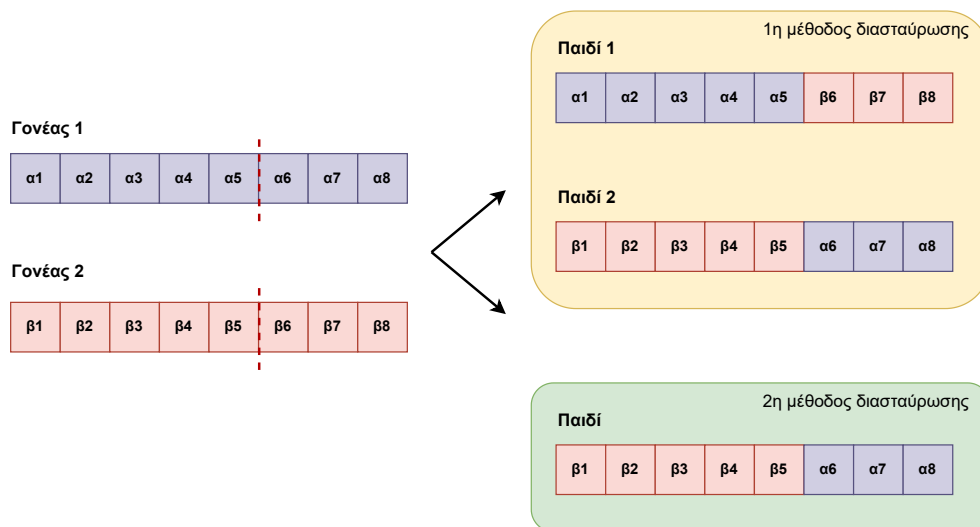
1.2.2 Επιλογή παραμέτρων γενετικού αλγορίθμου

Η επιλογή του μεγέθους του πληθυσμού, καθώς και των τιμών των αντίστοιχων ποσοστών επιλογής χρωμοσωμάτων έχει γίνει μετά από πολλές δοκιμές, κάποιες από τις οποίες θα παρουσιαστούν και στην ανάλυση που ακολουθεί στην Ενότητα 2. Οι τελικές τιμές των παραμέτρων που επιλέχθηκαν είναι:

Αριθμός γκαουσιανών συναρτήσεων: 15

²Η συγκεκριμένη μέθοδος διασταύρωσης εφαρμόζεται μόνο στην περίπτωση όπου τα χρωμοσώματα αναπαριστώνται με πραγματικές τιμές όπως στην εργασία αυτή. Εάν χρησιμοποιηθεί δυαδική αναπαράσταση τότε δεν η μέθοδος αυτή δε μπορεί να αξιοποιηθεί.

³Στην προσέγγιση αυτή, υπάρχει ένα μειονέκτημα στη χρήση του δεκαδικού συστήματος, καθώς δεν μπορεί να αναπαραστηθεί απόλυτα η μετάλλαξη του δυαδικού συστήματος. Η νέα τιμή γονιδίου επιλέγεται τυχαία αλλά **εντός** των επιτρεπτών ορίων σύμφωνα με τους περιορισμούς, κάτι το οποίο δεν λαμβάνεται υπόψη στην αναστροφή ενός δυαδικού ψηφίου.



Σχήμα 1.2: Μέθοδοι διασταύρωσης

Μέγιστος αριθμός γενεών: 10.000

Μέγεθος πληθυσμού: 100

Περιορισμοί γονιδίων:

$A : [f_{min}, f_{max}] = [-0.99, 0.84]$

$c_i : [-3, 3]$

$\sigma_i : [0.1, 1.3]$

Πληθυσμός νέας γενιάς:

