ΤΕΧΝΙΚΕΣ ΒΕΛΤΙΣΤΟΠΟΙΗΣΗΣ, 3Η ΕΡΓΑΣΙΑ

Γιτόπουλος Γιώργος, 9344, ΤΗΜΜΥ ΑΠΘ, 2020

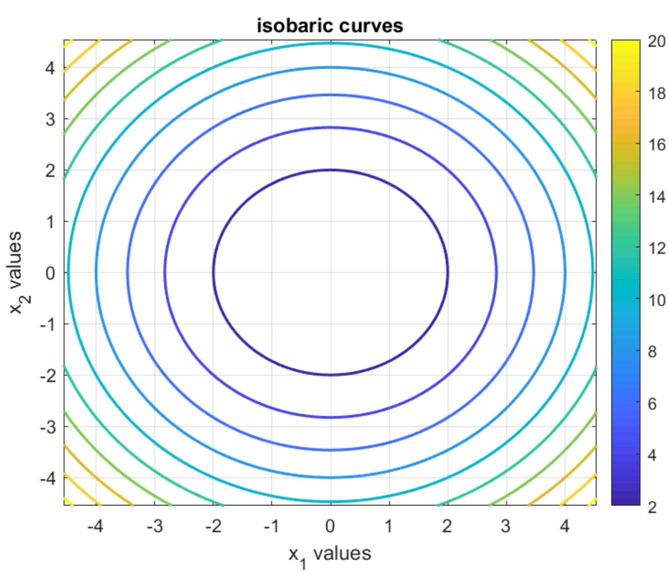
georgito@ece.auth.gr

Σε αυτή την εργασία, θα ασχοληθούμε με τη Μέθοδο Μέγιστης Καθόδου χωρίς περιορισμούς και με μια επέκτασή της, τη Μέθοδο Μέγιστης Καθόδου με προβολή. Η Μέθοδος Μέγιστης Καθόδου αποτελεί έναν αλγόριθμο τοπικής ελαχιστοποίησης παραγωγίσιμων συναρτήσεων πολλών μεταβλητών. Μάλιστα, θα ασχοληθούμε με κυρτές συναρτήσεις, για τις οποίες η ελαχιστοποίηση είναι ολική. Η βασική ιδέα της μεθόδου είναι η εξής: μέσω διαδοχικών βημάτων στην αντίθετη κατεύθυνση του διανύσματος κλίσης της συνάρτησης στο εκάστοτε σημείο, μπορεί να επιτευχθεί η ελαχιστοποίηση της συνάρτησης ($x_{\kappa+1} = x_\kappa - \gamma_\kappa \nabla f(x_\kappa),\ \gamma_\kappa > 0$). Αυτό μπορεί να γίνει εύκολα κατανοητό, αν θυμηθούμε πως για μία κυρτή συνάρτηση, η νόρμα της κλίσης μηδενίζεται μόνο στο ολικό ελάχιστο. Έτσι λοιπόν, ο αλγόριθμος επιδιώκει να μηδενίσει τη νόρμα της κλίσης, ή να την περιορίσει σε ένα διάστημα ακρίβειας της επιλογής μας.

Ο αλγόριθμος υλοποιήθηκε (με και χωρίς προβολή) σε Matlab. Σύντομη περιγραφή του κώδικα, γίνεται στο τέλος της αναφοράς. Για τη μελέτη της μεθόδου θα χρησιμοποιήσουμε την τετραγωνική συνάρτηση:

$$f : \mathbb{R}^2 \to \mathbb{R},\ f(x) = \frac{1}{2}x_1^2 + \frac{1}{2}x_2^2$$

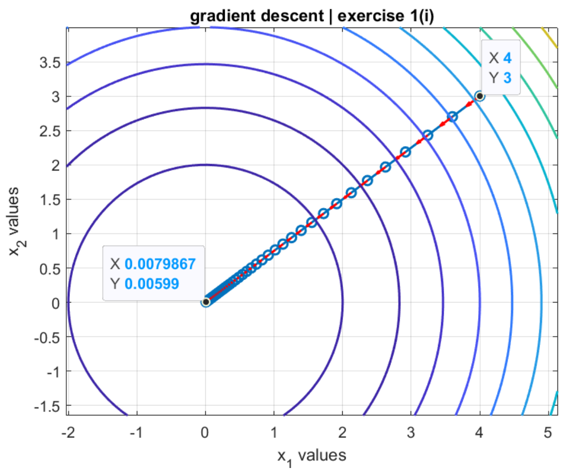
Η συνάρτηση αυτή παρουσιάζει μεγάλο ενδιαφέρον και αποτελεί χρήσιμο μαθηματικό εργαλείο. Οι ισοβαρείς καμπύλες της έχουν τη μορφή: $\frac{1}{2}x_1^2 + \frac{1}{2}x_2^2 = c$ ή ισοδύναμα $x_1^2 + x_2^2 = 2c$ και παριστάνουν ομόκεντρους κύκλους ακτίνας $\sqrt{2c}$ με κέντρο το $O(0,\ 0)$. Είναι προφανές πως το ολικό ελάχιστο της f παρουσιάζεται για $x_1 = 0,\ x_2 = 0$ με $f(0,0) = 0$. Η τέλεια συμμετρία της f (όλες οι ισοβαρείς ομόκεντροι κύκλοι) μας δίνει το πλεονέκτημα, αφού μελετήσουμε ένα διάνυσμα x να μπορούμε να επεκτείνουμε τα συμπεράσματά μας για όλα τα σημεία της ισοβαρούς στην οποία αυτό ανήκει. Επίσης, το γεγονός ότι η συνάρτηση είναι γνήσια θετική για κάθε $x \neq (0,0)$ και ελαχιστοποιείται - μηδενίζεται μόνο για $x = (0,0)$, σε συνδυασμό με την προηγούμενη ιδιότητα, την καθιστούν, όπως είπαμε χρήσιμο εργαλείο σε προβλήματα όπως η ελαχιστοποίηση και η ευστάθεια συναρτήσεων. Χαρακτηριστικό παράδειγμα αποτελεί η ευρεία χρήση της στα Συστήματα Αυτομάτου Ελέγχου ως υποψήφια συνάρτηση Lyapunov για την απόδειξη της ευστάθειας ενός συστήματος. Τέλος, ισχύει $\nabla f(x) = x$, κάτι που απλοποιεί σημαντικά τους υπολογισμούς σε προβλήματα που χρησιμοποιούν το διάνυσμα κλίσης, όπως το πρόβλημα της συγκεκριμένης εργασίας.

