ΤΕΧΝΙΚΕΣ ΒΕΛΤΙΣΤΟΠΟΙΗΣΗΣ

5η Εργαστηριακή Άσκηση

Ελαχιστοποίηση συνάρτησης πολλών μεταβλητών
Μέθοδοι Φραγμού - Ποινής

ΑΝΤΡΕΑΣ ΧΑΤΖΗΘΩΜΑΣ

ΑΕΜ: 8026

<u>Εισαγωγή</u>

Στην 5η εργαστηριακή άσκηση ασχοληθήκαμε με το πρόβλημα ελαχιστοποίησης μιας δοσμένης συνάρτησης πολλών μεταβλητών $f: \mathbb{R}^n \to \mathbb{R}$ χρησιμοποιώντας την (εσωτερική) μέθοδο φραγμού αλλά και την (εξωτερική) μέθοδο ποινής.

Για την εργαστηριακή άσκηση θεωρούμε τις συναρτήσεις $f_1, f_2: \mathbb{R}^2 \to \mathbb{R}$ με αναλυτικούς τύπους:

$f_1(x) = x_1\, x_2 + 2(x_1 - x_2)^2$

$f_2(x) = (x_1 - x_2)^2$

και τα προβλήματα ελαχιστοποίησης:

$min\, f_1(x)$		$min\, f_2(x)$
$3 \leq x_1 \leq 30$	και	$x_1 \leq -1$
$-25 \leq x_2 \leq -5$		$x_2 \leq -1$

(<u>Σημείωση</u>: η $f_2(x)$ στην εκφώνηση της άσκησης δίνεται ως $g(x)$ αλλά μετονομάστηκε σε $f_2(x)$ έτσι ώστε να μην την μπερδεύουμε με τους περιορισμούς οι οποίοι στο βιβλίο αναγράφονται ως $g_i(x)$)

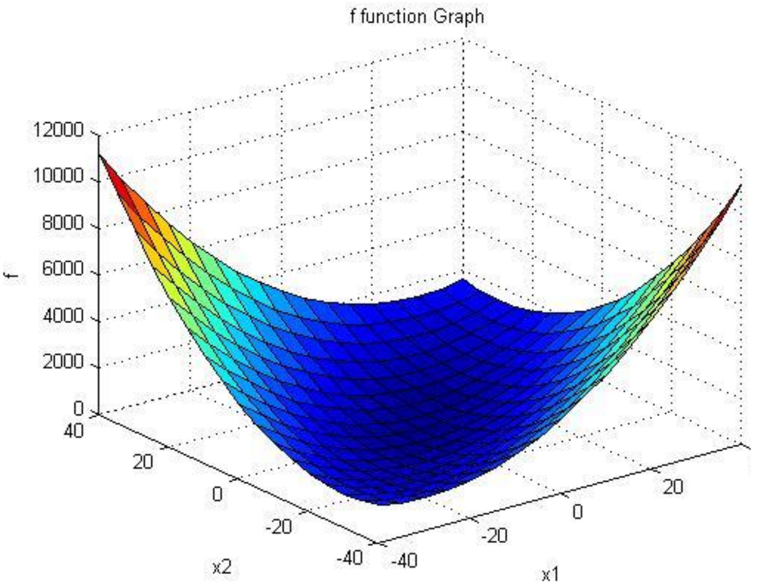
Αφού μελετήσαμε προσεκτικά την κατάλληλη θεωρία από το βιβλίο, προχωρήσαμε στην επίλυση των προβλημάτων θεωρητικά χρησιμοποιώντας το θεώρημα Karush-Kuhn-Tucker αλλά και αλγοριθμικά χρησιμοποιώντας τις μεθόδους φραγμού/ποινής.

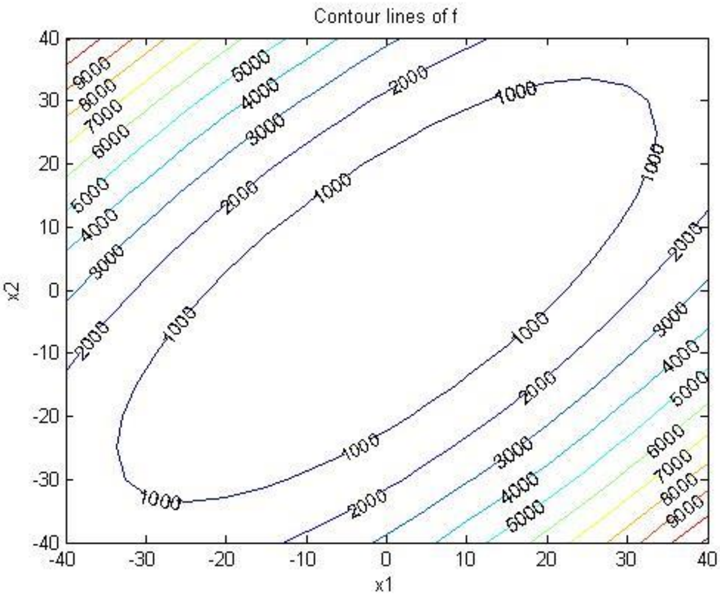
Για τη δική μας ευκολία, δημιουργήσαμε τα παρακάτω scripts:

- **f1.m & f2.m**: επιστρέφουν την τιμή των συναρτήσεων f1, f2 σε ένα σημείο x.
- **grad_f1.m & grad_f2.m:** επιστρέφουν το διάνυσμα κλίσης των συναρτήσεων f1, f2 σε ένα σημείο x.
- **g1.m & g2.m:** επιστρέφουν την τιμή των περιορισμών σε ένα σημείο x του 1ου και 2ου προβλήματος αντίστοιχα.
- **grad_g1 & grad_g2:** επιστρέφουν το διάνυσμα κλίσης των περιορισμών σε ένα σημείο x.