Βέλτιστα χρωμοσώματα προηγούμενης γενιάς: 20%

Τυχαία χρωμοσώματα προηγούμενης γενιάς: 10%

Διασταύρωση χρωμοσωμάτων: 70%

Πιθανότητα μετάλλαξης: 0.1%

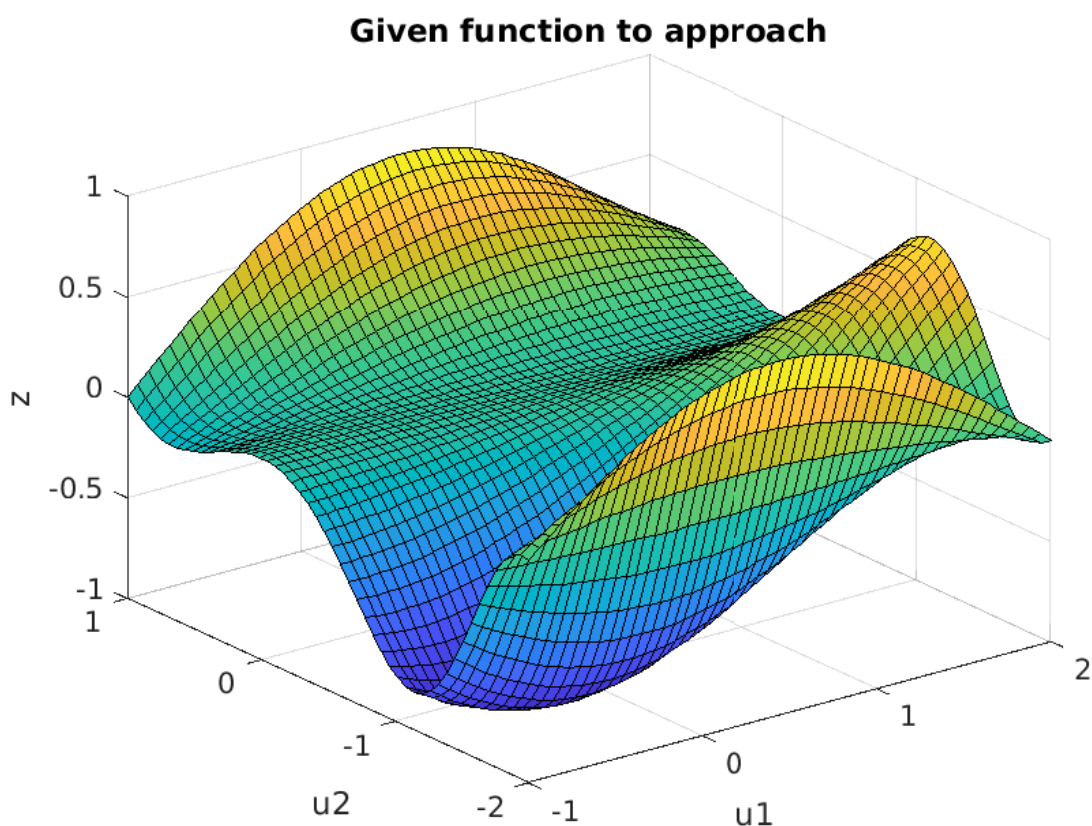
Ως αντικειμενική συνάρτηση για την αξιολόγηση των γονιδίων έχουν χρησιμοποιηθεί το **απόλυτο σφάλμα** $e_{abs} = \sum |f - \bar{f}|$ και το **μέσο τετραγωνικό σφάλμα** $e_{ms} = \sum (f - \bar{f})^2$, όπου f η 1.3 και \bar{f} η προσέγγιση της αναλυτικής συνάρτησης. Για να είναι όσο πιο αντικειμενική η αξιολόγηση, το σφάλμα υπολογίζεται χρησιμοποιώντας ένα σύνολο 400 τιμών (u_1, u_2) και με χρήση της συνάρτησης 1.3. Το βέλτιστο αποτέλεσμα προέκυψε με **χρήση του μέσου τετραγωνικού σφάλματος**.

Κεφάλαιο 2

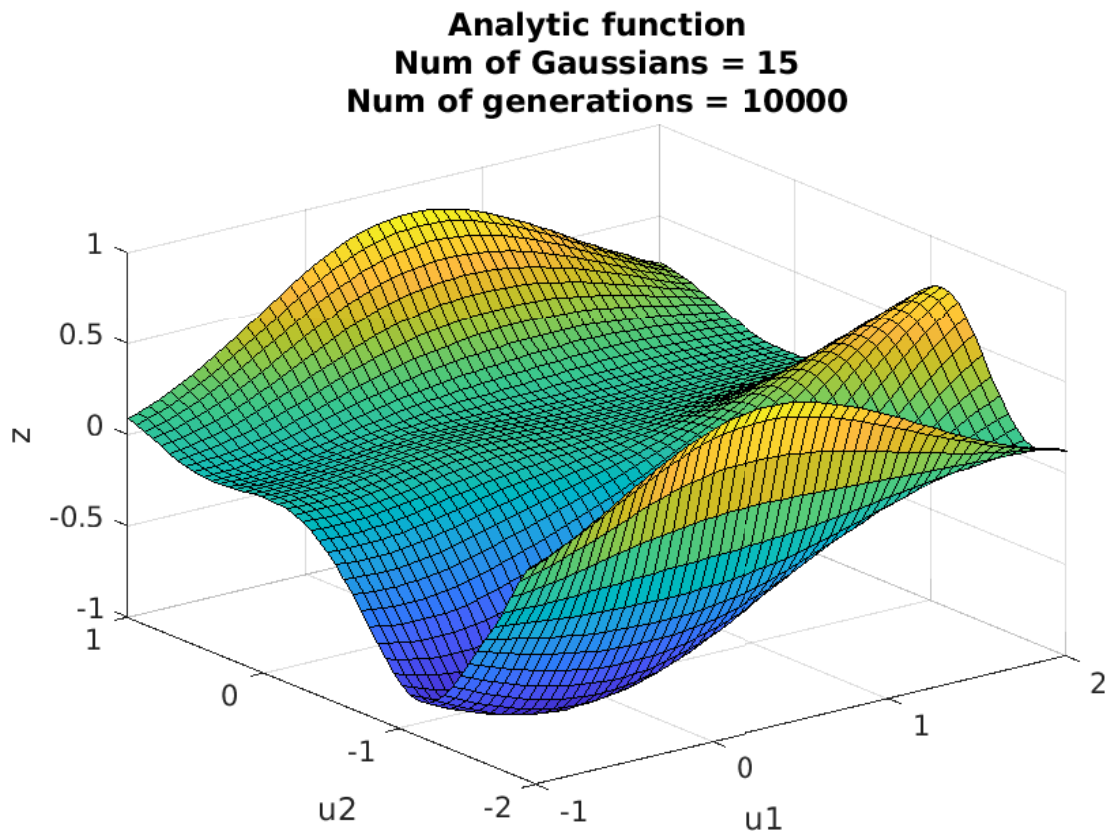
Πειραματικά Αποτελέσματα

Στην ενότητα αυτή θα παρουσιαστεί αρχικά το βέλτιστο αποτέλεσμα που προέκυψε μετά από πειραματική εφαρμογή του γενετικού αλγορίθμου. Στη συνέχεια, θα παρουσιαστούν και οι διάφορες δοκιμές που έγιναν, ώστε να εξάγουμε συμπεράσματα για την επιρροή κάποιων παραμέτρων στο αποτέλεσμα του γενετικού αλγορίθμου και να οδηγηθούμε όσο το δυνατόν πιο κοντά στη βέλτιστη λύση, δηλαδή τη βέλτιστη προσέγγιση της αναλυτικής μορφής της f .