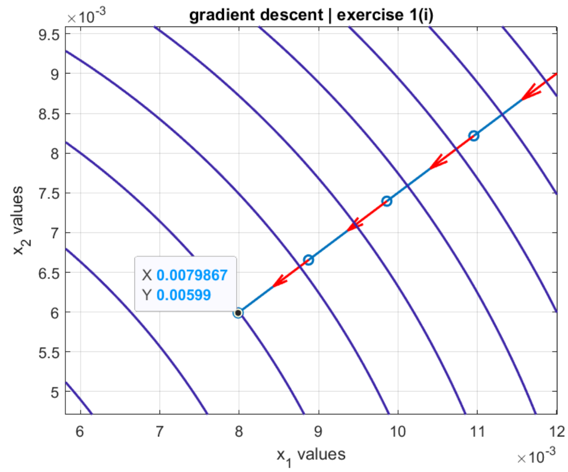


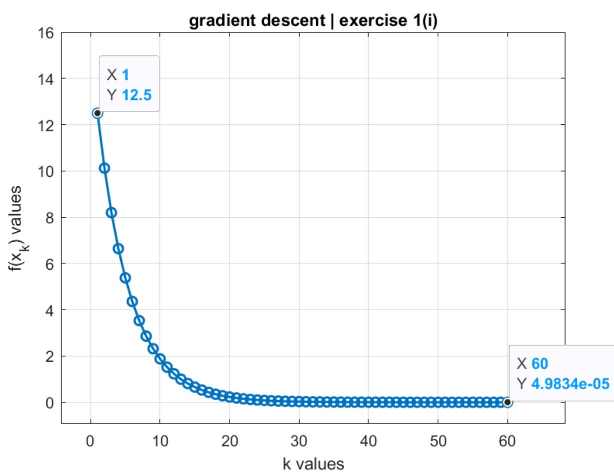
Παρακάτω θα παρουσιάσουμε τα αποτελέσματα των πειραμάτων που έγιναν στα πλαίσια των τεσσάρων θεμάτων της εργασίας.

Θέμα 1

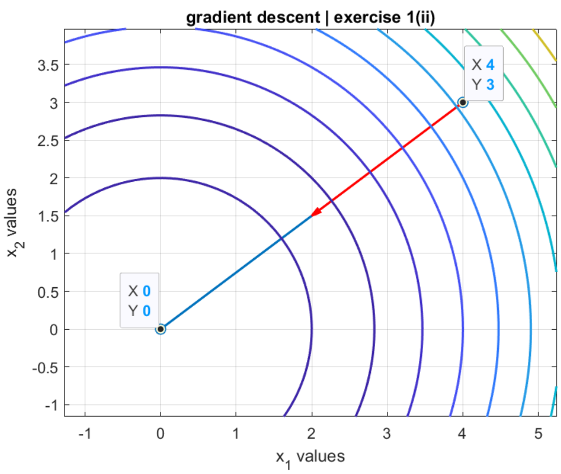
Υλοποιήσαμε τη μέθοδο της μέγιστης καθόδου (χωρίς περιορισμούς) με ακρίβεια $\varepsilon = 0.01$, σημείο εκκίνησης (4, 3) και μελετήσαμε τη λειτουργία του αλγορίθμου για τις διαφορετικές τιμές του γ_κ: (i) 0.1, (ii) 1, (iii) 2, (iv) 10.