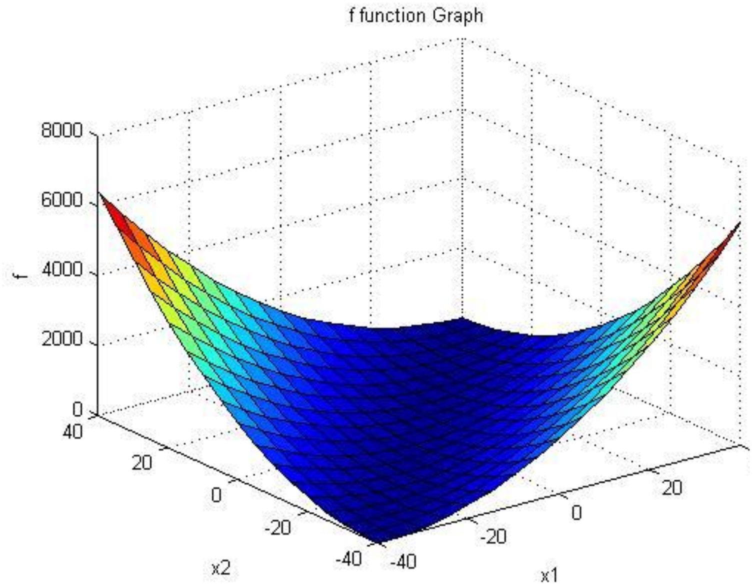
Ακολούθως, δημιουργήσαμε τη συνάρτηση plot_f(f), της οποίας η υλοποίηση βρίσκεται στο αρχείο **plot_f.m**. Καλώντας τη συνάρτηση plot_f(@f1) δημιουργείται το γράφημα και το διάγραμμα ισοβαρών καμπυλών της 1ης συνάρτησης. Αντίστοιχα, καλούμε τη συνάρτηση plot_f(@f2) για την 2η συνάρτηση. Τα διαγράμματα φαίνονται παρακάτω:

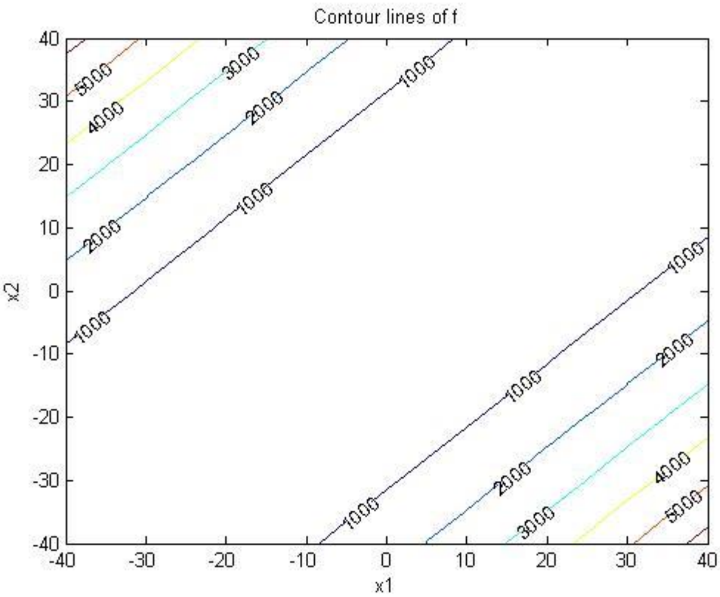
$$f_1(x) = x_1\,x_2 + 2(x_1 - x_2)^2$$





$$f_2(x) = (\,x_1 - x_2\,)^2$$





A. Karush-Kuhn-Tucker

Αφού μελετήσαμε την αντίστοιχη θεωρία του βιβλίου προχωρήσαμε στη θεωρητική επίλυση των προβλημάτων. Το θεώρημα Karush-Kuhn-Tucker λύθηκε χειρόγραφα και ακολούθως το αντιγράψαμε σε ηλεκτρονική μορφή όπως θα δούμε παρακάτω.

Πρόβλημα 1ο

Η αντικειμενική συνάρτηση του προβλήματος είναι : $f_1(x) = x_1x_2 + 2(x_1 - x_2)^2$

Η $f_1(x)$ είναι συνεχώς παραγωγίσιμη και η κλίση της είναι: $\nabla f_1(x) = \begin{bmatrix} 4x_1 & -3x_2 \\ -3x_1 & 4x_2 \end{bmatrix}$

Ο Εσσιανός πίνακας της συνάρτησης είναι : $H_{f_1}(x) = \begin{bmatrix} 4 & -3 \\ -3 & 4 \end{bmatrix}$.

Παρατηρούμε ότι εσσιανός πίνακας είναι συμμετρικός και οι ιδιοτιμές του είναι $\lambda_1 = 1\ \kappa\alpha\iota\ \lambda_2 = 7$. Επομένως ο πίνακας είναι θετικά ορισμένος.

Οπότε καταλήγουμε στο συμπέρασμα ότι η $f_1(x)$ είναι γνήσια κυρτή συνάρτηση.

Οι περιορισμοί του προβλήματος είναι :

- $g_1(x) = 3 - x_1 \leq 0$
- $g_2(x) = x_1 - 30 \leq 0$
- $g_3(x) = -25 - x_2 \leq 0$
- $g_4(x) = x_2 + 5 \leq 0$

Οι $g_i(x)$ είναι γραμμικές συναρτήσεις των x_1, x_2 άρα είναι κυρτές συναρτήσεις.

