



ΑΡΙΣΤΟΤΕΛΕΙΟ
ΠΑΝΕΠΙΣΤΗΜΙΟ
ΘΕΣΣΑΛΟΝΙΚΗΣ

Πολυτεχνική Σχολή

Τμήμα Ηλεκτρολόγων Μηχανικών και Μηχανικών
Υπολογιστών

Project 2021-2022

Τζομίδης Νικόλαος-Φώτιος
ΑΕΜ: (9461)
tzomidis@ece.auth.gr

Τεχνικές Βελτιστοποίησης (7ο εξάμηνο)

13 Φεβρουαρίου 2022

Περιεχόμενα

Περιεχόμενα	i
Λίστα Σχημάτων	ii
Λίστα Πινάκων	iii
1 Εισαγωγή	1
1.1 Γενετικοί Αλγόριθμοι	1
1.2 Διαδικασίες του Γεντικού Αλγορίθμου	2
1.2.1 Κωδικοποίηση Γονιδίων	2
1.2.2 Χρωμοσώματα	2
1.2.3 Συνάρτηση Ικανότητας(Fitness Function)	3
1.2.4 Διασταυρώσεις(Crossover)	3
1.2.5 Μεταλλάξεις(Mutations)	4
2 Προσέγγιση και αναλυτική έκφραση συνάρτησης	5
3 Σχόλια - Παρατηρήσεις	9

Λίστα Σχημάτων

1.1	Διαδικασία crossover	3
2.1	Τρισδιάστατη αναπαράσταση της f	6
2.2	Τρισδιάστατη αναπαράσταση της προσέγγισης μας μέσω Γκαουσιανών	7

Λίστα Πινάκων

2.1	Πίνακας παραμέτρων των 15 γκαουσιανών	8
-----	---	---

Κεφάλαιο 1

Εισαγωγή

Αντικείμενο της εργασίας είναι η σχεδίαση ενός γενετικού αλγορίθμου για την προσέγγιση μιας συνάρτησης $f(u_1, u_2)$ της οποίας δεν γνωρίζουμε την αναλυτική μορφή παρά μόνο το ότι είναι συνεχής. Η ζητούμενη προσέγγιση θα γίνει με την σχεδίαση ενός γενετικού αλγορίθμου και θα βασιστεί σε γραμμικό συνδυασμό Γκαουσιανών συναρτήσεων της μορφής:

$$G(u_1, u_2) = e^{-[\frac{(u_1 - c_1)^2}{2\sigma_1^2} + \frac{(u_2 - c_2)^2}{2\sigma_2^2}]}$$

1.1 Γενετικοί Αλγόριθμοι

Οι Γενετικοί Αλγόριθμοι αποτελούν μια μέθοδο αναζήτησης βέλτιστων λύσεων και χρησιμοποιούνται ευρέως σε προβλήματα όπου η αναλυτική έκφραση της συνάρτησης, όπως και στην δική μας περίπτωση, είναι άγνωστη.

Η λειτουργία τους είναι εμπνευσμένη από την βιολογία και την θεωρία της εξέλιξης. Πιο συγκεκριμένα υπάρχει ο πληθυσμός που αποτελείται από το σύνολο των λύσεων όπου κάθε τέτοια λύση ονομάζεται *χρωμόσωμα* και κάθε παράμετρος του κωδικοποιούνται ως δυαδικοί αριθμοί. Αυτός ο πληθυσμός υφίσταται μια προσομοιωμένη γενετική εξέλιξη χρησιμοποιώντας διάφορους γενετικούς τελεστές όπως η επιλογή, η διασταύρωση και η μετάλλαξη. Στην πράξη δηλαδή από το σύνολο λύσεων του πληθυσμού εκκίνησης ακολουθείται μια σειρά τυχαίων ανασυνδυασμών και μεταλλάξεων των *χρωμοσωμάτων*. Οι πιο 'ικανές' από αυτές τις λύσεις συνεχίζουν να εξελίσσονται και να

ανασυνδυάζονται έως ότου να 'επιβιώσουν' οι καλύτερες από αυτές. Το κριτήριο για την εύρεση αυτών των καλύτερων χρωμοσωμάτων ονομάζεται *συνάρτηση ικανότητας*. Η όλη διαδικασία έχει διάφορες τεχνικές υλοποίησης που διαφέρουν από πρόβλημα σε πρόβλημα.