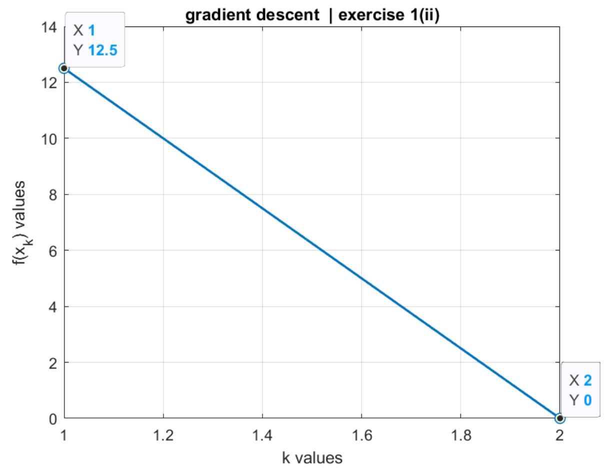




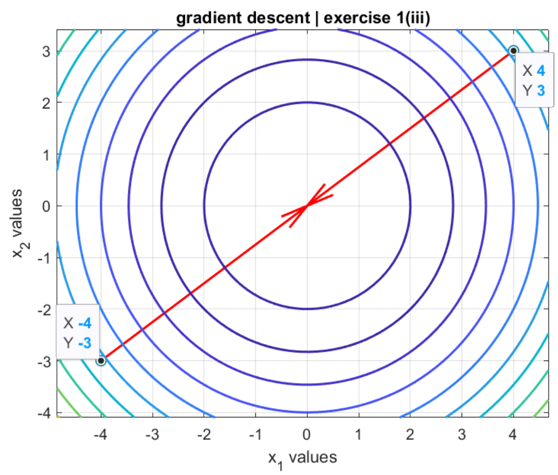


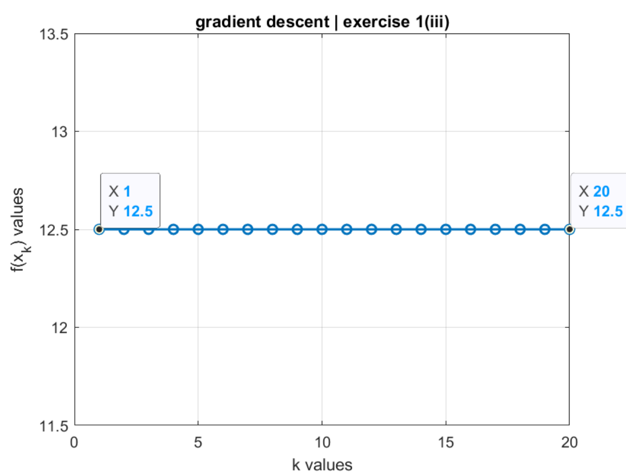
i) $\gamma_\kappa = 0.1$*:* Παρατηρούμε πως ο αλγόριθμος συγκλίνει στο σημείο ελαχίστου, ωστόσο η μικρή τιμή του γ_κ κάνει τη σύγκλιση ιδιαίτερα αργή (max_k = 60), σε σχέση με τις υπόλοιπες τιμές, όπως θα δούμε στη συνέχεια.



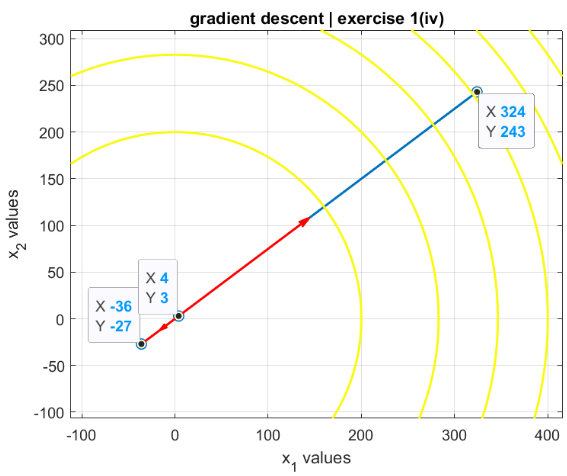


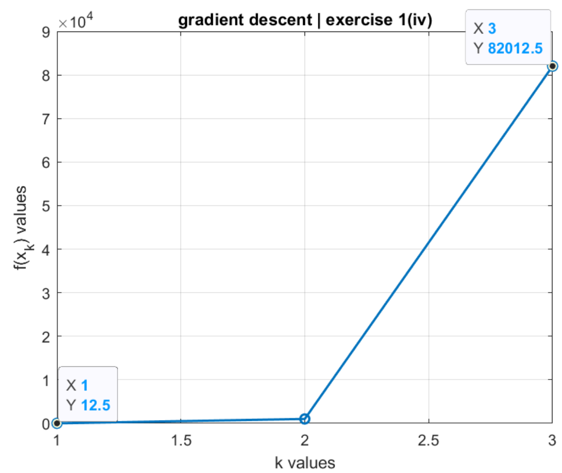
ii) $\gamma_\kappa = 1$: Ο αλγόριθμος τερματίζει με μόλις μία επανάληψη και μάλιστα καταλήγει ακριβώς στο σημείο ολικού ελαχίστου. Φαίνεται πως για αυτή τη τιμή, η αποδοτικότητα του αλγορίθμου μεγιστοποιείται.





iii) $\gamma_\kappa = 2$: Σε αυτή την περίπτωση, ο αλγόριθμος δεν τερμάτιζε με τη συνθήκη τερματισμού που ορίσθηκε και δημιουργήθηκε μια νέα συνθήκη, ώστε να τερματίσει μετά από συγκεκριμένο αριθμό επαναλήψεων. Από τα διαγράμματα παρατηρούμε, αρχικά, πως η τιμή της συνάρτησης παραμένει σταθερή μετά από κάθε επανάληψη. Επίσης, βλέπουμε πως το διάνυσμα x_k ταλαντεύεται μεταξύ των θέσεων (4, 3) και (-4, -3) σε κάθε επανάληψη, παραμένοντας έτσι στην ίδια ισοβαρή. Εφόσον οι τιμές της συνάρτησης είναι σταθερές, το γ_κ = 2 αποτελεί κρίσιμη τιμή και οδηγεί τον αλγόριθμο σε κατάσταση οριακής ευστάθειας, χωρίς να υπάρχει σύγκλιση στο ελάχιστο.





iv) γ_κ*= 10:* Σε αυτή την περίπτωση, έχοντας ξεπεράσει την κρίσιμη τιμή του γ_κ, ο αλγόριθμος έχει μεταβεί σε περιοχή αστάθειας. Βλέπουμε μάλιστα, πως από τις πρώτες κιόλας επαναλήψεις, η τιμή της συνάρτησης αυξάνεται εκθετικά, καθώς το διάνυσμα x_k κινείται προς πολύ απομακρυσμένες ισοβαρείς.

