ΑΡΙΣΤΟΤΕΛΕΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΟΝΙΚΗΣ
ΤΗΜΜΥ ΤΟΜΕΑΣ ΗΛΕΚΤΡΟΝΙΚΗΣ ΚΑΙ ΥΠΟΛΟΓΙΣΤΩΝ

`

ΤΕΧΝΙΚΕΣ ΒΕΛΤΙΣΤΟΠΟΙΗΣΗΣ

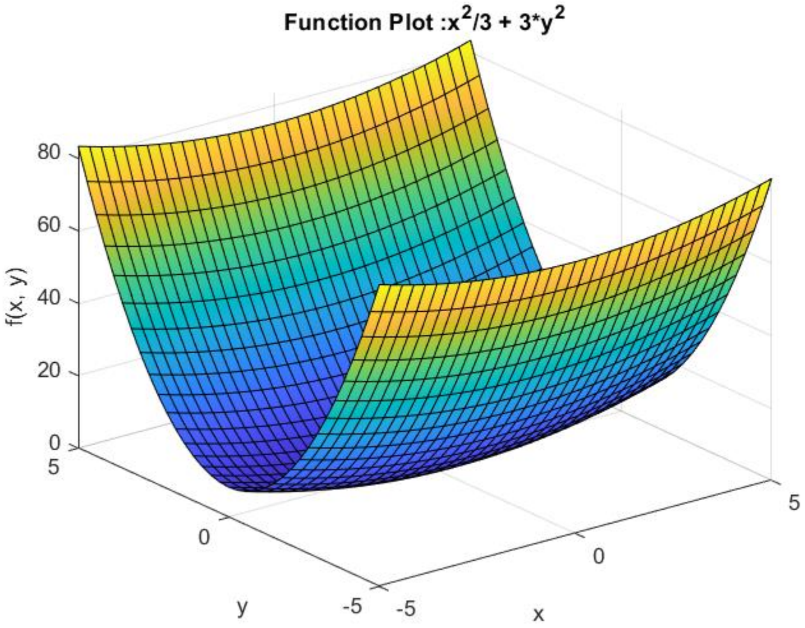
3Η ΕΡΓΑΣΙΑ

Ιωάννης Δεϊρμεντζόγλου 10015

ΑΕΜ 10015 Email deirmentz@ece.auth.gr

Εισαγωγή

Στόχος της τρίτης εργασίας είναι ελαχιστοποίηση μιας συνάρτησης 2 μεταβλητών με περιορισμούς. Πιο συγκεκριμένα, συνάρτηση προς ελαχιστοποίηση είναι η $f(x_1, x_2) = \frac{1}{3}x_1^2 + 3x_2^2$ f : $R^2 \rightarrow R$, η οποία απεικονίζεται και στο σχήμα 1. Αρχικά, μελετάται η συμπεριφορά της μεθόδου της μέγιστης καθόδου, της προηγούμενης εργασίας με τέσσερα διαφορετικά βήματα γ_κ και τυχαίο σημείο εκκίνησης του αλγορίθμου, 2) η δοκιμή της Μεθόδου Μέγιστης Καθόδου με προβολή σε τρεις διαφορετικές περιπτώσεις (διαφορετικά γ_κ, s_κ, σημεία εκκίνησης).



Από την τρισδιάστατη αναπαράσταση της συνάρτησης παρατηρούμε πώς το γεωμετρικό της σχήμα είναι ένα ελλειπτικό παραβολοειδές ενώ μπορούμε να δούμε και πως εμφανίζει ελάχιστο στο (0,0).

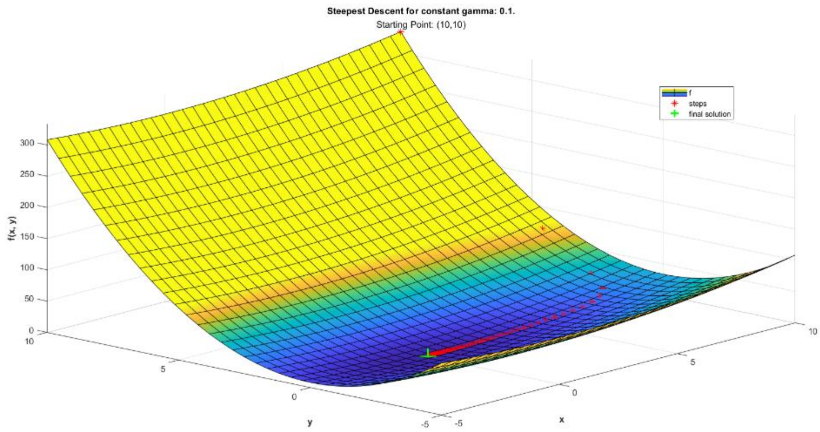
Θέμα 1 : Μέθοδος Μέγιστης Καθόδου χωρίς περιορισμούς