1.2 Διαδικασίες του Γεντικού Αλγορίθμου

1.2.1 Κωδικοποίηση Γονιδίων

Το πρόβλημα μας δεν επιτρέπει την αναπαράσταση των γονιδίων από bits 0, 1 όπως συνηθίζεται σε γενετικούς αλγορίθμους. Για αυτό τον λόγο έγινε χρήση πραγματικών αριθμών βασιζόμενοι στο γεγονός ότι τα χρωμοσώματα μας θα αποτελούνται από τα πλάτη τα κέντρα και τις τυπικές αποκλίσεις του γραμμικού συνδυασμού των γκαουσιανών συναρτήσεων. Το καθένα από αυτά όμως θα κινείται σε συγκεκριμένα όρια:

- Πλάτος: $A \in (f_{min}, f_{max})$, με f_{min} και f_{max} η ελάχιστη και η μέγιστη τιμή της συνάρτησης μας αντίστοιχα.
- Κέντρα (c_1, c_2) , με πεδία ορισμού βάσει της εκάστοτε συνάρτησης μας. Π.χ για πεδίο ορισμού $u_1 = (\alpha, \beta)$, $u_2 = (\gamma, \delta)$ θα έχουμε για τα κέντρα $c_1 \in (\alpha - 3, \beta + 3)$, $c_2 \in (\gamma - 3, \delta + 3)$.
- Τέλος, τυπική απόκλιση $\sigma \in (0.2, 1.3)$

1.2.2 Χρωμοσώματα

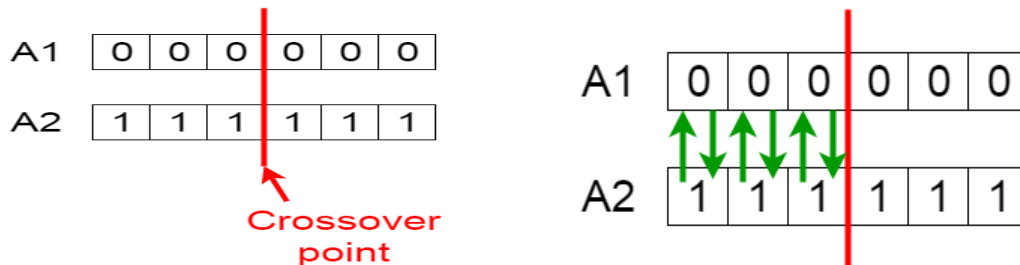
Πέρα από την κωδικοποίηση για το κάθε γονίδιο των χρωμοσωμάτων θα πρέπει να ορίσουμε αρχικό πληθυσμό λύσεων αλλά και τον αριθμό των 'καλύτερων χρωμοσωμάτων' που θα μεταφέρουμε στην επόμενη γενιά. Στην συνέχεια πρέπει να ορίσουμε και το ποσοστό *διασταυρώσεων (crossover)* προσέχοντας να διατηρούμε σταθερό τον αριθμό καθόλη την διάρκεια της διαδικασίας. Ενώ τέλος θα καθορίσουμε και την πιθανότητα *μετάλλαξης (mutation)* σε κάποια από τα επιμέρους γονίδια μας.

1.2.3 Συνάρτηση Ικανότητας(Fitness Function)

Κάθε λύση ελέγχεται για το πόσο κοντά φέρνει την αντίδραση του συστήματος μας στην επιθυμητή μέσω μιας συνάρτησης η οποία ονομάζεται *συνάρτηση ικανότητας(Σ.Ι)*. Οι λύσεις που σύμφωνα με την Σ.Ι θα βρίσκονται πιο κοντά στην επιθυμητή αναπαράγονται στην επόμενη γενιά. Εμείς για την προσέγγιση της συνάρτησης θα χρησιμοποιήσουμε το σφάλμα για την απόκλιση από της πραγματική τιμή, πιο συγκεκριμένα θα χρησιμοποιήσουμε το μέσο τετραγωνικό σφάλμα(mean squared error) όπου για $f(x_i)$ πραγματική τιμή της συνάρτησης και $\bar{f}(x_i)$ τη προσέγγιση μας θα είναι: $e_{mse} = (\frac{1}{n}) \cdot \sum (f(x_i) - \bar{f}(x_i))^2$ με n τον αριθμό των σημείων που χρησιμοποιήσαμε στην προσέγγιση.