Προκειμένου να μπορέσουμε και γραφικά να αξιολογήσουμε το αποτέλεσμα της συνάρτησης, η γραφική αναπαράσταση της 1.3 βρίσκεται στο Σχήμα 2.1.



Σχήμα 2.1: Συνάρτηση που προσεγγίζεται



Σχήμα 2.2: Προσέγγιση συνάρτησης f

2.1 Βέλτιστη προσέγγιση της αναλυτικής μορφής της συνάρτησης f

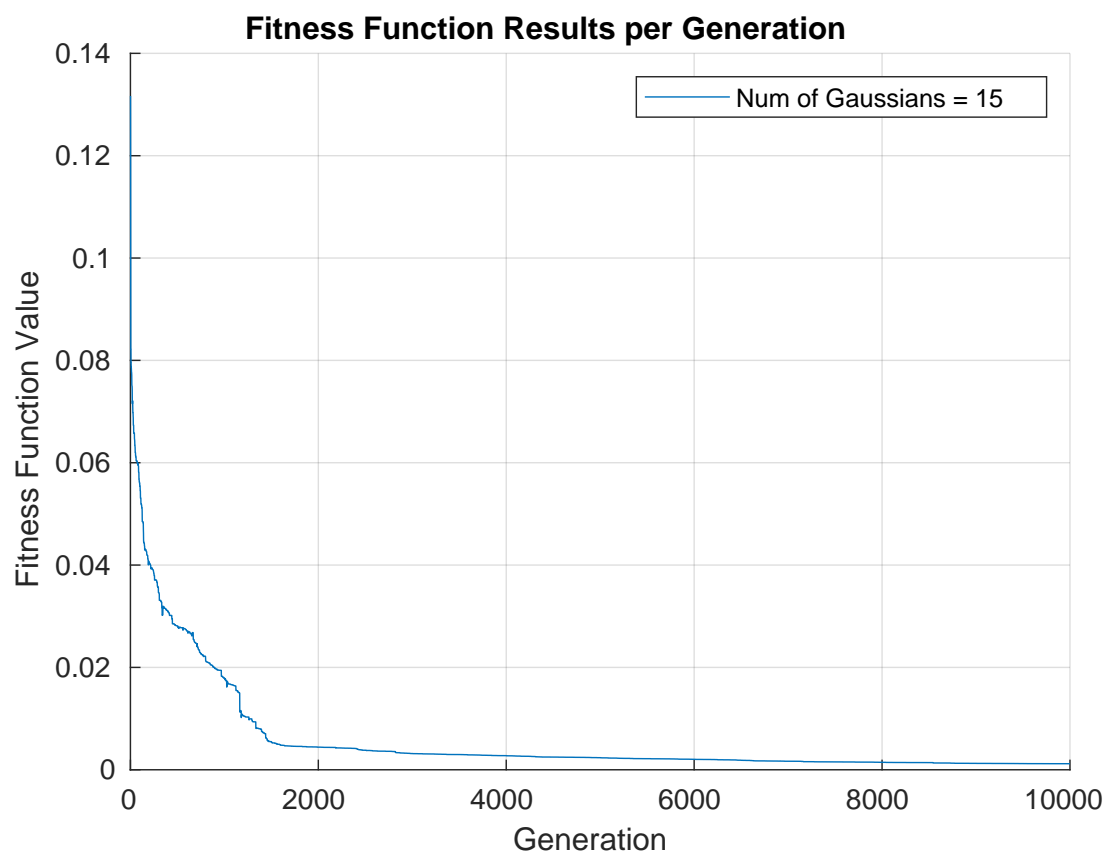
Η βέλτιστη προσέγγιση προέκυψε με χρήση των παραμέτρων που αναφέρονται στην Υποενότητα 1.2.2. Η γραφική αναπαράσταση της αναλυτικής έκφρασης της f και το σφάλμα που προέκυψε σε κάθε γενιά βρίσκονται στα Σχήματα 2.2 και 2.3 αντίστοιχα.

Ουσιαστικά, αφού βρέθηκαν οι βέλτιστες τιμές των A , c_i , σ_i για κάθε μία από τις 15 Γκαουσιανές συναρτήσεις υπολογίστηκε η αναλυτική έκφραση ως το άθροισμα αυτών. Το σφάλμα μπορούμε να παρατηρήσουμε από το Σχήμα 2.3 πως μειώθηκε εκθετικά ανά τις γενιές, ενώ τελικό σφάλμα προσέγγισης ήταν 0.0012, οπότε μπορούμε να πούμε πως οδηγηθήκαμε σε μια καλή προσέγγιση.

2.2 Δοκιμές για μελέτη επιρροής παραμέτρων

Παρακάτω ακολουθούν κάποιες από τις δοκιμές που πραγματοποιήθηκαν προκειμένου να επιλέξουμε τις βέλτιστες τιμές κάποιων παραμέτρων. Οι δοκιμές που έγιναν αφορούν:

- το πλήθος των Γκαουσιανών συναρτήσεων
- την πιθανότητα μετάλλαξης γονιδίων



Σχήμα 2.3: Σφάλμα ανά γενεά της τελικής προσέγγισης

- το ποσοστό των βέλτιστων γονιδίων που μεταδίδονται από μία γενιά στην επόμενη
- το μέγεθος του πληθυσμού

Προκειμένου να εξάγουμε συμπεράσματα για τις παραμέτρους αυτές, ο αλγόριθμος έχει επαναληφθεί **4 φορές** για κάθε μια από τις περιπτώσεις, ώστε να μην επηρεαστούμε μόνο από μια τυχαία επανάληψη. Έτσι υπολογίζεται για κάθε περίπτωση ο μέσος όρος του σφάλματος και τα αποτελέσματα παρουσιάζονται στα διαγράμματα που ακολουθούν. Προφανώς, θεωρούμε πως οι βέλτιστες τιμές των παραμέτρων είναι αυτές για τις οποίες προκύπτει το ελάχιστο σφάλμα.