Στη συνέχεια, θα επιχειρήσουμε να επαληθεύσουμε θεωρητικά τις παραπάνω παρατηρήσεις. Θα ξεκινήσουμε από την υπόθεση ότι για κάποιο κ και γ_κ, έχουμε $f(x_{\kappa+1}) < f(x_\kappa)$ (1), με $x_{\kappa+1} = x_\kappa - \gamma_\kappa \nabla f(x_\kappa)$.

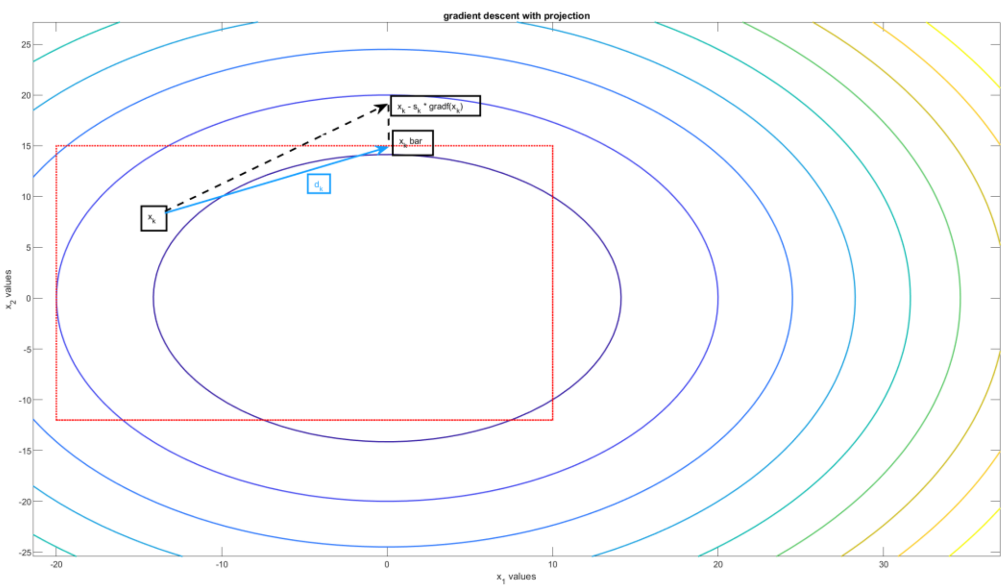
Θέτουμε $y = x_{\kappa+1} = x_\kappa - \gamma_\kappa x_\kappa = x_\kappa(1-\gamma_\kappa)$.

$$(1) \Leftrightarrow \frac{1}{2}y_1^2 + \frac{1}{2}y_2^2 < \frac{1}{2}x_1^2 + \frac{1}{2}x_2^2 \Leftrightarrow x_{\kappa 1}^2(1-\gamma_\kappa)^2 + x_{\kappa 2}^2(1-\gamma_\kappa)^2 < x_{\kappa 1}^2 + x_{\kappa 2}^2$$
$$\Leftrightarrow (1-\gamma_\kappa)^2\left(x_{\kappa 1}^2 + x_{\kappa 2}^2\right) < x_{\kappa 1}^2 + x_{\kappa 2}^2 \Leftrightarrow (1-\gamma_\kappa)^2 < 1 \Leftrightarrow 1 - 2\gamma_\kappa + \gamma_\kappa^2 < 1$$
$$\Leftrightarrow 0 < \gamma_\kappa < 2$$

Επαληθεύονται, λοιπόν, οι πρακτικές παρατηρήσεις μας, καθώς η τιμή της συνάρτησης μειώνεται σε κάθε επανάληψη αν και μόνο αν το γ_κ να βρίσκεται εντός των ορίων 0 και 2 (σύγκλιση). Προφανώς για γ_κ = 0 ή γ_κ = 2 έχουμε κρίσιμες τιμές και οριακή ευστάθεια, αφού $x_{\kappa+1} = \pm x_\kappa \Rightarrow f(x_{\kappa+1}) = f(x_\kappa)$. Για γ_κ > 2 οι τιμές της συνάρτησης αυξάνουν και έχουμε αστάθεια, ενώ η μέγιστη αποδοτικότητα επιτυγχάνεται για εκείνο το γ_κ που ελαχιστοποιεί την $f(x_\kappa - \gamma_\kappa \nabla f(x_\kappa))$, δηλαδή $f(x_\kappa - \gamma_\kappa \nabla f(x_\kappa)) = 0 \Leftrightarrow x_\kappa - \gamma_\kappa x_\kappa = 0\ \forall x_\kappa \in \mathbb{R}^2 \Leftrightarrow \gamma_\kappa = 1$, όπως είδαμε και στα παραπάνω πειράματα.

