

ΤΕΧΝΙΚΕΣ ΒΕΛΤΙΣΤΟΠΟΙΗΣΗΣ-PROJECT

Αποστολίδου Αθηνά

AEM: 10400

Θέμα : Γενετικοί Αλγόριθμοι

Σκοπός της εργασίας είναι να γίνει εκτίμηση της τιμής μίας δοθείσας συνάρτησης δύο μεταβλητών από έναν γραμμικό συνδυασμό το πολύ 15 Gaussian συναρτήσεων. Για αυτές τις συναρτήσεις Gaussian θέλουμε να βρούμε τον συντελεστή τους, τη μέση τιμή τους και την διακύμανση τους.

1)Γενικές πληροφορίες

Οι γεννητικοί αλγόριθμοι είναι μια κατηγορία εξελικτικών αλγορίθμων που εμπνέονται από τη φυσική εξέλιξη και τη γενετική. Χρησιμοποιούνται για την επίλυση προβλημάτων βελτιστοποίησης και αναζήτησης. Σε έναν γενετικό αλγόριθμο, ένας αρχική πληθυσμός ατόμων παριστάνει πιθανές λύσεις για το πρόβλημα. Η διαδικασία εξέλιξης περιλαμβάνει τρεις βασικές λειτουργίες: επιλογή, διασταύρωση και μετάλλαξη. Κατά τη διάρκεια της εξέλιξης, τα άτομα που έχουν υψηλότερη προσαρμοστικότητα (βάσει ενός κριτηρίου fitness) έχουν μεγαλύτερη πιθανότητα να επιλεγούν για διασταύρωση, προκειμένου να δημιουργήσουν νέα άτομα. Η διαδικασία αυτή επαναλαμβάνεται κατάλληλο αριθμό φορών, με την ελπίδα ότι η πληθυσμιακή εξέλιξη θα οδηγήσει σε βέλτιστες λύσεις για το πρόβλημα.

2)Ορισμοί

- Πληθυσμός : Είναι το σύνολο των λύσεων
- Χρωμόσωμα : Μια επιμέρους λύση
- Γονίδια : Τα χαρακτηριστικά κάθε χρωμοσώματος
- Γενιά : Ο πληθυσμός συνεχώς ανανεώνεται με τις διαδικασίες της επιλογής, της διασταύρωσης και της μετάλλαξης. Κάθε ανανέωση του πληθυσμού μας δίνει μια γενιά.
- Fitness function : Είναι συνάρτηση μέσο της οποίας εκτιμάται η καταλληλότητα μιας λύσης. Ανάλογα με το πρόβλημα επιλέγεται ένας τρόπος ποσοτικοποίησης της ικανότητας του κάθε χρωμοσώματος

- Επιλογή : Η μέθοδος με την οποία επιλέγεται ποια χρωμοσώματα θα περάσουν στην επόμενη γενιά
- Διασταύρωση : Με την μέθοδο αυτή συνδυάζονται τα γονίδια 2 χρωμοσωμάτων
- Μετάλλαξη : Γίνεται τυχαία αλλαγή ενός γονιδίου σε κάποια χρωμοσώματα