Επιπλέον, προκειμένου να ελέγξουμε εάν τα συμπεράσματά μας επηρεάζονται από το μέγιστο αριθμό γενεών, η παραπάνω διαδικασία έχει ολοκληρωθεί για μέγιστο αριθμό γενεών ίσο με 5.000 και 10.000. Ωστόσο, δεν παρατηρείται κάποια σημαντική διαφορά που να επηρεάζει τη σύγκριση των περιπτώσεων.

Για τις παρακάτω προσομοιώσεις, οι τιμές των παραμέτρων που χρησιμοποιήθηκαν, **τροποποιώντας κάθε φορά την παράμετρο που εξετάζεται**, είναι οι εξής:

Αριθμός γκαουσιανών συναρτήσεων: 15

Μέγιστος αριθμός γενεών: [5.000, 10.000]

Μέγεθος πληθυσμού: 30

Περιορισμοί γονιδίων:

$A : [f_{min}, f_{max}] = [-0.99, 0.84]$

$c_i : [-3, 3]$

$\sigma_i : [0.1, 1.3]$

Πληθυσμός νέας γενιάς:

Βέλτιστα χρωμοσώματα προηγούμενης γενιάς: 30%

Τυχαία χρωμοσώματα προηγούμενης γενιάς: 10%

Διασταύρωση χρωμοσωμάτων: 70%

Πιθανότητα μετάλλαξης: 0.1%

2.2.1 Παράμετροι υπό μελέτη

Πλήθος γκαουσιανών συναρτήσεων

Όπως έχουμε ήδη αναφέρει, η αναλυτική έκφραση της 1.1 μπορεί να προσεγγιστεί ως γραμμικός συνδυασμός συναρτήσεων της μορφής 1.2. Έχοντας στη διάθεσή μας το πολύ 15 τέτοιες συναρτήσεις, στόχος είναι να βρούμε το **βέλτιστο πλήθος** αυτών ώστε να καταλήξουμε σε μια προσέγγιση με ικανοποιητικά μικρό σφάλμα. Έτσι, ολοκληρώθηκαν δοκιμές για χρήση **5, 7, 10, 15 γκαουσιανών συναρτήσεων**, και όπως παρατηρούμε από τα Διαγράμματα 2.4 και 2.5, το ελάχιστο σφάλμα προκύπτει με χρήση **15 γκαουσιανών συναρτήσεων**.

Χρησιμοποιώντας περισσότερες συναρτήσεις, είναι ευκολότερο να προσεγγιστεί η συνάρτηση με μεγαλύτερη ακρίβεια, καθώς προσφέρεται μεγαλύτερη ευελιξία. Κάτι που αξίζει να σημειώσουμε, είναι ότι για 5000 γενιές, όπως φαίνεται και στο Σχήμα 2.4 η χρήση 10 γκαουσιανών οδηγεί σε ελάχιστα μικρότερο σφάλμα, ωστόσο αυτό αλλάζει στι 10.000 επαναλήψεις. Επιπλέον, ενώ η απόκλιση του τελικού σφάλματος που επιτυγχάνεται με χρήση 5 και 7 γκαουσιανών είναι μηδαμινή στις 5.000 επαναλήψεις,

μπορούμε να παρατηρήσουμε ότι αυξάνεται σημαντικά στις 10.000.

Ποσοστό κληρονομησης βέλτιστων γονιδίων

Στον γενετικό αλγόριθμο, κάθε χρωμόσωμα του πληθυσμού έχει πιθανότητα επιβίωσης ανάλογη της ικανότητάς του. Έτσι, αξιολογώντας το κάθε χρωμόσωμα, επιλέγεται ένα ποσοστό των χρωμοσωμάτων με την καλύτερη ικανότητα να μεταβιβαστεί στην επόμενη γενιά. Έγιναν δοκιμές για ποσοστό ίσο με **20%, 30% και 50%** και τα αποτελέσματα παρουσιάζονται στα Σχήματα 2.6 και 2.7.