Μέθοδος Μέγιστης Καθόδου με προβολή

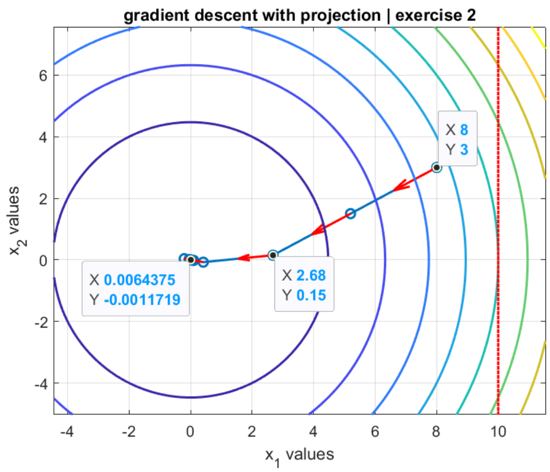
Η Μέθοδος της Μέγιστης Καθόδου είναι ένας χρήσιμος αλγόριθμος ελαχιστοποίησης για προβλήματα χωρίς περιορισμούς. Αν, ωστόσο, το πρόβλημά μας περιέχει περιορισμούς, για παράδειγμα αν πρέπει τα διανύσματα $x_k \in \mathbb{R}^2$ να βρίσκονται συνεχώς εντός ενός κυρτού συνόλου $X \subset \mathbb{R}^2$, η μέθοδος αυτή δε θα μας βοηθήσει να ολοκληρώσουμε την επίλυσή του. Αυτή την αδυναμία της μεθόδου την καλύπτουμε με τη χρήση προβολής. Υπολογίζοντας κάθε φορά το διάνυσμα κατεύθυνσης $d_k = \bar{x}_k - x_k$ (ισχύει $x_{k+1} = x_k + \gamma_\kappa d_k$), φροντίζουμε ώστε το πέρας του ($\bar{x}_k \in \mathbb{R}^2$) να βρίσκεται εντός του συνόλου $\boldsymbol{X}$, δηλαδή το διάνυσμα κατεύθυνσης να ανήκει στο σύνολο των εφικτών κατευθύνσεων. Για να επιτευχθεί αυτός ο έλεγχος, ορίζουμε το $\bar{x}_k$ ως την προβολή ενός διανύσματος που θα προέκυπτε από τη μέθοδο μέγιστης καθόδου με βήμα s_k (συνήθως μεγαλύτερο από το γ_κ). Έτσι, αν το διάνυσμα αυτό βρεθεί εκτός του $\boldsymbol{X}$, η προβολή του το επαναφέρει στα όρια του συνόλου. Αντίθετα, αν βρεθεί εντός του $\boldsymbol{X}$, η προβολή διατηρεί αυτή την τιμή. Το $\bar{x}_k$, λοιπόν βρίσκεται πλέον σίγουρα εντός του $\boldsymbol{X}$ (ή στα όριά του) και το διάνυσμα κατεύθυνσης d_k είναι εφικτό. Η χρήση της προβολής που αναφέραμε δικαιολογεί και το όνομα αυτής της επέκτασης της αρχικής μεθόδου, που ονομάζεται Μέθοδος Μέγιστης Καθόδου με προβολή. Για σταθερά γ_κ και s_k, αν το διάνυσμα κατεύθυνσης δεν πάει ποτέ να ξεφύγει εκτός του $\boldsymbol{X}$, τα x_k θα κινούνται σε σταθερή διεύθυνση. Αν όμως η χρήση της προβολής προκαλέσει την επαναφορά του διανύσματος στο όριο του $\boldsymbol{X}$, η διεύθυνση κίνησης σχεδόν σε κάθε περίπτωση θα αλλάξει (εξαίρεση αποτελεί η περίπτωση που το d_k εξέρχεται από το $\boldsymbol{X}$ σε κάποια από τις 'γωνίες' του συνόλου). Αφού βρεθεί το $\bar{x}_k$ με τη μέθοδο που περιγράψαμε και βλέπουμε στο παρακάτω σχήμα, το x_{k+1} προκύπτει αν κινηθούμε στο d_k με βήμα γ_κ.

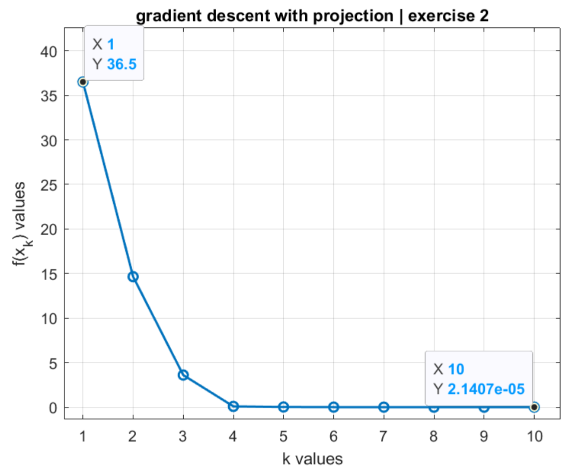


Αφού αναλύσαμε τη μέθοδο μέγιστης καθόδου με προβολή, είμαστε πλέον σε θέση να παρουσιάσουμε και να σχολιάσουμε τα αποτελέσματα των επόμενων πειραμάτων.

Θέμα 2

Υλοποιήσαμε τη μέθοδο της μέγιστης καθόδου με προβολή με ακρίβεια $\varepsilon = 0.01$, σημείο εκκίνησης (8, 3), $\gamma_\kappa = 0.1$ και $s_k = 15$.