Εφόσον υπάρχουν σημεία του πεδίου ορισμού της $f_1(x)$ που ικανοποιούν τους περιορισμούς, το πρόβλημα μας είναι κυρτό και υπερ–αποτελούμενο.

Οι συνθήκες Karush-Kuhn-Tucker είναι οι ακόλουθες:

1. $\lambda_i \geq 0 \, , i = 1, \dots ,4$

2. $\lambda_1 g_1(x^*) = 0 => \lambda_1(3 - x_1^*) = 0$

$\lambda_2 g_2(x^*) = 0 => \lambda_2(x_1^* - 30) = 0$

$\lambda_3 g_3(x^*) = 0 => \lambda_3(-x_2^* - 25) = 0$

$\lambda_4 g_4(x^*) = 0 => \lambda_4(x_2^* + 5) = 0$

3. $4x_1^* - 3x_2^* - \lambda_1 + \lambda_2 = 0$

$-3x_1^* + 4x_2^* - \lambda_3 + \lambda_4 = 0$

Λύνοντας το παραπάνω σύστημα καταλήγουμε σε 9 λύσεις από τις οποίες μόνο η μια τηρεί τους περιορισμούς. Οι λύσεις φαίνονται στον παρακάτω πίνακα:

Λύση	$\boldsymbol{x_1^*}$	$\boldsymbol{x_2^*}$	$\boldsymbol{\lambda_1}$	$\boldsymbol{\lambda_2}$	$\boldsymbol{\lambda_3}$	$\boldsymbol{\lambda_4}$	**Περιορισμοί**
1η	-3.75	-5	0	0	0	8.75	Όχι ($x_1 \leq 3$)
2η	0	0	0	0	0	0	Όχι ($x_2 \geq -5$)
3η	-18.75	-25	0	0		0	Όχι ($\lambda_2 \leq 0$)
4η	30	22.5	0	-52.5	0	0	Όχι ($\lambda_2 \leq 0$)
5η	30	-5	0	-135	0		Όχι ($\lambda_2 \leq 0$)
6η	30	-25	0	-195		0	Όχι ($\lambda_2 \leq 0$)
7η	3	-2.25		0	0	0	Όχι ($x_2 \geq -5$)
8η	**3**	**-5**	**27**	**0**	**0**	**29**	**Ναι**
9η	3	-25	87	0	-109	0	Όχι ($\lambda_3 \leq 0$)

Όπως βλέπουμε και από τον πίνακα, μόνο η λύση 8 ικανοποιεί τους περιορισμούς του προβλήματος. Επομένως έχουμε:

$[x_1 \; x_2] = [3 - 5]$ και $[\lambda_1 \; \lambda_2 \; \lambda_3 \; \lambda_4] = [27 \; 0 \; 0 \; 29]$.

Άρα το ελάχιστο της $f_1(x)$ είναι το (3,-5).

Πρόβλημα 2ο

Η αντικειμενική συνάρτηση του προβλήματος είναι : $f_2(x) = (x_1 - x_2)^2$

Η $f_1(x)$ είναι συνεχώς παραγωγίσιμη και η κλίση της είναι: $\nabla \mathrm{f}_2(x) = \begin{bmatrix} 2x_1 - 2x_2 \\ -2x_1 + 2x_2 \end{bmatrix}$

Ο Εσσιανός πίνακας της συνάρτησης είναι: $H_{f_2}(x) = \begin{bmatrix} 2 & -2 \\ -2 & 2 \end{bmatrix}$

Παρατηρούμε ότι εσσιανός πίνακας είναι συμμετρικός και οι ιδιοτιμές του είναι $\lambda_1 = 0 \; και \; \lambda_2 = 4$. Επομένως ο πίνακας είναι θετικά ημι-ορισμένος.

Οπότε καταλήγουμε στο συμπέρασμα ότι η $f_2(x)$ είναι κυρτή συνάρτηση.

Οι περιορισμοί του προβλήματος είναι :

- $g_1(x) = x_1 + 1 \leq 0$
- $g_2(x) = x_2 + 1 \leq 0$

Οι $g_i(x)$ είναι γραμμικές συναρτήσεις των x_1, x_2 άρα είναι κυρτές συναρτήσεις.

Εφόσον υπάρχουν σημεία του πεδίου ορισμού της $f_2(x)$ που ικανοποιούν τους περιορισμούς, το πρόβλημα μας είναι κυρτό και υπερ–αποτελούμενο.

Οι συνθήκες Karush-Kuhn-Tucker είναι οι ακόλουθες:

1. $\lambda_i \geq 0 \, , i = 1{,}2$

2. $\lambda_1 c_1(x^*) = 0 => \lambda_1(x_1^* + 1) = 0$

$\lambda_2 c_2(x^*) = 0 => \lambda_2(x_2^* + 1) = 0$

3. $2x_1^* - 2x_2^* + \lambda_1 = 0$

$-2x_1^* + 2x_2^* + \lambda_2 = 0$

Λύνοντας το παραπάνω σύστημα καταλήγουμε στις εξής λύσεις:

Λύση	$\boldsymbol{x_1}$	$\boldsymbol{x_2}$	$\boldsymbol{\lambda_1}$	$\boldsymbol{\lambda_2}$	**Περιορισμοί**
$1^η$	-1	-1	0	0	Ναι
$2^η$	z	z	0	0	Ναι $\forall \, z \leq -1$

Όπως βλέπουμε και από τον πίνακα, έχουμε ελάχιστο στο $[x_1 \; x_2] = [-1 - 1]$ με $[\lambda_1 \; \lambda_2] = [0 \; 0]$.