3) Το πρόβλημα μας

- Ως πληθυσμός επιλεχτήκαν τα 50 χρωμοσώματα
- Κάθε χρωμόσωμα θα αποτελείται από $5 \cdot n$ γονίδια όπου n ο αριθμός των Gaussian συναρτήσεων που θα χρησιμοποιηθούν. Τα διάφορα γονίδια θα παίρνουν τυχαίες πραγματικές τιμές. Το εύρος των τυχαίων τιμών που θα λάβει κάθε γονίδιο εξαρτάται από ποιο χαρακτηριστικό μοντελοποιεί.
- $$G(u1, u2) = A * e^{-\left(\frac{(u1-c1)^2}{2 \cdot \sigma1^2} + \frac{(u2-c2)^2}{2 \cdot \sigma2^2}\right)}$$
 1. Το πλάτος A θα παίρνει τιμές από το μέγιστο της f που θέλουμε να εκτιμήσουμε μέχρι το ελάχιστο.
 2. Τα $c1$ και $c2$ έχουν εύρος τιμών $(u1_min - 1, u1_max + 1)$ και $(u2_min - 1, u2_max + 1)$ αντίστοιχα.
 3. Οι διακυμάνσεις $\sigma1$ και $\sigma2$ παίρνουν τιμές από 0.3 έως 1,1.
 4. Το πρώτο στη σειρά χρωμόσωμα αντιστοιχεί στην τιμή του πλάτους, το επόμενο στο $c1$, το 3^ο στο $c2$, το 4^ο στο $\sigma1$, το 5^ο στο $\sigma2$ και η σειρά αυτή επαναλαμβάνεται ανάλογα με τον αριθμό των Gaussian συναρτήσεων.
- Στην επιλογή χρωμοσωμάτων της νέας γενιάς κρατήσαμε τα 10 καλύτερα από την προηγούμενη γενιά και συμπληρώσαμε τα υπόλοιπα με διασταύρωση τυχαίων χρωμοσωμάτων με πιθανότητα επιλογή που εξαρτάται από την fitness function. Η διαδικασία αυτή έγινε χρησιμοποιώντας τον τροχό της τύχης.
- Ως fitness function ορίστηκε η αντίστροφη τιμή του μέσου τετραγωνικού σφάλματος.
- Για να τερματίσει ο αλγόριθμός μας όταν η τιμή της fitness function γίνει μεγαλύτερη του 250.
- Στη διαδικασία της διασταύρωσης παίρνουμε τον μέσο όρο των αντιστοίχων γονιδίων από τα δυο χρωμοσώματα που διασταυρώνουμε. Αξίζει να σημειωθεί ότι δεν επιλέξαμε να κάνουμε διασταύρωση παίρνοντας αυτούσια κομμάτια και από τα δύο χρωμοσώματα καθώς κάτι τέτοιο δεν θα έβρισκε καλή

εφαρμογή στην περίπτωση μας καθώς τα γονίδια που κωδικοποιούν την ίδια Gaussian έχουν μεγάλη σύνδεση ενώ τα γονίδια που κωδικοποιούν διαφορετικές Gaussian όχι.

- Για την επιλογή του αριθμού των Gaussian καθώς και για τις παραμέτρους της μετάλλαξης θα γίνει αναφορά στην συνέχεια.

4)Εξαγωγή δεδομένων ελέγχου

Η συνάρτηση που μας δόθηκε για έλεγχο είναι η :

$$f(u1,u2) = \sin(u1 + u2) * \sin(u2^2)$$

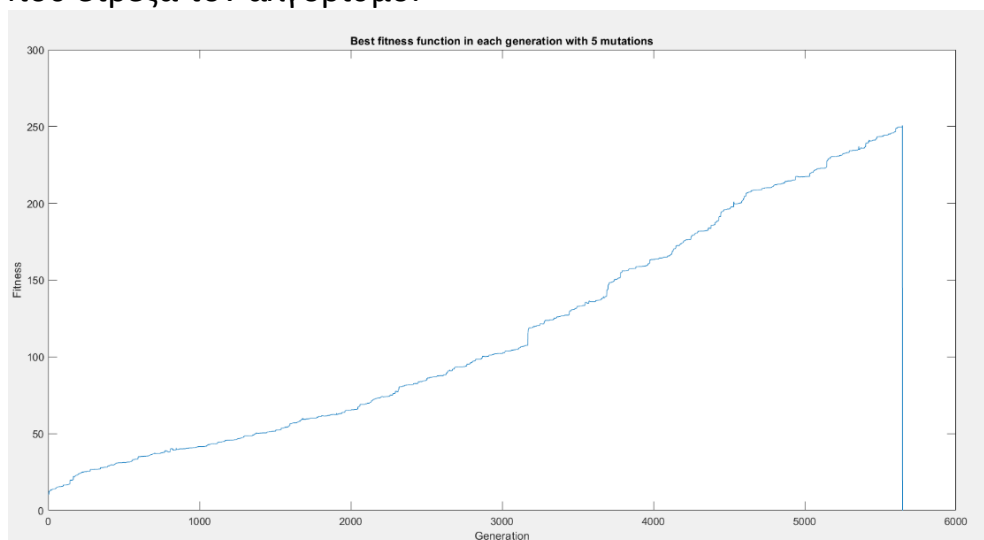
$$\text{Με } u1 \in (-1,2) \text{ και } u2 \in (-2,1)$$

Από αυτή την συνάρτηση θα υπολογίζω την τιμή της σε 900 σημεία την οποία θα συγκρίνω με την τιμή του αθροίσματος των Gaussian στα σημεία αυτά. Ακόμα, στους υπολογισμούς μου λαμβάνω υπ' όψη το ελάχιστό της, το μέγιστό της καθώς και το πεδίο ορισμού της.