Η σχετικά μεγάλη τιμή του s_k οδηγεί από την πρώτη επανάληψη το d_k εκτός του X με αποτέλεσμα η χρήση της προβολής να το επαναφέρει και να αλλάζει τη διεύθυνσή του. Κάθε φορά που το διάνυσμα κατεύθυνσης ξεφεύγει από τα όρια, η διεύθυνσή αλλάζει, όπως αναφέρθηκε και προηγουμένως.

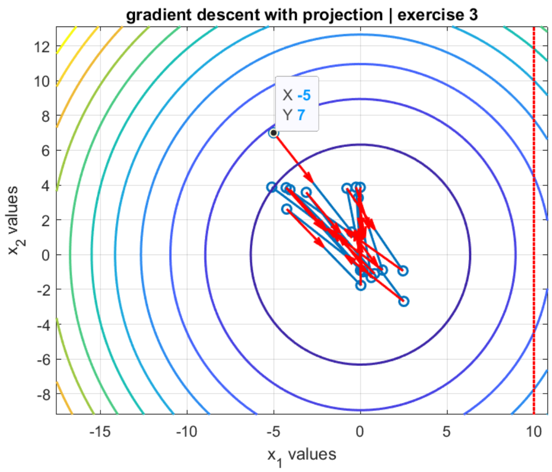
Σε σύγκριση με το Θέμα 1(i), η ύπαρξη του s_k επιταχύνει σημαντικά τον αλγόριθμο, αφού ουσιαστικά αυξάνει το μέσο βήμα σύγκλισης, 'προσέχοντας' ταυτόχρονα να μην ξεφύγει εκτός του συνόλου X. Επίσης, στο Θέμα 1(i) παρατηρούμε τη σταθερή διεύθυνση του διανύσματος κατεύθυνσης, σε αντίθεση με το Θέμα 2, όπου όπως είπαμε η διεύθυνση αλλάζει.

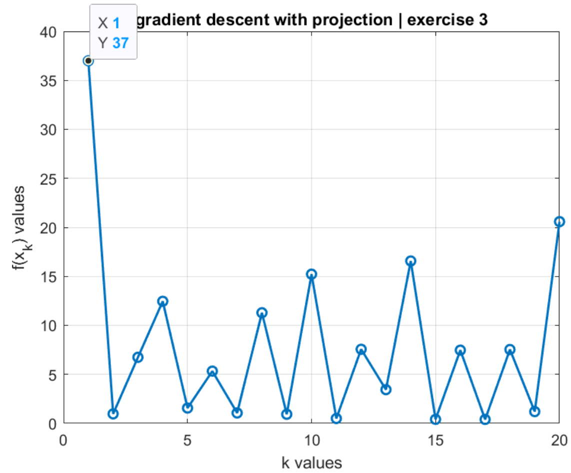
Σε σύγκριση με το Θέμα 1(iv), μπορούμε να παρατηρήσουμε πως ενώ η μεγάλη τιμή του s_k αυξάνει σημαντικά το τελικό βήμα (ουσιαστικά το βήμα είναι $\gamma_\kappa s_k$), ο αλγόριθμος δεν μπορεί να αποκλίνει σε πολύ απομακρυσμένες ισοβαρείς (ακόμη και για μεγαλύτερο βήμα), καθώς η προβολή διατηρεί πάντα τις τιμές εντός των επιθυμητών ορίων. Αυτό δε συμβαίνει στο Θέμα 1(iv), όπου δεν υπάρχει κανένας περιορισμός, με αποτέλεσμα η μέθοδος να αποτυγχάνει λόγω μεγάλης τιμής του βήματος.

Τα αποτελέσματα, λοιπόν, είναι αναμενόμενα και οι παράγοντες που τα προκαλούν αναλύθηκαν.

Θέμα 3

Επαναλάβαμε το προηγούμενο πείραμα με ακρίβεια $\varepsilon = 0.02$, σημείο εκκίνησης (-5, 7), $\gamma_\kappa = 0.3$ και $s_k = 20$.



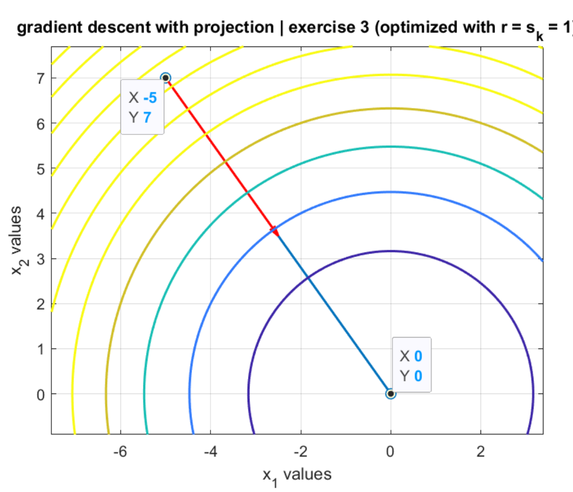


Ο αλγόριθμος αδυνατούσε να τερματίσει στις 6000 επαναλήψεις και παρουσιάζονται τα διαγράμματα για τις πρώτες 20 για να είναι πιο ευδιάκριτα. Παρατηρούμε πως το x_k 'εγκλωβίζεται' εντός του X λόγω των περιορισμών και της χρήσης της προβολής, ενώ η μεγάλη τιμή του s_k, οδηγεί σε επανειλημμένες αλλαγές στη διεύθυνση κίνησης, καθώς το d_k τείνει συνεχώς να ξεφύγει από τα όρια και η προβολή το επαναφέρει στα άκρα του συνόλου, με αποτέλεσμα να αλλάζει τη διεύθυνσή του.