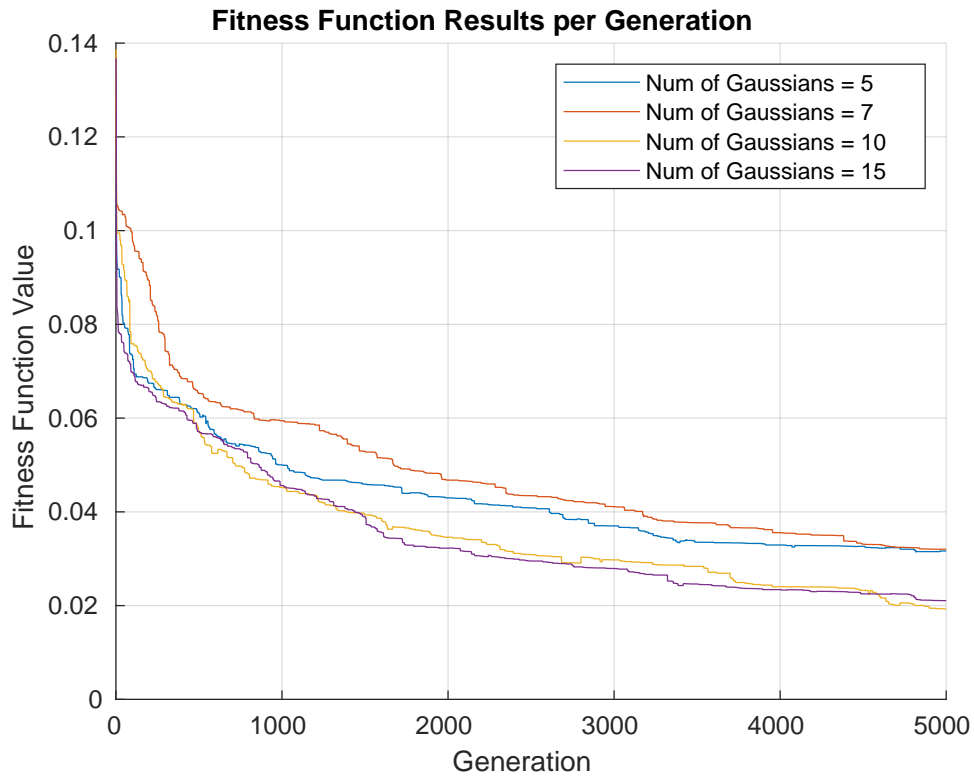
Παρόλο που στις 5.000 γενιές δεν φαίνεται να επηρεάζει σημαντικά το ποσοστό αυτό, στις 10.000 επαναλήψεις μπορεί να εξαχθεί το συμπέρασμα ότι το **βέλτιστο ποσοστό είναι 20%**.

Πιθανότητα μετάλλαξης

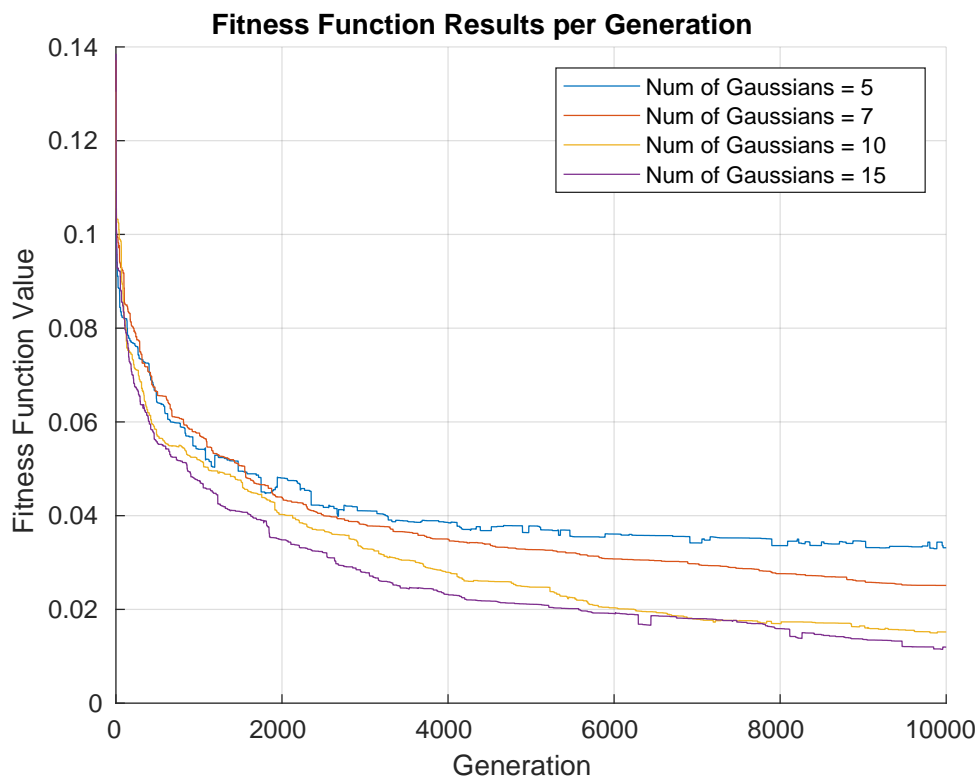
Η μετάλλαξη αποτελεί έναν από τους σημαντικότερους τελεστές των γενετικών αλγορίθμων, καθώς εμποδίζουν τον εγκλοβισμό σε συγκεκριμένες περιοχές του χώρου αναζήτησης. Στα διαγράμματα που ακολουθούν έγιναν δοκιμές για τιμές της πιθανότητας μετάλλαξης **0.01, 0.1, 0.2**. Παρατηρείται, ότι **όσο μεγαλύτερη η πιθανότητα, τόσο μικρότερο και το τελικό σφάλμα προσέγγισης**. Ωστόσο, καθώς η τιμή 0.2 είναι μια στην πραγματικότητα σχετικά μεγάλη τιμή για πιθανότητα μετάλλαξης, επιλέχθηκε στον τελικό αλγόριθμο να χρησιμοποιηθεί η τιμή 0.1.