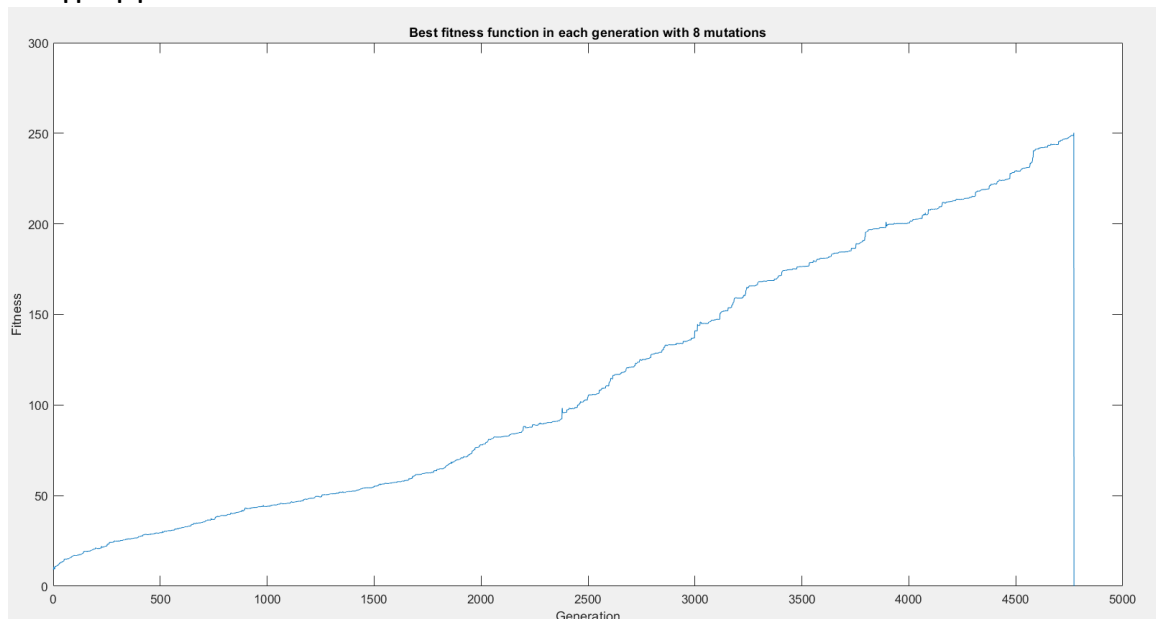
5) Κάνω διάφορες δοκιμές για να δω ποια είναι μια ικανοποιητική συχνότητα μεταλλάξεων για να χρησιμοποιηθεί:

Εξετάζω την συχνότητα των μεταλλάξεων. Αυτό θα γίνει θεωρώντας αρχικά ότι σε κάθε μετάλλαξη αλλάζει ένα γονίδιο και ότι ο αριθμός των Gaussian είναι σταθερός και ίσος με 10.Θα κάνω δοκιμές για 5,8 και 10 μεταλλάξεις. Επειδή υπάρχει ο παράγοντας της τύχης στη διαδικασία θα τρέξω την κάθε περίπτωση 4 φορές.