Παρατηρούμε επίσης ότι έχουμε άπειρα ελάχιστα της μορφής $[x_1 \; x_2] = [z \; z] \;\; \forall \, z \leq -1$. Αυτό οφείλεται στο ότι η συνάρτηση $f_2(x)$ είναι κυρτή και όχι γνησίως κυρτή όπως είχαμε στο προηγούμενο πρόβλημα.

B. Μέθοδος Φραγμού (Barrier Method)

Αφού μελετήσαμε την αντίστοιχη θεωρία του βιβλίου, υλοποιήσαμε στο Matlab την μέθοδο Φραγμού.

Η υλοποίηση του αλγόριθμου της μεθόδου βρίσκεται στο αρχείο με το όνομα **barrier_alg.m.**

Ακολούθως, για τη μέθοδο φραγμού, δημιουργήσαμε τις παρακάτω συναρτήσεις τις οποίες χρησιμοποιήσαμε στον κώδικα μας έτσι ώστε να διευκολυνθεί η υλοποίηση του αλγορίθμου:

- ***barrier_B.m:*** *Υπολογίζει κ' επιστρέφει την τιμή της συνάρτησης B(x) στο x.*
- ***barrier_grad_B.m****: Υπολογίζει κ' επιστρέφει το grad της συνάρτησης B(x) στο x.*
- ***grad_aux_barrier.m****: Υπολογίζει κ' επιστρέφει το grad της Φ(x,r) στο x.*

Επίσης, από προηγούμενες εργαστηριακές ασκήσεις χρησιμοποιήθηκε η μέθοδος μέγιστης καθόδου η οποία προσαρμόστηκε στα νέα δεδομένα της άσκησης. Η υλοποίηση της βρίσκεται στο αρχείο **steepest_descent.m**.

Εντός του αλγόριθμου καλούμε τη συνάρτηση **contour_f** έτσι ώστε να δημιουργήσουμε το γράφημα και να απεικονίσουμε το αρχικό και τελικό σημείο καθώς και όλα τα ενδιάμεσα.

Η αρχική τιμή του r_k είναι η $r_0 = 0.2\,\frac{f(x_0)}{B(x_0)}$ και σε κάθε επανάληψη η νέα τιμή του r_k υπολογίζεται από τον τύπο $r_{k+1} = 0.8\, r_k$. Το βήμα που χρησιμοποιείται στην μέθοδο μέγιστης καθόδου είναι σταθερό και ίσο με 0.0001 .

Για να τρέξουμε τον αλγόριθμο για το 1° πρόβλημα καλούμε τη συνάρτηση barrier_alg με τα εξής ορίσματα :

X = barrier_alg([x_{01} x_{02}], e , gamma, @f1, @grad_f1, @g1, @grad_g1).

Όπου ε = 0.001, $\gamma_\kappa = 0.0001$ και [x_{01} x_{02}] = το αρχικό σημείο.

Αντίστοιχα για το 2° πρόβλημα:

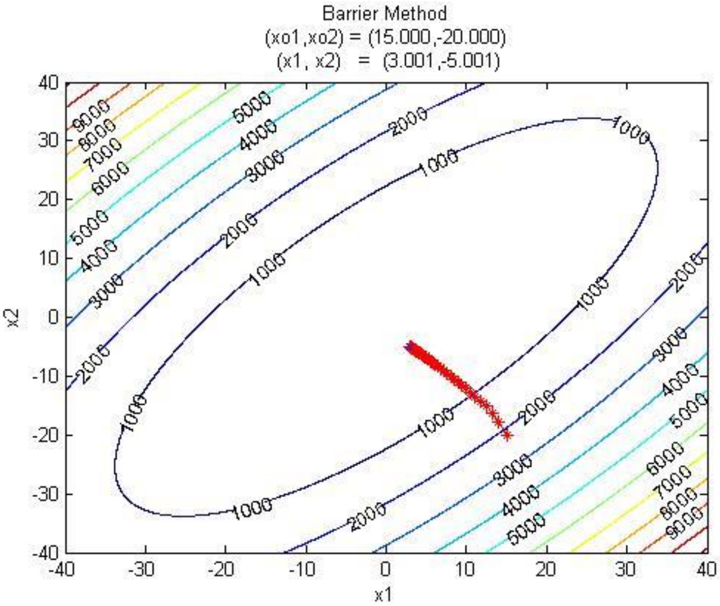
X = barrier_alg([x_{01} x_{02}], e , gamma, @f1, @grad_f2, @g2, @grad_g2).

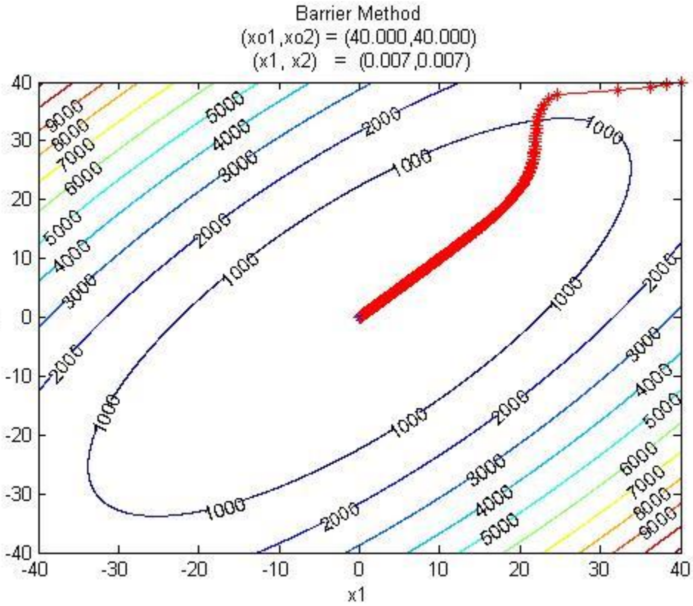
<u>Πρόβλημα 1ο</u>

Αρχικά, τρέχουμε τον αλγόριθμο με αρχικό σημείο [15 -20] το οποίο ικανοποιεί τους περιορισμούς.