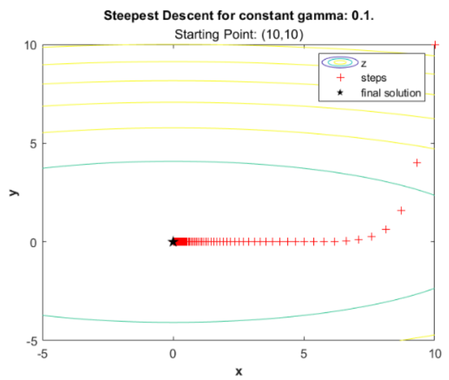
Ο κώδικας βρίσκεται στο αρχείο Task1.m .

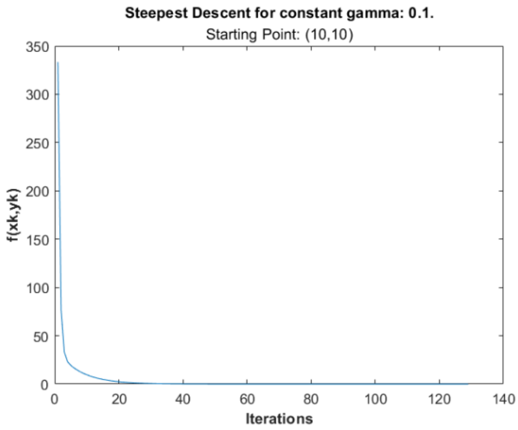
Στο πρώτο θέμα χρησιμοποιήθηκε η μέθοδος της μέγιστης καθόδου της προηγούμενης εργασίας για ε=0.001και βήματα i) γκ=0.1 ii) γκ=0.3 iii) γκ=3 iv) γκ=5 και ως αρχικό σημείο επιλέχθηκε το (10,10).

Παρακάτω , παρουσιάζονται διαγράμματα που δείχνουν τις θέσεις των σημείων μετα από κάθε επανάληψή και πως αλλάζει η τιμή της συνάρτησης συναρτήσει του αριθμού των επαναλήψεων. Σημειώνεται πως για την περιπτωση οπού $γ_κ$=3 ο αλγόριθμος σταμάτησε εξαναγκασμένα στις 100 επαναλήψεις για να μπορέσουν να παρουσιαστούν τα αντίστοιχα διαγράμματα .