1.2.4 Διασταυρώσεις(Crossover)

Στην διαδικασία του *crossover* μετά την επιλογή των 'ικανότερων ατόμων' μέσω της Σ.Ι κάθε ζευγάρι αυτών το ονομάζουμε *γονείς* και επιλέγουμε τυχαία ένα σημείο όπου θα γίνει το crossover. Π.χ:



A5

1	1	1	0	0	0
---	---	---	---	---	---

A6

0	0	0	1	1	1
---	---	---	---	---	---

(c) Νέο παιδί

Σχήμα 1.1: Διαδικασία crossover

Για επιλογή σημείου διασταύρωσης στο σημείο 3 των γονέων θα γίνει ανταλλαγή των γονιδίων μεταξύ τους από όπου θα προκύψει το νέο παιδί στο (c).

Εμείς για την προσέγγιση της συνάρτησης μας όμως δεν θα χρησιμοποιήσουμε την κλασσική μέθοδο διασταύρωσης μιας και η κωδικοποίηση μας δεν είναι δυαδική και τα αποτελέσματα δεν ήταν τα καλύτερα δυνατά. Αντί αυτής για δύο γονείς έστω A και B το νέο παιδί θα αποτελείται από το ημίθροισμα των όρων των δύο γονέων, δηλαδή με $\frac{A+B}{2}$.

1.2.5 Μεταλλάξεις(Mutations)

Σε κάποια από τα νέα 'παιδιά' μερικά από τα γονίδια τους θα υποστούν *μετάλλαξη(mutation)* δηλαδή για δυαδικό τύπο γονιδίων θα είχαμε εναλλαγή από 0 σε 1 και αντίστροφα. Στην δική μας συνάρτηση η μετάλλαξη ενός τυχαίου γονιδίου θα γίνει με εναλλαγή της τιμής του απο μια τιμή μέσα στα αντίστοιχα πραγματικά εύρη τιμών του που ορίσαμε παραπάνω. Για αυτή την διαδικασία ορίζουμε την πιθανότητα αυτής της μετάλλαξης στην αρχή του αλγόριθμου μας.

Κεφάλαιο 2

Προσέγγιση και αναλυτική έκφραση συνάρτησης

Για την αξιολόγηση του αλγορίθμου βάσει της εκφώνησης θα χρησιμοποιηθεί, για την παραγωγή των δεδομένων εισόδου-εξόδου, η:

$$f(u_1, u_2) = \sin(u_1 + u_2) \cdot \sin(u_2^2)$$

με $u_1 \in [-1, 2]$ και $u_2 \in [-2, 1]$.

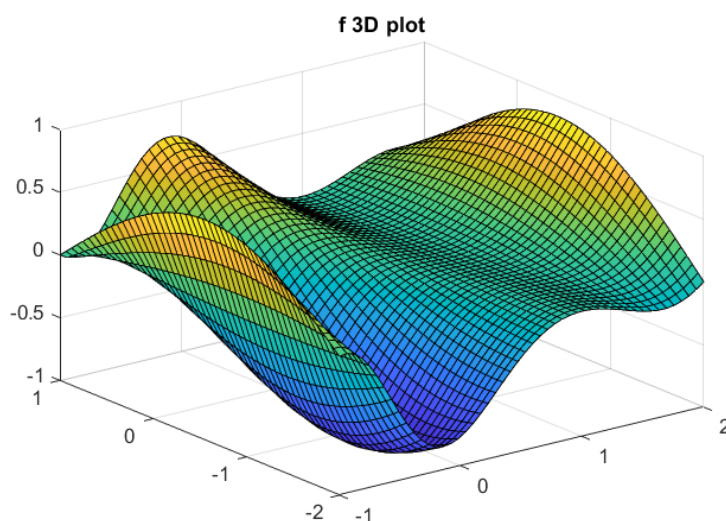
Όπως είδαμε και στην εισαγωγή θα πρέπει να χρησιμοποιήσουμε γραμμικό συνδυασμό Γκαουσιανών συναρτήσεων της εξής μορφής:

$$G(u_1, u_2) = e^{-[\frac{(u_1 - c_1)^2}{2\sigma_1^2} + \frac{(u_2 - c_2)^2}{2\sigma_2^2}]}$$