Ακολούθως, τρέχουμε τον αλγόριθμο με αρχικό σημείο το οποίο δεν ικανοποιεί τους περιορισμούς.

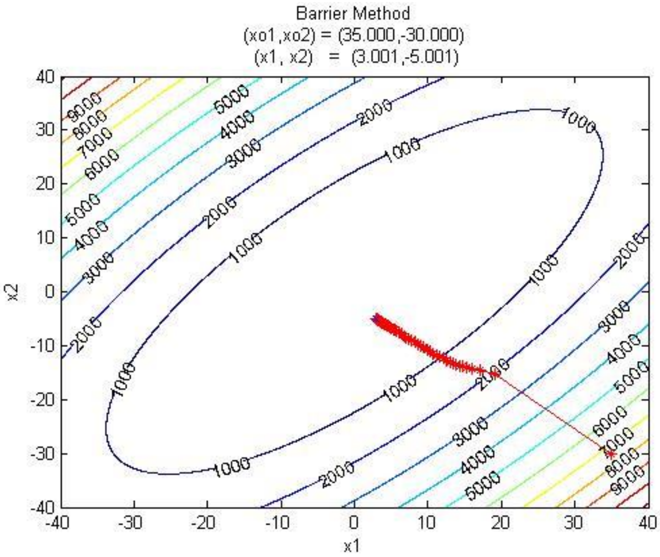
Τα διαγράμματα φαίνονται παρακάτω:





Από το 1ο διάγραμμα βλέπουμε ότι όταν το αρχικό σημείο τηρεί τους περιορισμούς τότε ο αλγόριθμος συγκλίνει στο ελάχιστο, δηλαδή στο (3,-5) το οποίο προβλέψαμε χρησιμοποιώντας το θεώρημα Karush-Kuhn-Tucker.

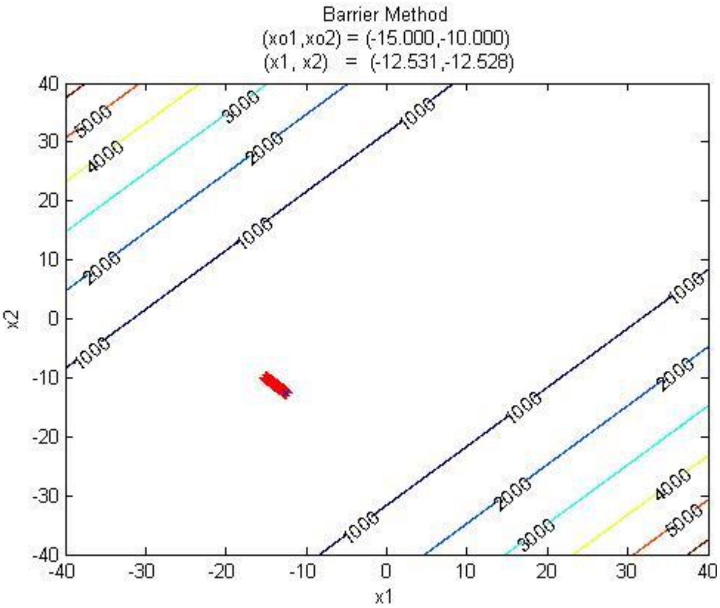
Στο 2ο διάγραμμα βλέπουμε ότι για αρχικό σημείο που βρίσκεται μακριά από τους περιορισμούς τότε ο αλγόριθμος αποκλίνει και καταλήγει στο ολικό ελάχιστο της συνάρτησης.

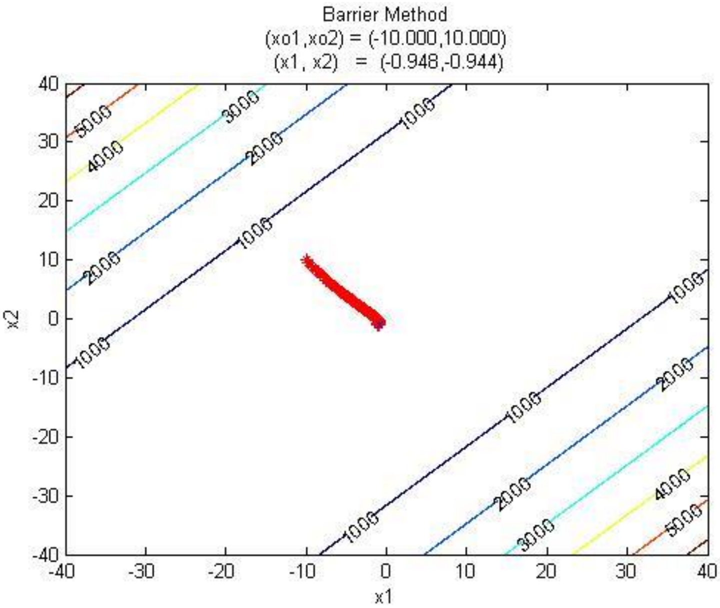


Ακολούθως τρέξαμε τον αλγόριθμο για ένα σημείο το οποίο ήταν εκτός των περιορισμών αλλά κοντά σε αυτούς. Όπως φαίνεται και από το διπλανό διάγραμμα ο αλγόριθμος σε αυτή την περίπτωση συγκλίνει στο ελάχιστο.