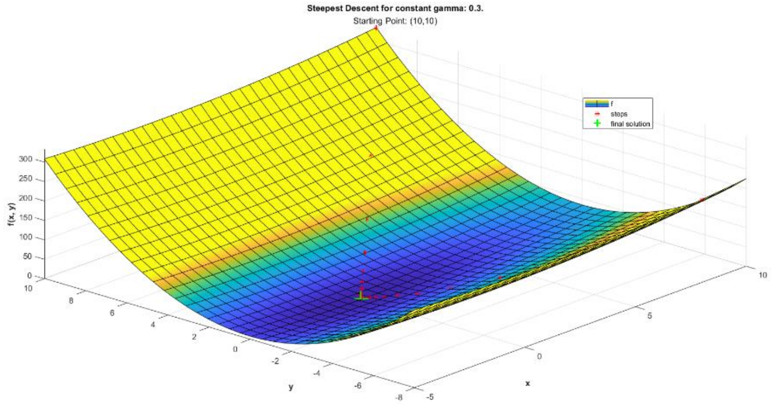
- **Steepest Descent γκ=0.1**

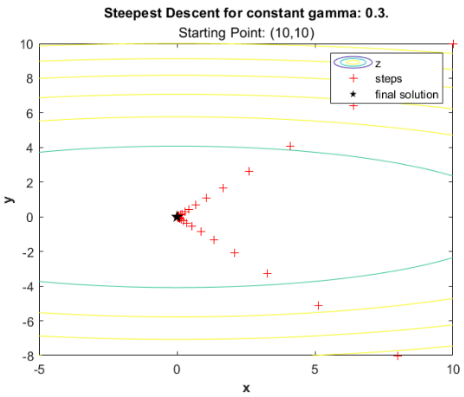


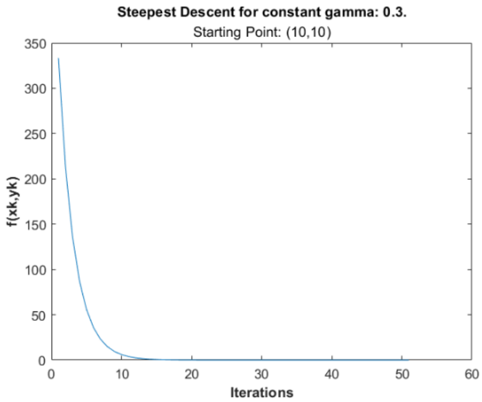




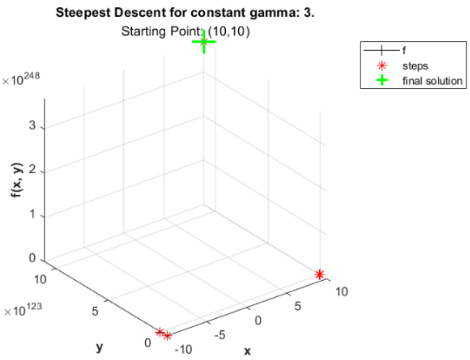
- **Steepest Descent γκ=0.3**

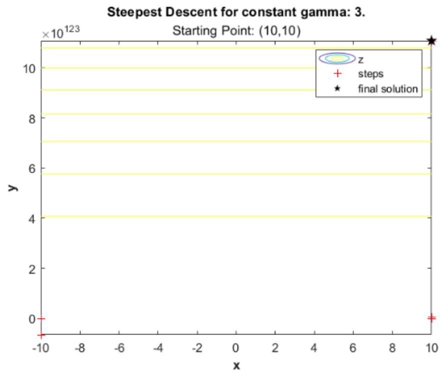


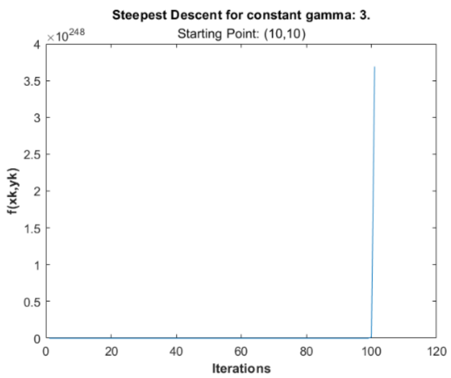




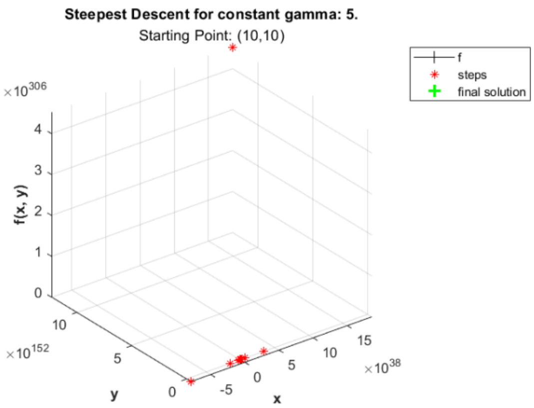
- **Steepest Descent γκ=3**

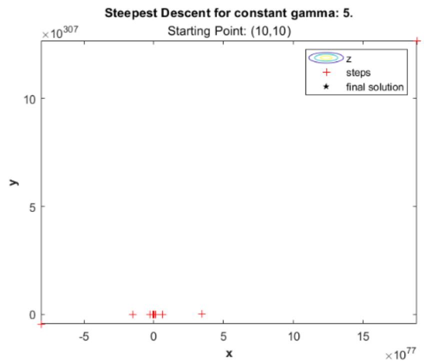


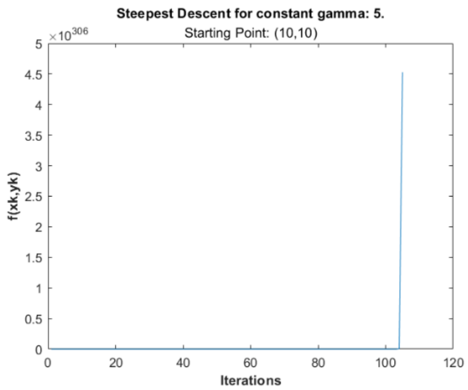




- **Steepest Descent γκ=5**







Παρατηρείται ότι στις περιπτώσεις (i) και (ii) η μέθοδος κατάφερε να βρει το ελάχιστο με πολύ καλή ακρίβεια, σε αντίθεση με τις περιπτώσεις (iii) και (iv) ο αλγόριθμος δεν συγκλίνει . Αυτό συμβαίνει γιατί η επιλογή του γκ πολλές φορές έχει περιορισμούς οι οποίοι καθορίζονται από την εκάστοτε συνάρτηση που τίθεται για ελαχιστοποίηση. Παρακάτω παρουσιάζεται η μαθηματική απόδειξη των περιορισμών που θέτει η συνάρτηση $f(x_1, x_2) = \frac{1}{3}x_1^2 + 3x_2^2$, f : $R^2 \rightarrow R$.

Απόδειξη

Αρχικά υπολογίζεται το gradient της συνάρτησης ως εξής :

$$\nabla f(x_1, x_2) = \begin{bmatrix} \frac{\partial f(x1,x2)}{\partial x1} \\ \frac{\partial f(x1,x2)}{\partial x2} \end{bmatrix} = \begin{bmatrix} \frac{2}{3}x_1 \\ 6x_2 \end{bmatrix}$$

Ουσιαστικά για να συγκλίνει ο αλγόριθμος πρέπει να ισχύει $\frac{|x_{k+1}|}{|x_k|} < 1$ (1) ,για να συγκλίνει δηλαδή η ακολουθία θέλουμε κάθε τιμή του επόμενου βήματος να είναι μικρότερη από την προηγούμενη, ώστε να συγκλίνει ο αλγόριθμος στο ελάχιστο που είναι το (0, 0) , κάτι το οποίο είναι προφανές καθώς η αντικειμενική συνάρτηση είναι ένα τετραγωνικό άθροισμα των x1 , x2 . (λυμένη άσκηση 5.5.2, σελίδα 169).