Από την τρισδιάστατη αναπαράσταση όμως της δοθείσας συνάρτησης για την παραγωγή των δεδομένων μας βλέπουμε ότι η συνάρτηση παίρνει και αρνητικές τιμές. Επομένως η χρησιμοποίηση Γκαουσιανών της παραπάνω μορφής δεν θα είχε πάρα πολύ νόημα μιας και θα ήταν συνδυασμός μόνο θετικών τιμών με αποτέλεσμα η προσέγγιση του γενετικού αλγορίθμου μας να μην θεωρηθεί ακριβής. Αντί αυτού θα γίνει χρήση Γκαουσιανών συναρτήσεων της μορφής:

$$G(u_1, u_2) = A \cdot e^{-[\frac{(u_1 - c_1)^2}{2\sigma_1^2} + \frac{(u_2 - c_2)^2}{2\sigma_2^2}]}$$

Όπου πρακτικά δηλαδή προσθέτουμε και έναν παράγοντα A του πλάτους ο οποίος θα δίνει και τις επιθυμητές αρνητικές τιμές.



Σχήμα 2.1: Τρισδιάστατη αναπαράσταση της f

Υπολογίζουμε λοιπόν τα f_{min} και f_{max} της συνάρτησης παραγωγής δεδομένων (fLimits.m) και ορίζουμε έτσι το εύρος τιμών των πλατών A που αναφέραμε παραπάνω και τελικά θα έχουμε.

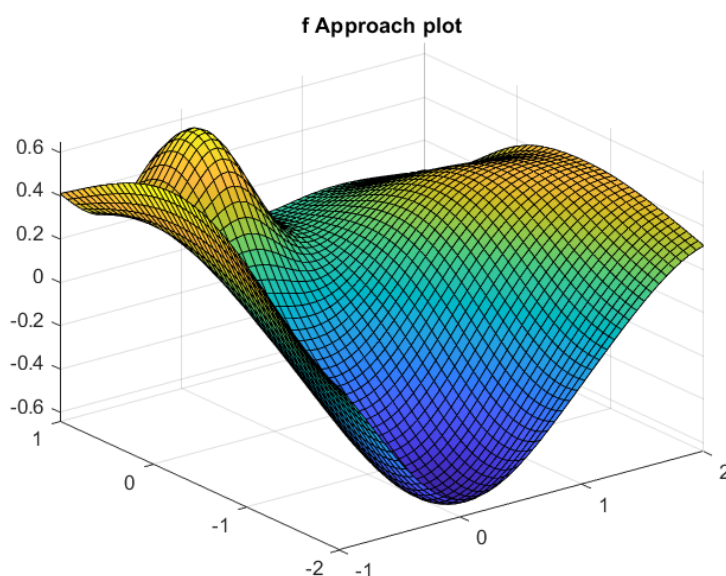
- Πλάτος: $A \in (-0.9996, 0.8415)$,
- Κέντρα (c_1, c_2) με $c_1 \in (-4, 5)$, $c_2 \in (-5, 4)$,
- Τυπική απόκλιση $\sigma_1, \sigma_2 \in (0.2, 1.3)$.

Ο πληθυσμός ορίζεται στα 50 χρωμοσώματα, με τον πληθυσμό της επόμενης γενιάς να αποτελείται από το 30% των καλύτερων της προηγούμενης, το 10% τυχαίων λύσεων και το υπόλοιπο 60% δημιουργείται μέσω του crossover.

Έπειτα από αρκετές δοκιμές και πειραματισμούς πάνω στον κώδικα καταλήξαμε σε πιθανότητα μετάλλαξης ίση με 33% δηλαδή μεταλλάσσουμε το 1/3 του πληθυσμού μας. Ενώ παρατηρήθηκε η μεγάλη σημασία που έχει στον γενετικό αλγόριθμο η διαδικασία της μετάλλαξης και πόσο αυτή επηρεάζει στην σύγκλιση του ανάλογα την επιλογή που θα κάνουμε.

Τέλος χρησιμοποιούμε και τις 15 γκαουσιανές που μας επιτρέπει η εκφώνηση για την επίτευξη του μικρότερου δυνατού σφάλματος.

Το καλύτερο μας αποτέλεσμα φαίνεται να προέρχεται από μετάβαση του $1/3$ του πληθυσμού όπως ήδη αναφέραμε και ενώ τρέχουμε τον αλγόριθμο μας για 8000 γενιές. Παρακάτω φαίνεται η γραφική παράσταση που πετύχαμε με την προσέγγιση μέσω του γραμμικού συνδυασμού των 15 γκαουσιανών. Για τον συγκεκριμένο 'καλύτερο πληθυσμό' μας είχαμε $mse = 0.0667$ αρκετά ικανοποιητικό δηλαδή κάτι που γίνεται και αντιληπτό από την τρισδιάστατη αναπαράσταση η οποία φαίνεται να 'μοιάζει' με την παραπάνω τρισδιάστατη αναπαράσταση της f που χρησιμοποιήσαμε για την παραγωγή των δεδομένων μας.