Σε σύγκριση με το Θέμα 1(i), παρατηρούμε πως ενώ η πρώτη μέθοδος διατηρεί σταθερή διεύθυνση και ακολουθεί ομαλή πορεία, η δεύτερη αλλάζει τη διεύθυνσή της και ακολουθεί περίπλοκη διαδρομή. Η μέθοδος με την προβολή παρουσιάζει αστάθεια και δε συγκλίνει στο στάσιμο σημείο (0, 0). Ωστόσο, το γεγονός ότι η προβολή διατηρεί πάντα το διάνυσμα εντός του συνόλου X, καθιστά υπαρκτή την πιθανότητα κάποια στιγμή μετά από πολλές επαναλήψεις να καταλήξουμε 'τυχαία' κοντά στη θέση ολικού ελαχίστου (0, 0) και η συνθήκη τερματισμού να ικανοποιηθεί. Έτσι, ο αλγόριθμος ίσως μετά από άγνωστο αριθμό επαναλήψεων να πετύχει τον στόχο του.

Σε σύγκριση με το Θέμα 1(v), οι δύο αλγόριθμοι έχουν ως κοινό χαρακτηριστικό την αστάθεια και μη σύγκλιση για σχετικά μεγάλες τιμές των παραμέτρων βήματος (γ_κ για τον πρώτο, γ_κ και s_k για τον δεύτερο). Ωστόσο, ενώ ο πρώτος αλγόριθμος αποκλίνει σε πολύ απομακρυσμένες ισοβαρείς και δεν πρόκειται να φτάσει στο ελάχιστο, ο δεύτερος διατηρεί τις τιμές του εντός των επιθυμητών ορίων και υπάρχει το ενδεχόμενο κάποια στιγμή να τερματίσει επιτυχώς.

Θα επιχειρήσουμε να εξαλείψουμε την αστάθεια και 'τυχαιότητα' του πειράματος και να το κάνουμε να συγκλίνει με αποδοτικό τρόπο στο στάσιμο σημείο. Η ιδέα είναι να καταλήξουμε όσο το δυνατό πιο άμεσα και με μεγαλύτερη ακρίβεια στο ελάχιστο. Μειώνοντας ελαφρώς τις δύο παραμέτρους του βήματος, θα καταλήξουμε σε ένα αποτέλεσμα της μορφής του Θέματος 2 με λίγες αλλαγές διεύθυνσης. Επίσης, αν μειωθεί αρκετά το s_k, το διάνυσμα d_k θα ανήκει πάντα στις εφικτές κατευθύνσεις και ο αλγόριθμος θα μετατραπεί σε απλή μέθοδο μέγιστης καθόδου με βήμα $\gamma_\kappa s_k$ (αν μειωθεί όμως πολύ το s_k θα προκαλέσει αργή σύγκλιση). Μάλιστα, με $s_k = 1$ απαλλασσόμαστε εντελώς από τους περιορισμούς και από τις αλλαγές



κατεύθυνσης. Με κατάλληλη επιλογή του γ_κ μπορούμε να λύσουμε το πρόβλημα χωρίς τις περιττές αλλαγές κατεύθυνσης. Πρέπει όμως να είμαστε πολύ προσεκτικοί σε αυτή την επιλογή, καθώς πλέον ουσιαστικά δεν υπάρχουν περιορισμοί και για μεγάλα γ_κ θα έχουμε απόκλιση σε μεγάλες ισοβαρείς. Στο Θέμα 1, αποδείξαμε πως η μέγιστη αποδοτικότητα της Μεθόδου Μέγιστης Καθόδου είναι για γ_κ = 1. Εφαρμόζοντας, λοιπόν, τη συνθήκη s_k = γ_κ =1 καταλήγουμε στο ελάχιστο με μόλις μία επανάληψη, όπως φαίνεται στο σχήμα. Να σημειωθεί πως το αποτέλεσμα αφορά αποκλειστικά την τετραγωνική συνάρτηση που χρησιμοποιούμε, η οποία λόγω της απλότητάς και των ιδιαίτερων χαρακτηριστικών που αναφέραμε ($\nabla f(x) = x$ και ελάχιστο στο (0, 0)) μας δίνει τη δυνατότητα να φτάσουμε κατευθείαν στο στάσιμο σημείο. Για άλλες συναρτήσεις και πιο σύνθετους περιορισμούς, η μέθοδος χωρίς περιορισμούς με σταθερό γ_κ θα μπορούσε να αποτύχει και θα έπρεπε να βρούμε τις σωστές παραμέτρους s_k και γ_κ , ώστε να την βελτιστοποιήσουμε (οι τιμές του Θέματος 2 θα ήταν μια καλή επιλογή).