Επομένως , $x_{k+1} = x_k - \gamma_k * \nabla f(x_{1k}, x_{2k}) = x_k - \begin{bmatrix} \frac{2}{3} * \gamma_k * x_{1k} \\ 6 * \gamma_k * x_{2k} \end{bmatrix} = \begin{bmatrix} 1 - \frac{2}{3} * \gamma_k \\ 1 - 6 * \gamma_k \end{bmatrix} * X_k$

Για να συγκλίνει ο αλγόριθμος πρέπει:

- $|1-2/3*\gamma\kappa|<1 \Rightarrow 0<\gamma\kappa<3$
- $|1-6*\gamma\kappa|<1 \Rightarrow 0<\gamma\kappa<1/3$

Τελικά $\mathbf{0<\gamma_k<1/3=0.333}$.
Έτσι εξηγείται γιατί συγκλίνει ο αλγόριθμος στις πρώτες δύο περιπτώσεις (0,1 , 0.3 < 1/3) και γιατί δεν συγκλίνει στις τελευταίες δύο περιπτώσεις (3, 5 > 1/3).

Θέμα 2 : Μέθοδος Μέγιστης Καθόδου με προβολή

Ο κώδικας βρίσκεται στο αρχείο με τίτλο Task2.m .

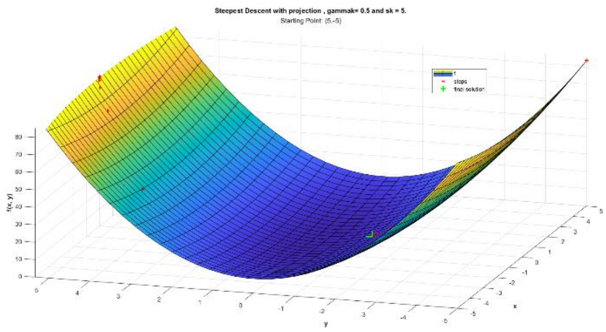
Η μέθοδος της μέγιστης καθόδου με προβολή λειτουργεί ως εξής: Ξεκινά με ένα εφικτό σημείο και ακολουθεί τον αλγόριθμος της μέγιστης καθόδου. Αν το νέο σημείο που προκύπτει είναι και αυτό εφικτό τότε συνεχίζει κανονικά με τον ίδιο αλγόριθμο. Ωστόσο, αν το νέο σημείο δεν είναι εφικτό τότε βρίσκει την προβολή του στο X (στο κυρτό σύνολο) και επαναλαμβάνει την ίδια διαδικασία. Για να συγκλίνει ο αλγόριθμος στο (0,0), που είναι μέσα στο σύνολο X στην περίπτωση μας, θα πρέπει τουλάχιστον στην τελευταία επανάληψη το σημείο να είναι μέσα στο σύνολο X, δηλαδή να είναι εφικτό. Τότε όμως για την έκφραση του x_{k+1} ισχύει:

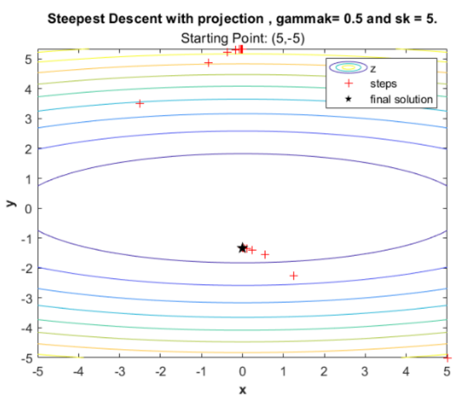
$x_{k+1} = x_k + \gamma_k * (\bar{x} - x_k)$, όπου $\bar{x} = \Pr_X\{ x_k - s_k * \nabla f(xk)\}$.

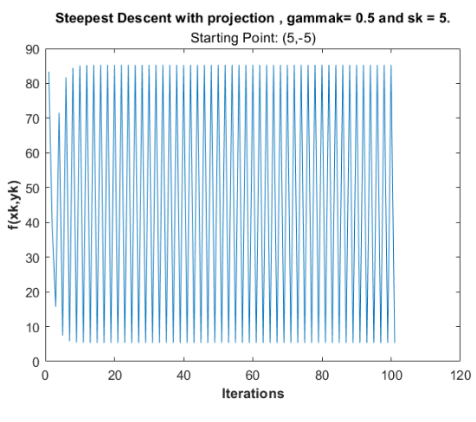
Όμως αν το x_k είναι εφικτό σημείο τότε $\bar{x} = xk - sk * \nabla f(xk)$

Τότε , από τις παραπάνω σχέσεις θα έχω $x_{k+1} = x_k - \gamma_k * s_k * \nabla f(x_k)$. Βγαίνει λοιπόν το συμπέρασμα πως το x_{k+1} υπολογίζεται από τον αλγόριθμο της μέγιστης καθόδου με $\gamma_k' = \gamma_k * s_k$ και για το γ_k' θα ισχύουν οι περιορισμοί του θέματος 1 ($0 < \gamma_k' < 1/3$).

Στο Θέμα 2 χρησιμοποιείται την μέθοδο μέγιστης καθόδου με προβολή με **sk = 5, γk = 0.5, ε = 0.01** και αρχικό σημείο εκκίνησης του αλγορίθμου το (5,-5). Λαμβάνοντας υπόψιν και τα παραπάνω θα είναι **γ'k = γk *sk = 0.5*5 = 2.5** δηλαδή **γ'κ > 1/3** και θα αναμένουμε να ο αλγοριθμος να αποκλίνει με αυτές τις αρχικές συνθήκες.