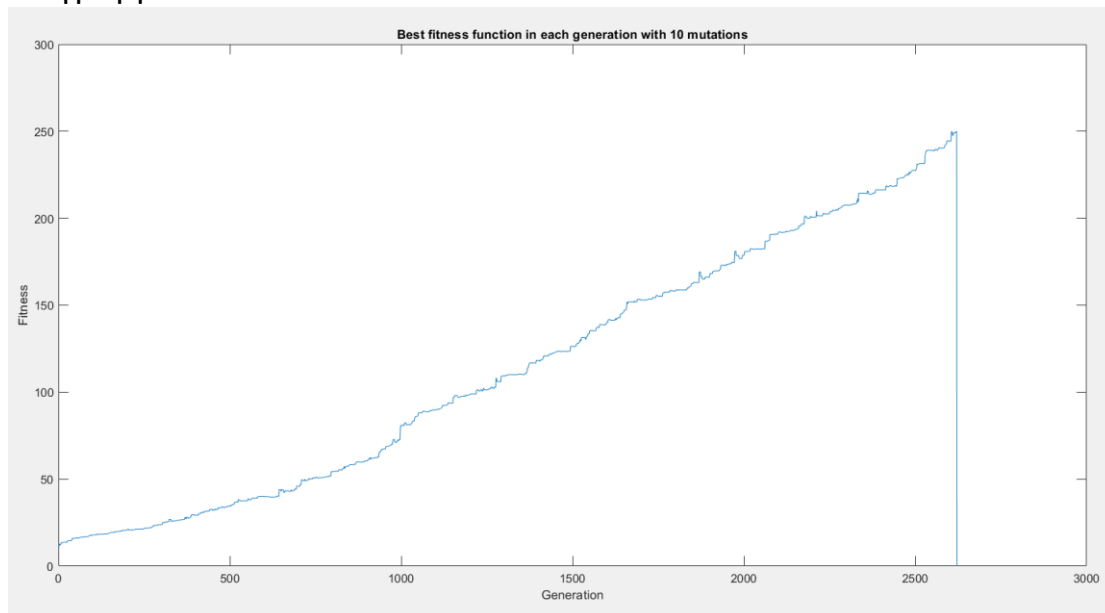
α)Για $m=5$. Την πρώτη φορά ο αλγόριθμος εφτάσε την προσδοκώμενη τιμή της fitness function. Την 2^η φορά έφτασε αυτή την τιμή στην γενιά 5831. Την 3^η φορά στην γενιά 5674 και την 4^η φορά στην 1530. Δίνεται ενδεικτικά το διάγραμμα από την 3^η φορά που έτρεξα τον αλγόριθμο.



β) Για $m = 8$. Την 1^η φορά χρειάστηκε 4815 γενιές. Την 2^η 2787 γενιές. Την 3^η 4772 γενιές και την 4^η 1520 γενιές. Ενδεικτικά ένα από τα διαγράμματα.



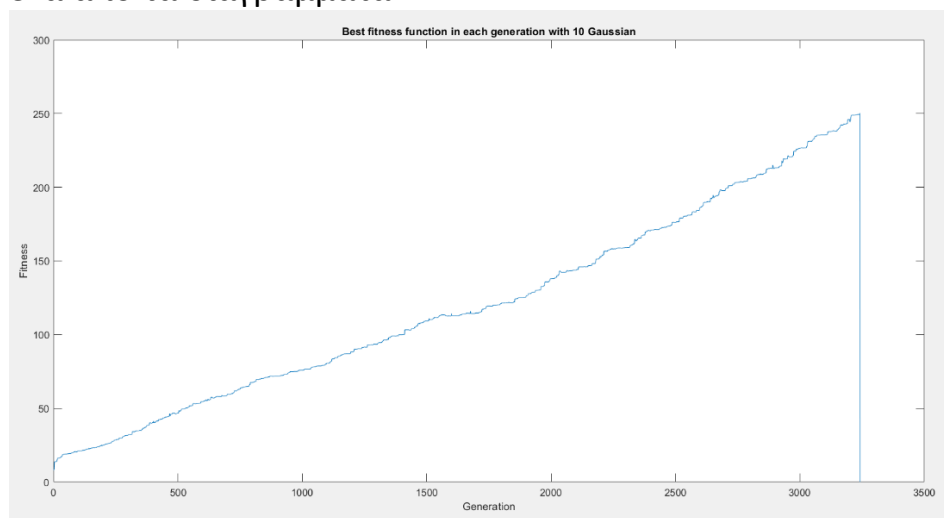
γ) Για $m = 10$. Την 1^η φορά χρειάστηκε 4265 γενιές. Την 2^η 4682 γενιές. Την 3^η 2662 γενιές και την 4^η 5264 γενιές. Ενδεικτικά ένα από τα διαγράμματα.