Μέγεθος πληθυσμού

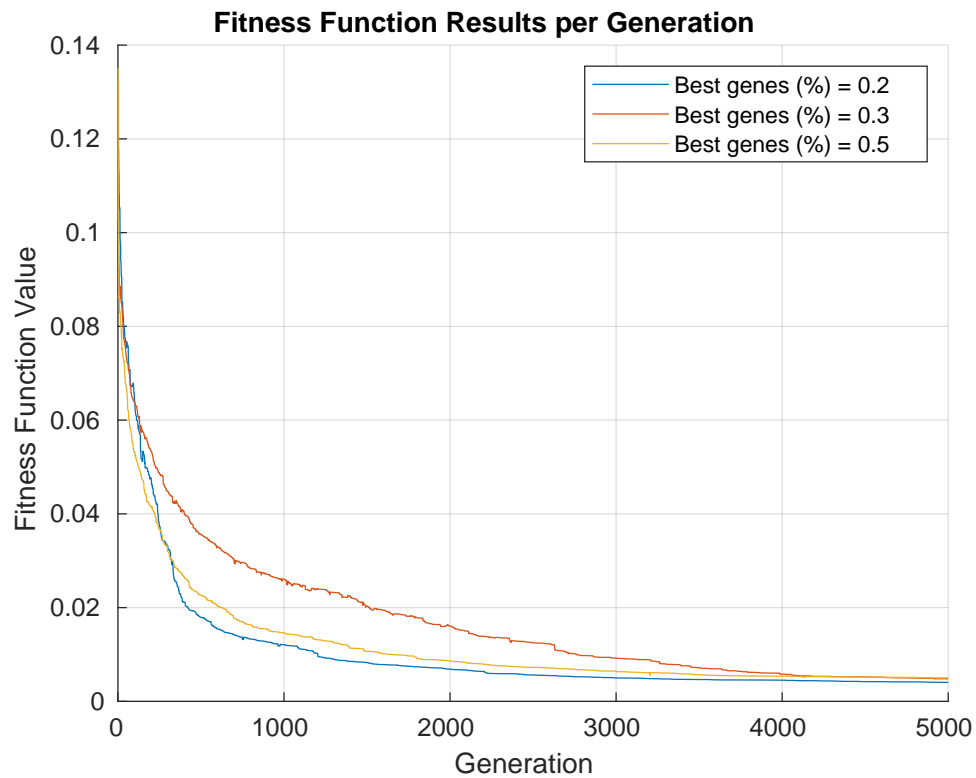
Το μέγεθος του πληθυσμού κάθε γενιάς είναι κι αυτό μια από τις βασικές παραμέτρους του γενετικού αλγορίθμου. Ωστόσο, από τις δοκιμές που έγιναν εξάγεται το συμπέρασμα πως η επιλογή του δεν είναι καθοριστική για το τελικό σφάλμα που προκύπτει **στο συγκεκριμένο πρόβλημα**. Μάλιστα, στις 5.000 επαναλήψεις παρατηρούμε ότι η διαφορά είναι ελάχιστη μεταξύ των διαφορετικών τιμών για το μέγεθος του πληθυσμού. Στις 10.000 επαναλήψεις, φαίνεται το ελάχιστο σφάλμα να προκύπτει για πληθυσμό ίσο με 100 χρωμοσώματα. Εφόσον όμως με την αύξηση του πληθυσμού, αυξάνονται οι ανάγκες σε μνήμη και η απαιτούμενη επεξεργαστική ισχύς, μπορούμε να μειώσουμε ακόμη και στο μισό τον πληθυσμό χωρίς να χάσουμε σημαντικά σε ακρίβεια.



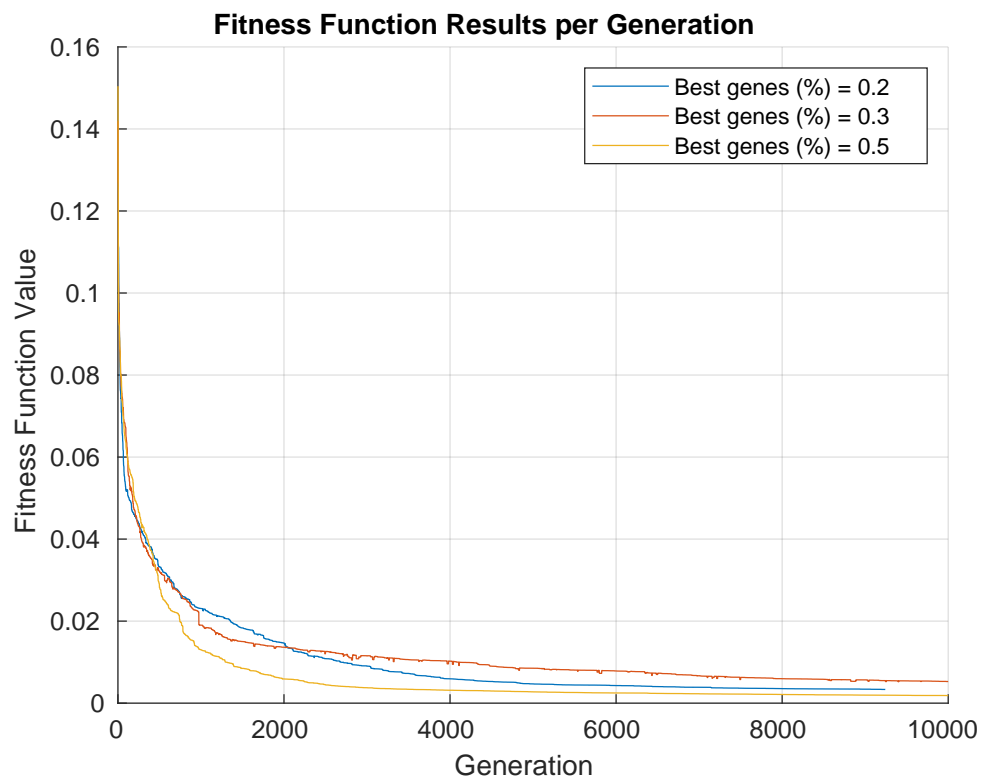
Σχήμα 2.4: Δοκιμές για διαφορετικό αριθμό Γκαουσιανών συναρτήσεων
Μέγιστος αριθμός γενεών = 5.000