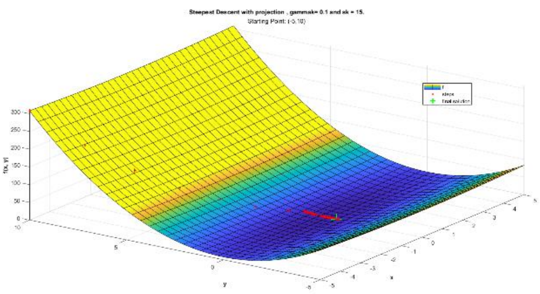


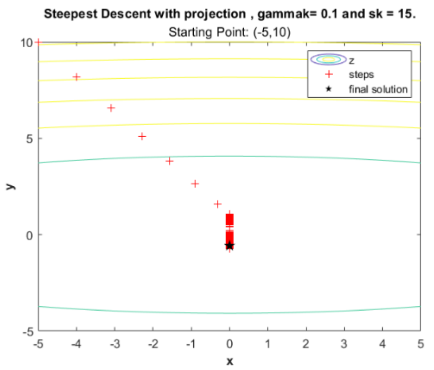
Παρατηρείται ότι σε αντίθεση με το πρώτο θέμα για μεγάλο βήμα, αυτή η μέθοδος δεν ξεφεύγει στο άπειρο, αλλά μένει μέσα στους αρχικούς περιορισμούς $-10 < x1 < 5$ και $-8 < x2 < 12$. Παρόλα αυτά δεν καταφέρνει να συγκλίνει στο ελάχιστο, αλλά ταλαντώνεται σε δύο τιμές επ'άπειρον (οι μέγιστες επαναλήψεις έχουν τεθεί 100).

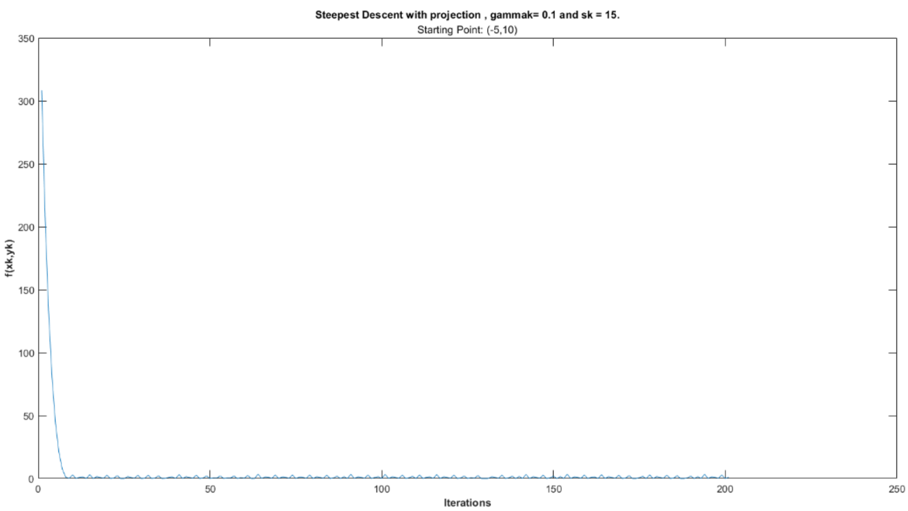
Θέμα 3 : Μέθοδος Μέγιστης Καθόδου με προβολή

Ο κώδικας βρίσκεται στο αρχείο με τίτλο Task3.m .

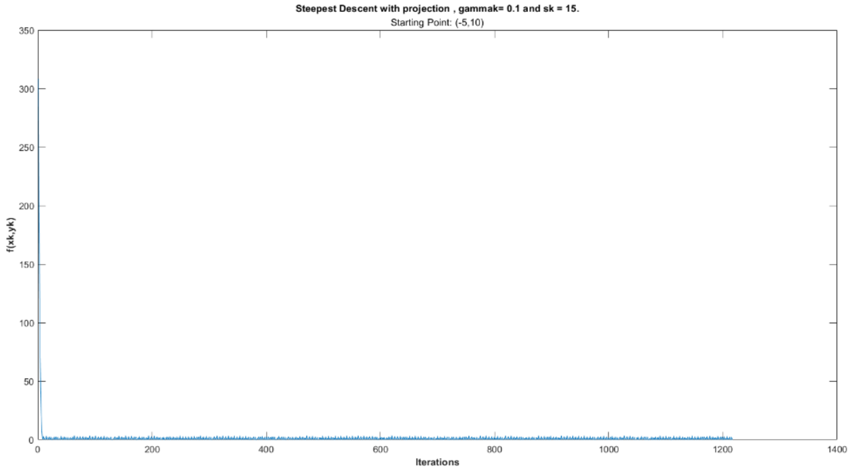
Στο Θέμα 3 χρησιμοποιούμε την μέθοδο μέγιστης καθόδου με προβολή με **sk = 15, γk = 0.1, ε = 0.01** και αρχικό σημείο εκκίνησης του αλγορίθμου το **(-5,10)**. Λαμβάνοντας υπόψιν ξανά τα παραπάνω θα είναι γ'k = γk * sk = 0.1*15 = 1.5 δηλαδή **γ'k > 1/3** οπότε αναμένεται η μέθοδος να μην συγκλίνει στο ελάχιστο για τις συγκεκριμένες αρχικές συνθήκες. Ο μέγιστος αριθμός επαναλήψεων έχει τεθεί στο 200 .