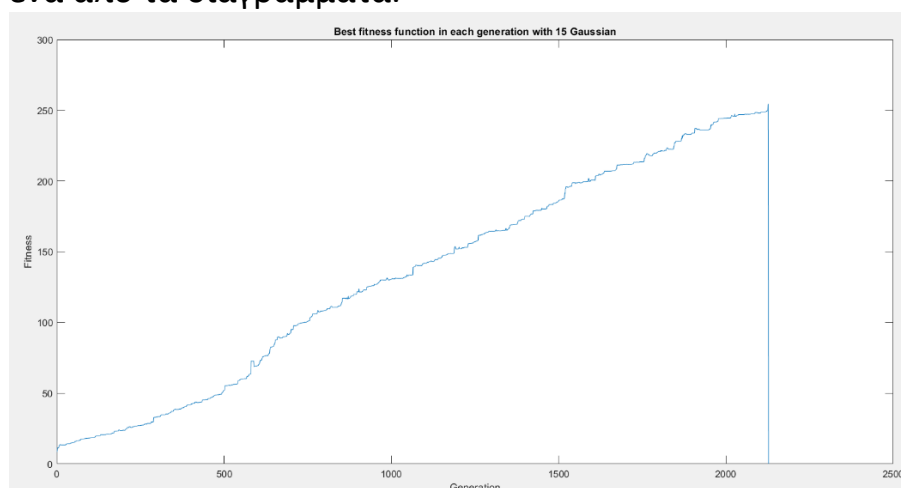
Από τα αποτελέσματα που παρουσιάστηκαν καθώς και από κάποιες ακόμα επαναλήψεις του αλγορίθμου φαίνεται ότι η τιμή του $m = 8$ εξασφαλίζει τα καλύτερα αποτελέσματα.

6) Έπειτα αναζητώ ποιος είναι ο βέλτιστος αριθμός των Gaussian συναρτήσεων που θα χρησιμοποιηθούν. Κρατώ σταθερό τον αριθμό των μεταλλάξεων στις 8. Εξετάζω για διαφορετικό αριθμό Gaussian πως θα επηρεαστεί η ταχύτητα που ο αλγόριθμος πιάνει την επιθυμητή ακρίβεια. Θα εξετάσω ενδεικτικά τι συμβαίνει στην περίπτωση των 5, 10 και 15 Gaussian:

- a) Για 5 Gaussian: Στην πλειοψηφία των φορών που τρέξαμε τον αλγόριθμο αυτός δεν συνέκλινε μέσα στις 10000 γενιές που είχαμε ορίσει.
- b) Για 10 Gaussian: Την 1η φορά χρειάστηκε 4647 γενιές. Την 2η 4676 γενιές. Την 3η 3241 γενιές και την 4η 4155 γενιές. Ενδεικτικά ένα από τα διαγράμματα.

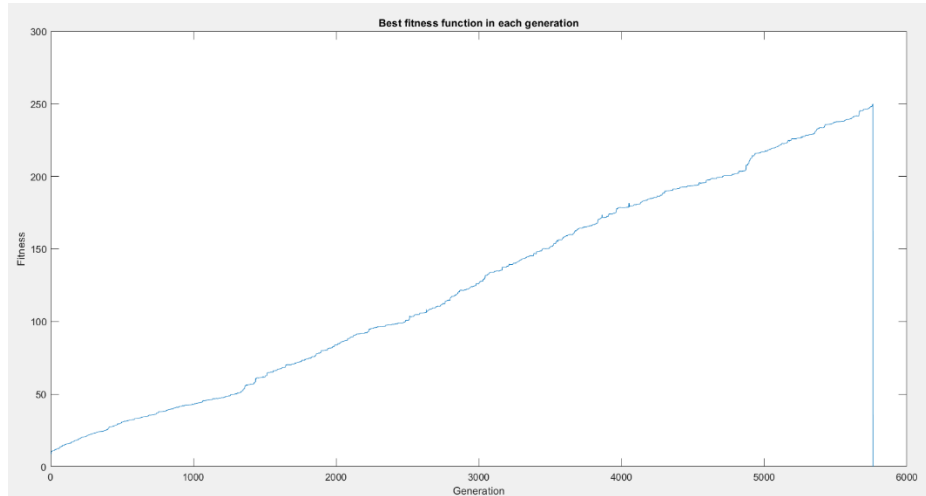


- c) Για 15 Gaussian: Την 1η φορά χρειάστηκε 4488 γενιές. Την 2η 5888 γενιές. Την 3η 2127 γενιές και την 4η 2652 γενιές. Ενδεικτικά ένα από τα διαγράμματα.



Από τα παραπάνω βλέπουμε ότι οι 15 Gaussian κάνουν τον αλγόριθμο να συγκλίνει πιο γρήγορα.