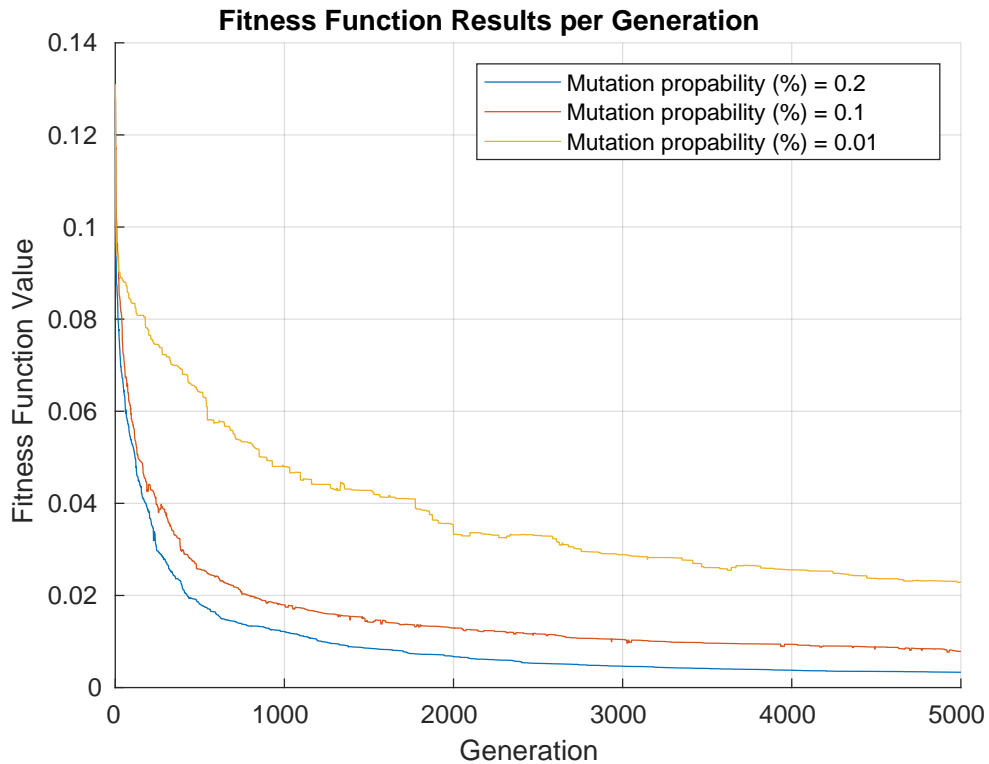
Σχήμα 2.5: Δοκιμές για διαφορετικό αριθμό Γκαουσιανών συναρτήσεων
Μέγιστος αριθμός γενεών = 10.000



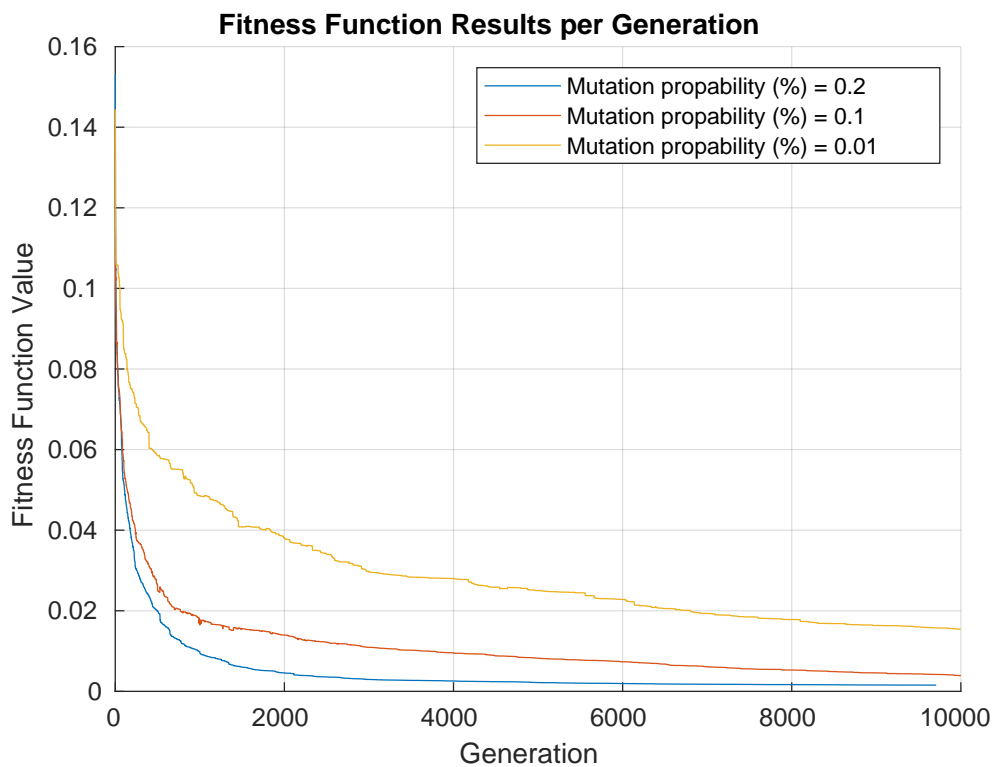
Σχήμα 2.6: Δοκιμές για διαφορετικό ποσοστό κληρονόμησης βέλτιστων γονιδίων
Μέγιστος αριθμός γενεών = 5.000