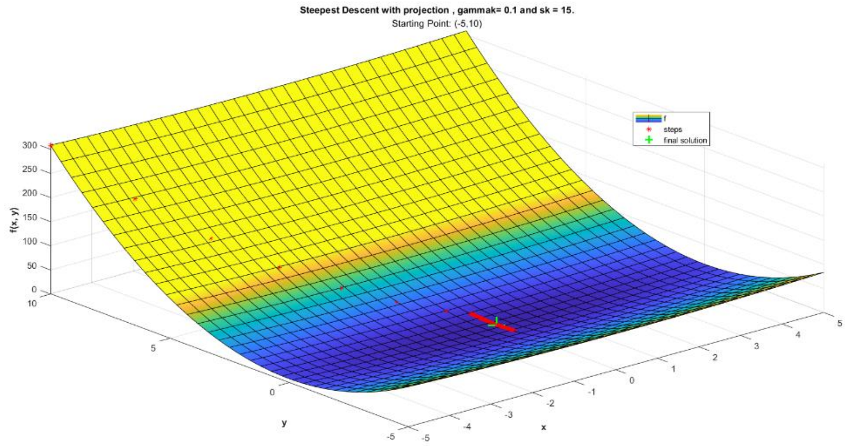


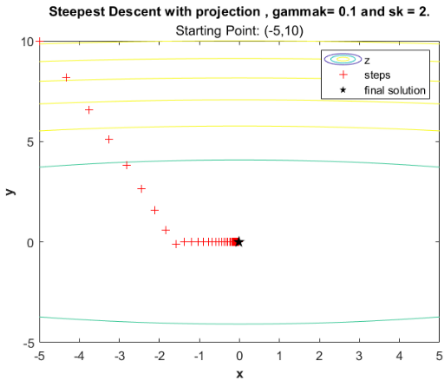


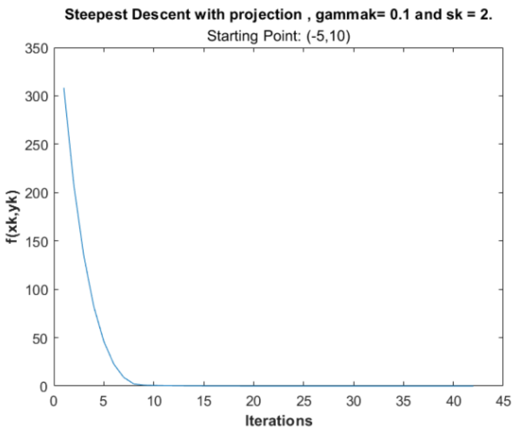
Χωρίς μέγιστο αριθμό επαναλήψεων ο αλγοριθμος τερματίζει στις 1216 επαναλήψεις και παρατίθεται το παρακάτω σχήμα :



Σε αυτήν την περίπτωση παρατηρείται πως λόγω του μεγάλου βήματος ο αλγοριθμος μετα από λίγες επαναλήψεις πολύ γρήγορα πλησιάζει σε μια περιοχή κοντά στο ολικό ελάχιστο –(0,0) , και έπειτα ταλαντώνεται γύρω από αυτή, και εν τέλει συγκλίνει (τερματίζει από μόνος του). Ένας τρόπος να λυθεί αυτό ίσως είναι η μείωση του s_k έτσι ώστε $s_k*\gamma_k < 1/3$. Αν τεθει το $s_k = 2$ έχουμε για $\gamma_k = 0.1$:





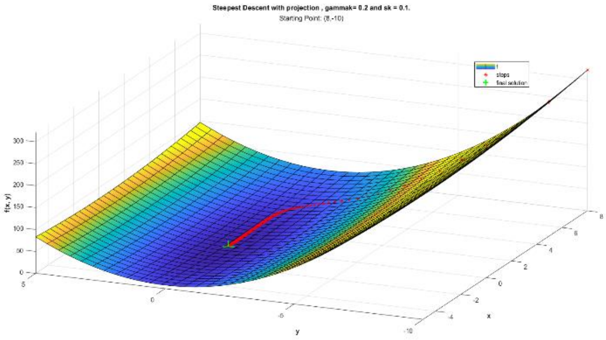


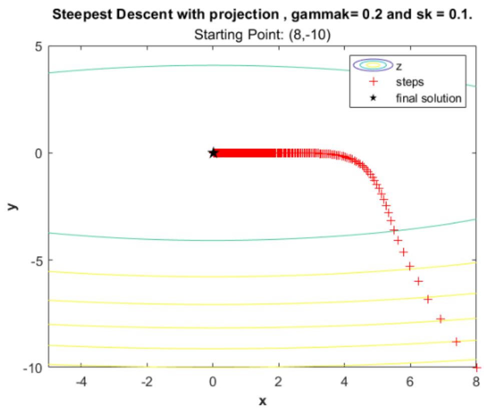
Θέμα 4 : Μέθοδος Μέγιστης Καθόδου με προβολή

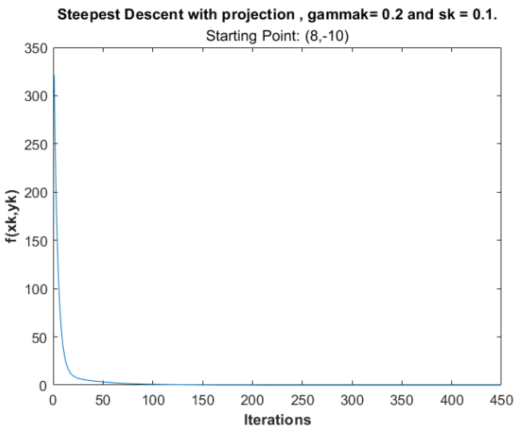
Ο κώδικας βρίσκεται στο αρχείο με τίτλο Task4.m .

Στο Θέμα 4 χρησιμοποιούμε την μέθοδο μέγιστης καθόδου με προβολή με sk = 0.1, γk = 0.2, ε = 0.01 και αρχικό σημείο εκκίνησης του αλγορίθμου το (8,-10). Λαμβάνοντας υπόψιν και τα παραπάνω θα είναι γ'k = γk *sk = 0.1*0.2 = 0.02. Πριν τρέξει ο αλγοριθμος για τις παραπάνω αρχικές συνθήκες παρατηρείται πως η τιμή x1 = 8 και x2 = -10 δεν ανήκουν στο κυρτό σύνολο των περιορισμών που ορίστηκε παραπάνω. Η πρώτη κλήση του αλγορίθμου θα χρησιμοποιήσει την προβολή και στο χ1 που θα γίνει 5 αλλά και στο χ2 που θα γίνει -8 οπότε το αρχικό σημείο θα γίνει το (5,-8) οδηγώντας δηλαδή τον αλγόριθμο μέσα στο διάστημα των περιορισμών έστω και αν αυτό είναι στο όριο του.

Παρακατω παρατίθενται τα διαγράμματα







Το αρχικό σημείο είναι μη εφικτό, αλλά η μέθοδος το φέρνει το σημείο μέσα στους περιορισμούς μετά από μερικές επαναλήψεις. Παρατηρείται πως επιτυγχάνεται σύγκλιση λόγω της επιλογής μικρού βήματος (0.02) .

Συμπεράσματα

Από τις παραπάνω ασκήσεις εξάγεται το συμπέρασμα πως ο αλγόριθμος της μέγιστης καθόδου με προβολή μπορεί να συγκλίνει στο ελάχιστο μιας συνάρτησης με αρκετά λιγότερα βήματα, αν γίνει καλή επιλογή των παραμέτρων $\gamma k\ \kappa\alpha\iota\ sk$, διότι σε αντίθετη περίπτωση μπορεί είτε να μην συγκλίνει είτε να χρειάζεται πολύ μεγάλο αριθμό επαναλήψεων . Το γινόμενο τους πρέπει να ικανοποιεί τυχόν περιορισμούς που θέτει η συνάρτηση στην περίπτωση της μέγιστης καθόδου. Επιπλέον, θα πρέπει να γνωρίζουμε το σύνολο Χ μέσα στο οποίο θα γίνει η αναζήτηση του ελαχίστου. Αν οι παραπάνω προϋποθέσεις ισχύουν ο αλγόριθμος της μέγιστης καθόδου με προβολή είναι αποδοτικότερος του αντίστοιχου για την ίδια συνάρτηση χωρίς περιορισμούς.