7) Θα εξετάσω τώρα τι θα γίνει αν αντί να κρατήσω τα 10 καλύτερα άτομα από την μια γενιά στην επόμενη να κρατήσω τα 20. Σε αυτή την περίπτωση έχουμε: την 1^η φορά 5654 γενιές, την 2^η 4686 γενιές, την 3^η 5763 γενιές και την 4^η 6256 γενιές. Ενδεικτικά ένα από τα διαγράμματα.



Οπότε μια τέτοια αλλαγή δεν μας συμφέρει.

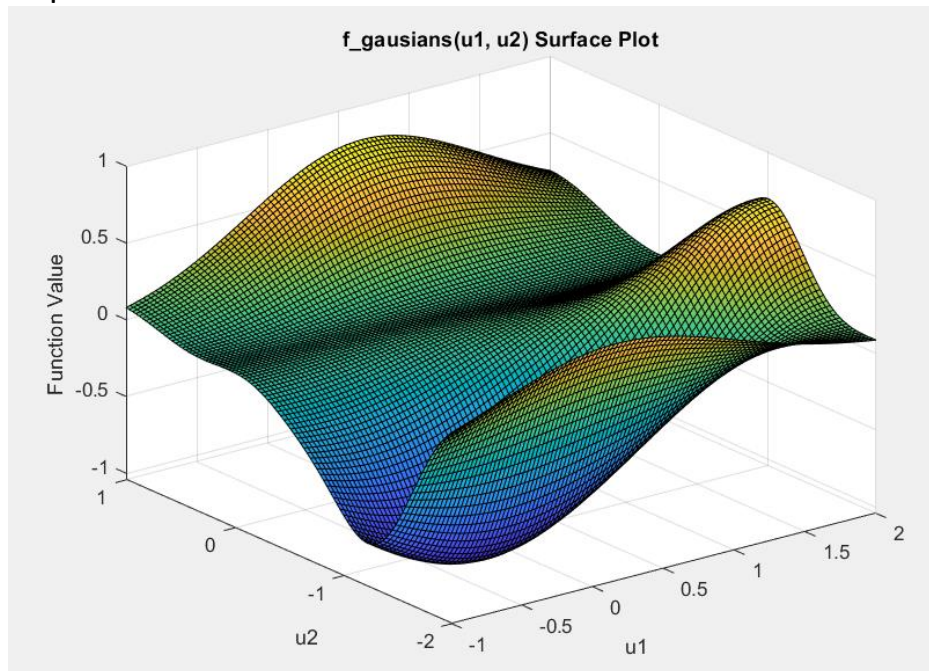
8) Η τελική μας λύση θα είναι με 8 μεταλλάξεις, 15 Gaussian και τα 10 άτομα της προηγούμενης γενιάς να επιβιώνουν στην επόμενη.

Στον παρακάτω πίνακα εμφανίζεται η λύση του αλγορίθμου.

Αριθμός Gaussian	Πλάτος	Κέντρο c1	Κέντρο c2	Διακύμανση S1	Διακύμανση s2
1 ^η	-0.16962	0.34697	0.35586	0.91664	0.37806
2 ^η	-0.65656	0.10997	-1.31089	1.00989	0.37897
3 ^η	-0.41744	-0.54784	-0.71097	0.83567	0.56529
4 ^η	0.05154	0.47603	-1.72462	0.93091	0.99463
5 ^η	0.35603	1.92563	-1.12468	0.47750	0.35449
6 ^η	0.00530	1.22257	-0.01889	0.73908	0.66055
7 ^η	0.50150	0.07103	-1.91929	0.86812	0.81875
8 ^η	0.17932	-0.42054	-0.12986	0.92013	0.32326
9 ^η	0.46520	1.81474	-1.00088	0.90941	0.32929
10 ^η	-0.21234	0.38831	0.34142	0.77661	0.38976
11 ^η	-0.24494	0.80119	0.10653	0.76588	0.48552
12 ^η	0.22013	0.16419	-1.89977	1.04236	0.79160
13 ^η	0.51542	0.45106	0.91473	0.74084	0.82552

14 ^η	0.44095	0.98199	0.91533	0.87097	0.61581
15 ^η	-0.79210	0.02553	-1.33613	0.98817	0.33893

Το γράφημα του γραφικού συνδυασμού των Gaussian φαίνεται παρακάτω:



Όπως διαπιστώνουμε είναι σχεδόν ταυτόσημο με το διάγραμμα της δοθείσας συνάρτησης πράγμα που επιβεβαιώνει την ορθότητα του αποτελέσματος μας.