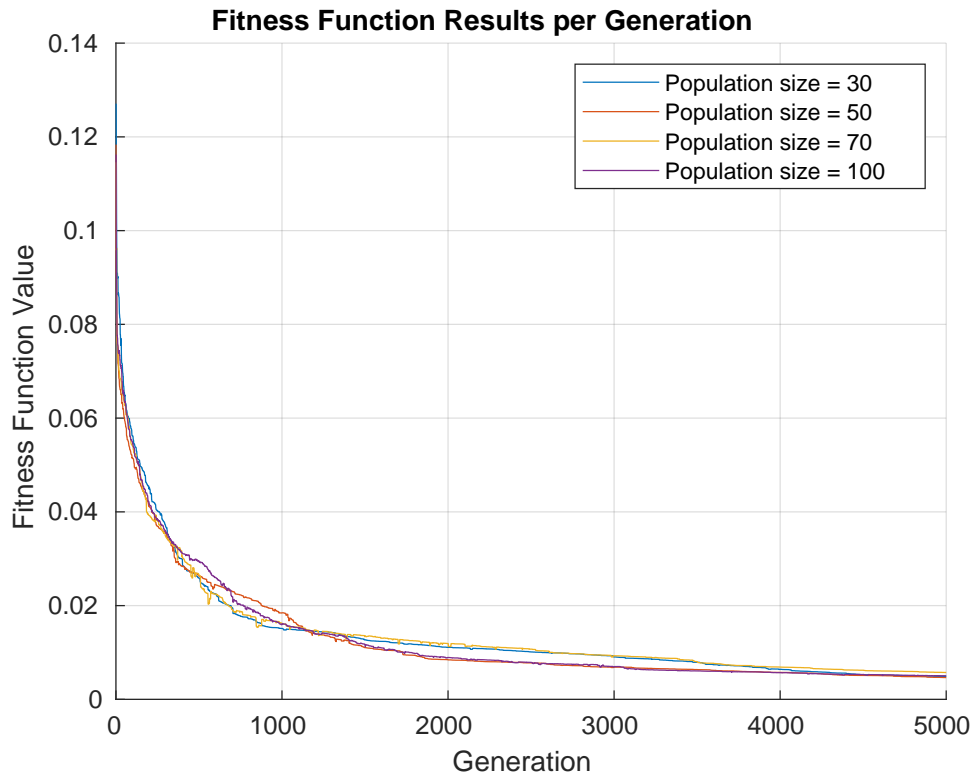
Σχήμα 2.7: Δοκιμές για διαφορετικό ποσοστό κληρονόμησης βέλτιστων γονιδίων
Μέγιστος αριθμός γενεών = 10.000



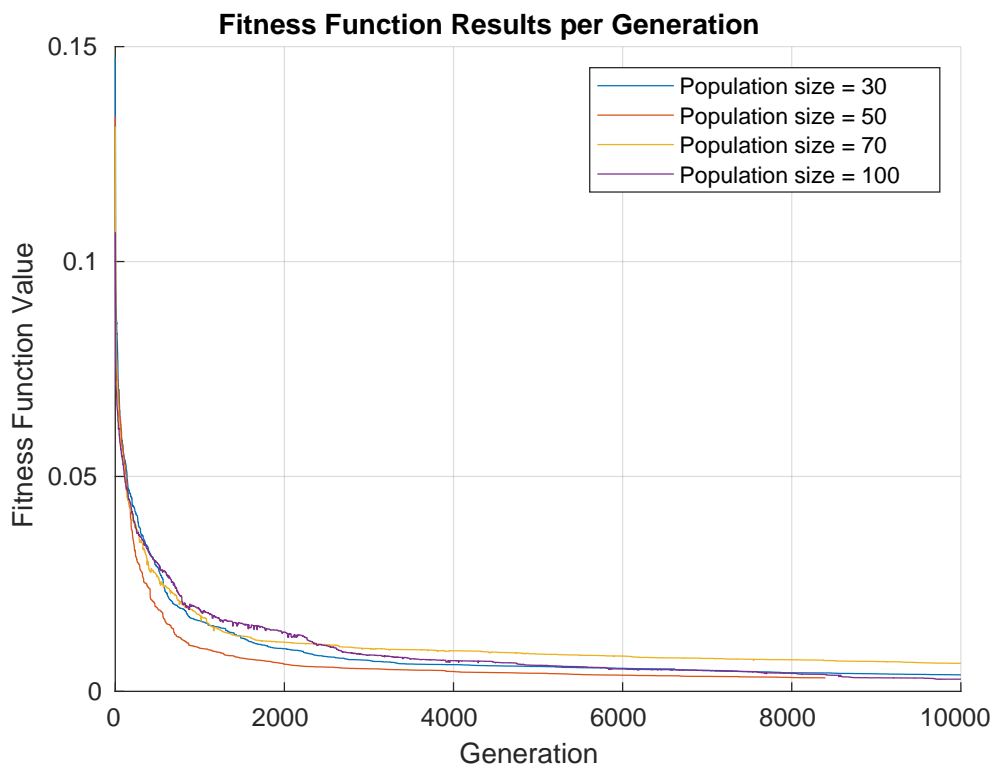
Σχήμα 2.8: Δοκιμές για διαφορετική πιθανότητα μετάλλαξης
Μέγιστος αριθμός γενεών = 5.000



Σχήμα 2.9: Δοκιμές για διαφορετική πιθανότητα μετάλλαξης
Μέγιστος αριθμός γενεών = 10.000



Σχήμα 2.10: Δοκιμές για διαφορετικό μέγεθος πληθυσμού
Μέγιστος αριθμός γενεών = 5.000



Σχήμα 2.11: Δοκιμές για διαφορετικό μέγεθος πληθυσμού
Μέγιστος αριθμός γενεών = 10.000