Πρόβλημα 2ο

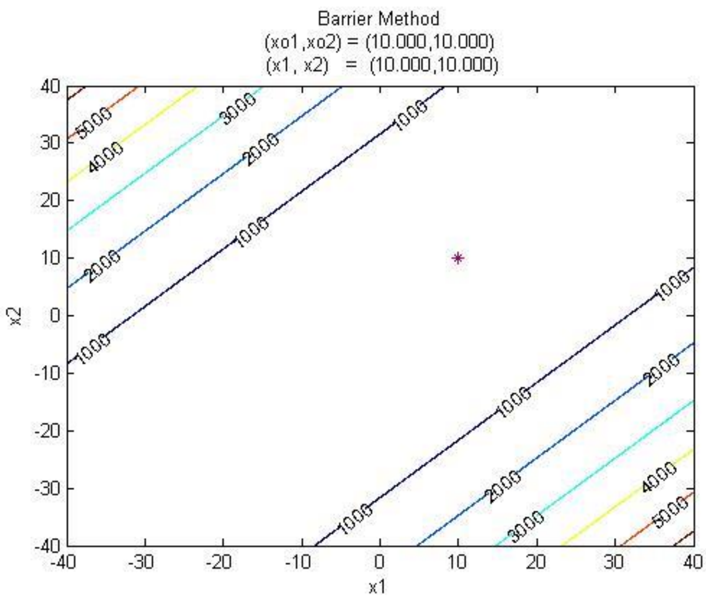
Αντίστοιχα με το 1ο πρόβλημα θα τρέξουμε τον αλγόριθμο για αρχικό σημείο εντός των περιορισμών και αρχικό σημείο εκτός των περιορισμών.





Όπως βλέπουμε από τα παραπάνω διαγράμματα ο αλγόριθμος συγκλίνει σε ελάχιστο της μορφής $[x_1\ x_2] = [z\ z] \quad \forall\, z \leq -1$ όταν το αρχικό σημείο είναι εντός των περιορισμών αλλά και όταν είναι εκτός των περιορισμών. Αυτό επαληθεύεται και από το θεώρημα Karush-Kuhn-Tucker.

Επίσης παρατηρήσαμε ότι λόγω της μορφής της $f_2(x)$ για οποιοδήποτε αρχικό σημείο $x_1 = x_2 = a$ η $f_2(a, a) = 0$ έτσι και το $r_0 = 0$ οπότε ο αλγόριθμος επιστρέφει το ίδιο το αρχικό σημείο, πράγμα το οποίο δεν παρατηρήθηκε στην $f_1(x)$. Αυτό φαίνεται και από το παρακάτω διάγραμμα για $x_1 = x_2 = 10$:



Γ. Μέθοδος Ποινής (Penalty Method)

Αφού μελετήσαμε την αντίστοιχη θεωρία του βιβλίου, υλοποιήσαμε στο Matlab την μέθοδο Ποινής.

Η υλοποίηση του αλγόριθμου της μεθόδου βρίσκεται στο αρχείο με το όνομα **penalty_alg.m.**

Ακολούθως, για τη μέθοδο ποινής, δημιουργήσαμε τις παρακάτω συναρτήσεις τις οποίες χρησιμοποιήσαμε στον κώδικα μας έτσι ώστε να διευκολυνθεί η υλοποίηση του αλγορίθμου:

- ***penalty_grad_f.m:*** *Υπολογίζει κ' επιστρέφει grad της συνάρτησης F(x,r) στο x.*
- ***penalty_grad_abs.m****: Υπολογίζει κ' επιστρέφει το grad του παράγοντα ποινής στο x.*

Επίσης, από προηγούμενες εργαστηριακές ασκήσεις χρησιμοποιήθηκε η μέθοδος μέγιστης καθόδου **steepest_descent.m**.

Εντός του αλγόριθμου καλούμε τη συνάρτηση **contour_f** έτσι ώστε να δημιουργήσουμε το γράφημα και να απεικονίσουμε το αρχικό και τελικό σημείο καθώς και όλα τα ενδιάμεσα.

Η αρχική τιμή του r_k ορίστηκε ως $r_0 = 3$ και σε κάθε επανάληψη η νέα τιμή του r_k υπολογίζεται από τον τύπο $r_{k+1} = 2\, r_k$. Το βήμα που χρησιμοποιείται στην μέθοδο μέγιστης καθόδου είναι σταθερό και ίσο με 0.00001.

Για να τρέξουμε τον αλγόριθμο για το 1^o πρόβλημα καλούμε τη συνάρτηση barrier_alg με τα εξής ορίσματα:

X = penalty_alg([x_{01} x_{02}], e , gamma, @f1, @grad_f1, @g1, @grad_g1).

Όπου ε = 0.001, $\gamma_\kappa =$ 0.00001 και [x_{01} x_{02}] = το αρχικό σημείο.

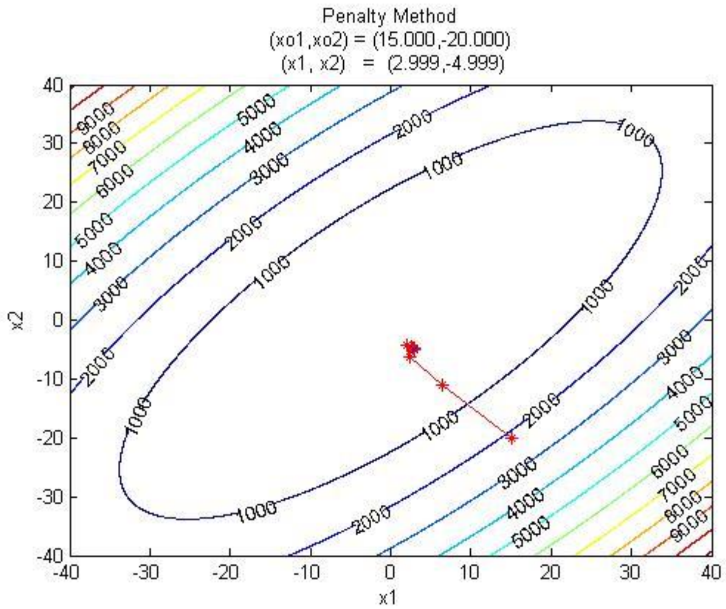
Αντίστοιχα για το 2^o πρόβλημα:

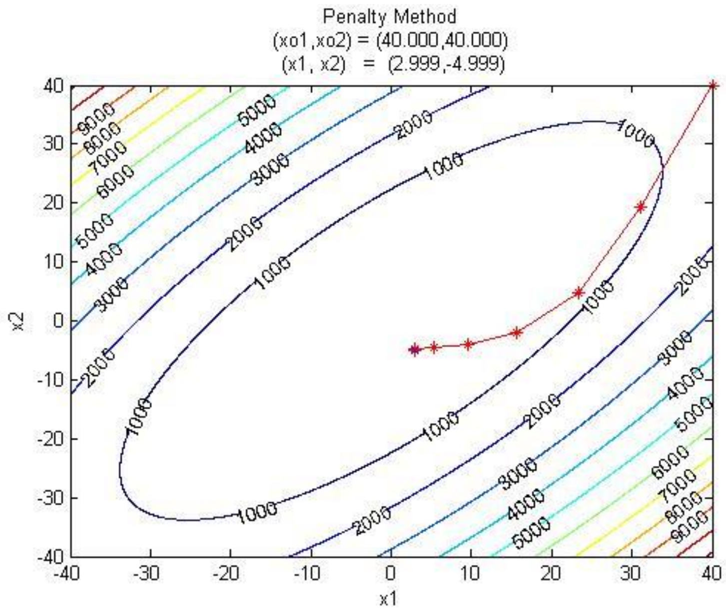
X = penalty_alg([x_{01} x_{02}], e , gamma, @f1, @grad_f2, @g2, @grad_g2).

Πρόβλημα 1ο

Τρέχουμε τον αλγόριθμο με αρχικό σημείο το οποίο ικανοποιεί τους περιορισμούς και ακολούθως, με αρχικό σημείο το οποίο δεν ικανοποιεί τους περιορισμούς.

Τα διαγράμματα που προκύπτουν για τις 2 περιπτώσεις αντίστοιχα είναι:

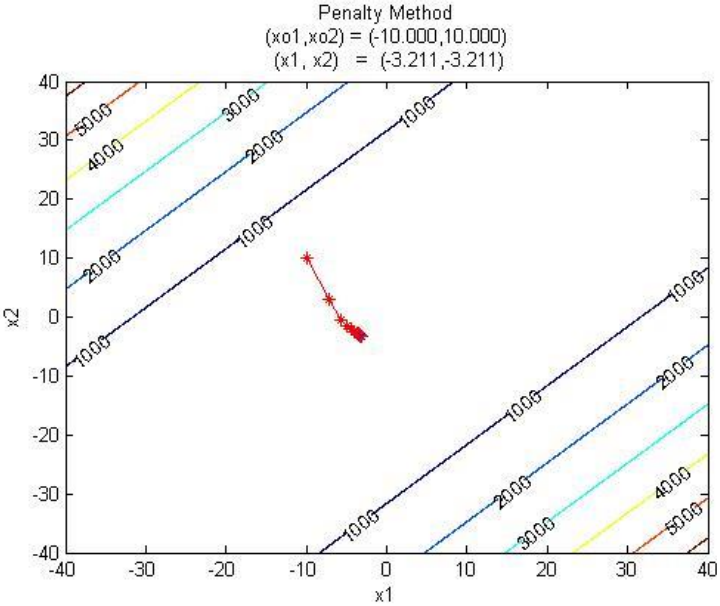


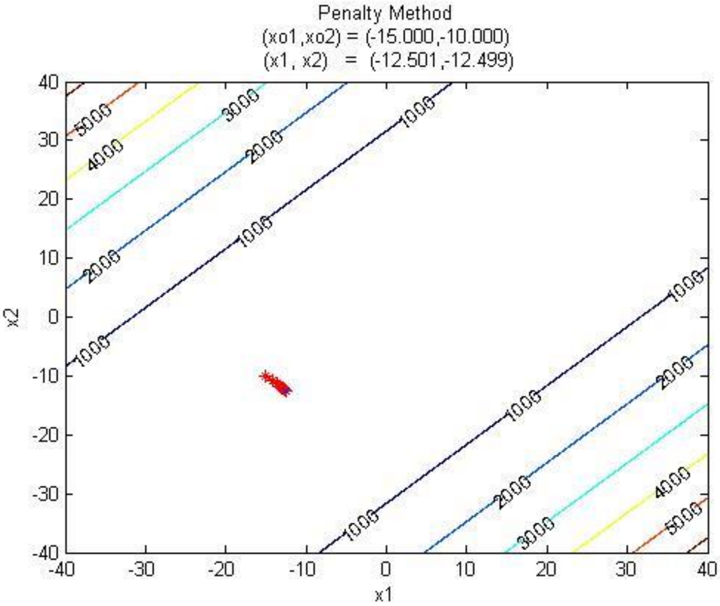


Και στις 2 περιπτώσεις παρατηρούμε ότι ο αλγόριθμος συγκλίνει ορθά στο ελάχιστο της συνάρτησης σε αντίθεση με την μέθοδο φραγμού με την οποία ο αλγόριθμος δε σύγκλινε για σημείο το οποίο βρισκόταν αρκετά έξω από τους περιορισμούς.