Σχήμα 2.2: Τρισδιάστατη αναπαράσταση της προσέγγισης μας μέσω Γκαουσιανών

Παρακάτω μπορούμε να δούμε τον πίνακα με τις τιμές των 15 γκαουσιανών συναρτήσεων που πήραμε από τον παραπάνω 'καλύτερο πληθυσμό' του γενετικού αλγορίθμου μας.

Γκαουσιανές	Πλάτη A	c1	c2	s1	s2
1st Gaussian	0.12121	0.38235	0.12121	0.12121	0.12121
2nd Gaussian	0.67669	1.6662	0.67669	0.67669	0.67669
3rd Gaussian	-0.11364	1.4648	-0.11364	-0.11364	-0.11364
4th Gaussian	1.3537	0.8014	1.3537	1.3537	1.3537
5th Gaussian	-0.15692	4.1251	-0.15692	-0.15692	-0.15692
6th Gaussian	0.36717	0.68604	0.36717	0.36717	0.36717
7th Gaussian	0.035194	-1.7188	0.035194	0.035194	0.035194
8th Gaussian	-0.28896	0.83211	-0.28896	-0.28896	-0.28896
9th Gaussian	0.54106	3.1217	0.54106	0.54106	0.54106
10th Gaussian	1.0336	-2.8502	1.0336	1.0336	1.0336
11th Gaussian	1.3223	1.0635	1.3223	1.3223	1.3223
12th Gaussian	-0.67534	-0.75786	-0.67534	-0.67534	-0.67534
13th Gaussian	0.31916	-0.68906	0.31916	0.31916	0.31916
14th Gaussian	0.67273	-2.1097	0.67273	0.67273	0.67273
15th Gaussian	0.16508	2.2142	0.16508	0.16508	0.16508

Πίνακας 2.1: Πίνακας παραμέτρων των 15 γκαουσιανών

Για αυτή την τελική λύση του γενετικού μας αλγορίθμου λοιπόν θα έχουμε μια χαμηλής πολυπλοκότητας αναλυτική έκφραση της συνάρτησης f το οποίο είναι και το ζητούμενο της παρούσας εργασίας. Πιο συγκεκριμένα με βάση τον παραπάνω πίνακα:

$$\text{Για } G_i(u_1, u_2) = A_i \cdot e^{-\left[\frac{(u_1 - c_{1_i})^2}{2\sigma_{1_i}^2} + \frac{(u_2 - c_{2_i})^2}{2\sigma_{2_i}^2}\right]}, \text{ όπου } i = 1, 2, \dots, 15$$

Η αναλυτική έκφραση της f γράφεται ως:

$$f(u_1, u_2) = \sum_{i=1}^{15} G_i(u_1, u_2)$$

Κεφάλαιο 3

Σχόλια - Παρατηρήσεις

Όπως είδαμε καταλήξαμε σε μια σχετικά καλή προσέγγιση για την ζητούμενη f , όμως σίγουρα ο αλγόριθμος που υλοποιήσαμε δεν μπορεί να θεωρηθεί 100% ακριβής και σωστός. Αντιμετωπίστηκαν προβλήματα στο fitness score των πληθυσμών τα οποία σε αρκετές δοκιμές φαίνεται να μην μειώνονται με την πάροδο των γενεών μας, κάτι δηλαδή που θα περιμέναμε σε έναν γενετικό αλγόριθμο. Αντί αυτού παρατηρήθηκε ταλαντώση του σφάλματος παρότι αυτή ήταν ανάμεσα σε μικρές τιμές οι οποίες θεωρητικά θα ήταν επιθυμητές. Επίσης σε δοκιμές του αλγορίθμου για ίδιες αρχικές συνθήκες εκκίνησης οι αποκλίσεις στα αποτελέσματα ήταν συχνά μεγάλες. Επιλογικά λοιπόν η παραπάνω υλοποίηση δεν είναι σε καμία περίπτωση τέλεια προσεγγίζει όμως σε έναν καλό βαθμό το ζητούμενο πρόβλημα και δίνει μια ικανοποιητική χαμηλής πολυπλοκότητας αναλυτική εκφραση για την f .