Πρόβλημα 2ο

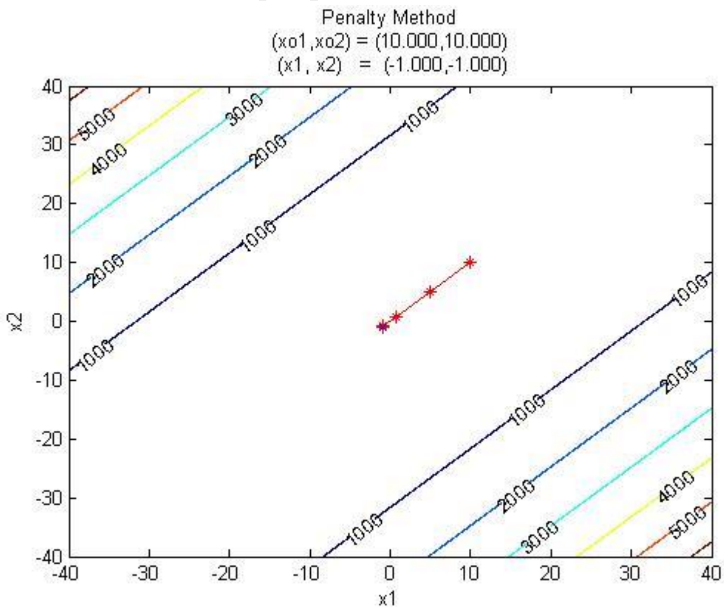
Αντίστοιχα με το 1ο πρόβλημα θα τρέξουμε τον αλγόριθμο για αρχικό σημείο εντός των περιορισμών και αρχικό σημείο εκτός των περιορισμών.





Όπως βλέπουμε από τα διαγράμματα, και στις 2 περιπτώσεις ο αλγόριθμος συγκλίνει σε ελάχιστο της μορφής $[x_1\ x_2] = [z\ z]\ \ \forall\, z \leq -1$.

Ακολούθως τρέξαμε τον αλγόριθμο για $x_1 = x_2 = 10$ και όπως φαίνεται και από το παρακάτω διάγραμμα, σε αντίθεση με την μέθοδο φραγμού, ο αλγόριθμος συγκλίνει σε ελάχιστο της μορφής $[x_1\ x_2] = [z\ z]\ \ \forall\, z \leq -1$ και συγκεκριμένα στο (-1, -1).



<u>Συμπερασματα – Παρατηρήσεις</u>

Οι μέθοδοι που αναπτύξαμε σε αυτή την εργαστηριακή άσκηση σύγκλιναν πάντα σε κάποιο ελάχιστο με εξαίρεση τη μέθοδο φραγμού για την $f_2(x)$ και για $x_1 = x_2 = a$ για τους λόγους που αναφέραμε πιο πάνω, αλλά ως γενικό συμπέρασμα μπορούμε να κρατήσουμε το ότι οι μέθοδοι σύγκλιναν πάντα σε κάποιο ελάχιστο.

Στις περιπτώσεις που το αρχικό σημείο ήταν έξω από τους περιορισμούς η μέθοδος φραγμού σύγκλινε στο ολικό ελάχιστο της συνάρτησης σε αντίθεση με τη μέθοδο ποινής στην οποία δεν συνέβαινε κάτι τέτοιο και σύγκλινε στο ελάχιστο. Οπότε για τη μέθοδο φραγμού πρέπει να λαμβάνονται υπόψη οι περιορισμοί έτσι ώστε να ρυθμίζεται σωστά για να παίρνουμε το επιθυμητό αποτέλεσμα, σε αντίθεση με τη μέθοδο ποινής η οποία έχει το πλεονέκτημα να συγκλίνει στο ελάχιστο για οποιοδήποτε αρχικό σημείο. Αυτό φαίνεται και από την αντίστοιχη θεωρία των μεθόδων.

Επίσης παρατηρήθηκε ότι για αρχικό σημείο εντός των περιορισμών, η μέθοδος φραγμού καθ’ όλη τη διάρκεια αναζήτησης μένει εντός των περιορισμών σε αντίθεση με τη μέθοδο ποινής η οποία μπορεί να κινηθεί και εκτός των περιορισμών, αλλά μόλις γίνει αυτό επανέρχεται εντός των περιορισμών.

Όσον αφορά την ταχύτητα των μεθόδων, δεν μπορούμε να βγάλουμε ασφαλή συμπεράσματα διότι ανάλογα με το πρόβλημα και τις παραμέτρους που έχουμε για το κάθε πρόβλημα, οι μέθοδοι συμπεριφέρονται αλλιώς. Επίσης μεγάλη σημασία έχει η μέθοδος ελαχιστοποίησης που επιλέχθηκε. Θα μπορούσαμε να βελτιώσουμε την ταχύτητα χρησιμοποιώντας άλλες μεθόδους που υλοποιήσαμε σε προηγούμενες εργαστηριακές ασκήσεις και ακόμη θα μπορούσαμε να χρησιμοποιήσουμε εσωτερική βελτιστοποίηση για το βήμα γ_κ που χρησιμοποιήθηκε στην ελαχιστοποίηση με τη μέθοδο μέγιστης καθόδου.

Τέλος, καταλήγουμε στο συμπέρασμα ότι , είναι καταλληλότερο να επιλέγουμε τη μέθοδο που θα χρησιμοποιήσουμε ανάλογα με το πρόβλημα που έχουμε και τους περιορισμούς αλλά λαμβάνοντας υπόψη και τις παραμέτρους που της μεθόδου που θα χρησιμοποιήσουμε.