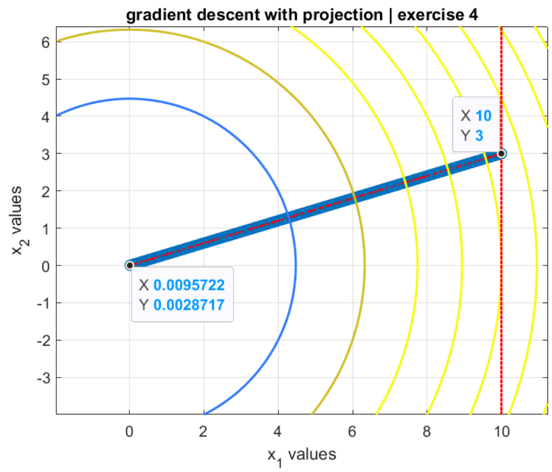
Συμπεραίνουμε λοιπόν, πως για τη συνάρτηση με την οποία ασχολούμαστε, οι περιορισμοί είναι περιττοί και η εύρεση του ελαχίστου μπορεί να γίνει με ένα βήμα. Η μέθοδος με χρήση προβολής δεν πετυχαίνει κάποια βελτίωση, καθώς η βέλτιστη επιλογή για λύση του προβλήματος είναι η απλή μέθοδος μέγιστης καθόδου με βήμα ίσο με 1. Για πιο σύνθετα προβλήματα - συναρτήσεις όμως, η χρήση της προβολής αποτελεί πολύ χρήσιμο εργαλείο για τη επίτευξη της ευστάθειας και της σύγκλισης στο ελάχιστο. Σε αυτές τις περιπτώσεις πρέπει να γίνει κατάλληλη επιλογή των δύο παραμέτρων για να μην οδηγηθούμε σε ανεπιθύμητες συμπεριφορές, όπως αυτή του πειράματος στο Θέμα 3.

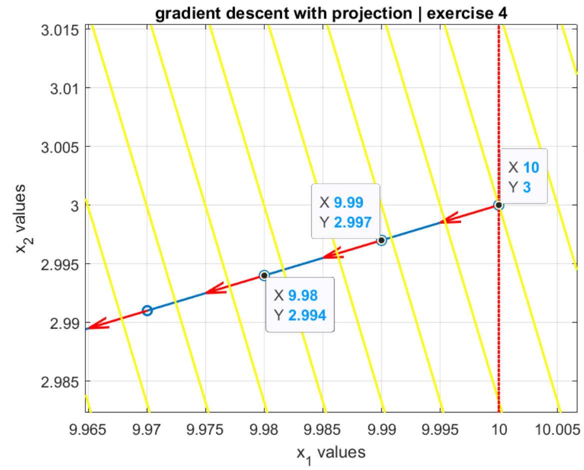
Θέμα 4

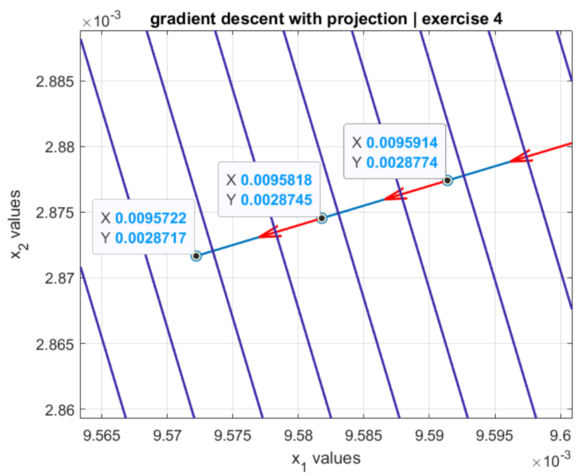
Θεωρούμε τώρα την ίδια μέθοδο, με παραμέτρους ε = 0.01, γ_κ = 0.01 και s_k = 0.1 και σημείο εκκίνησης το (11, 3). Η περίπτωση αυτή παρουσιάζει ιδιαίτερο ενδιαφέρον. Θα αποδείξουμε θεωρητικά πριν την εκτέλεση του αλγορίθμου, ότι ο αλγόριθμος πρόκειται να συγκλίνει στο ελάχιστο με πολύ μικρό βήμα.

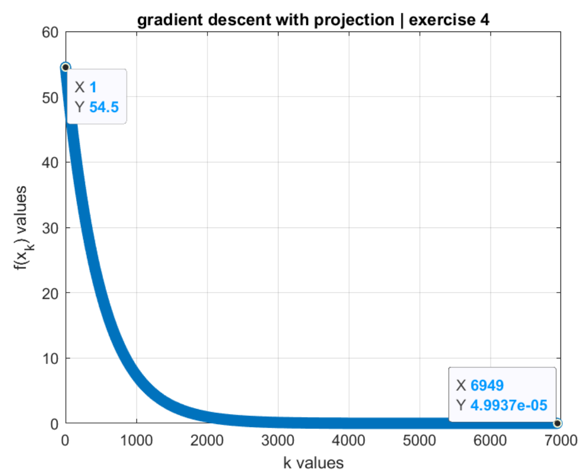
Αρχικά, παρατηρούμε πως το αρχικό σημείο βρίσκεται εκτός του X, επομένως πριν την πρώτη επανάληψη, θα χρησιμοποιηθεί η προβολή του στο X και θα γίνει (10, 3). Στη συνέχεια έχουμε $x_{\kappa+1} = x_\kappa + \gamma_\kappa(\bar{x} - x_\kappa)$, $\bar{x}$ = $\Pr_X\{x_\kappa - s_\kappa \nabla f(x_\kappa)\} = \Pr_X\{x_\kappa - s_\kappa x_\kappa\}$. Στο Θέμα 1(i), παρατηρήσαμε πως για $x_{\kappa+1} = x_\kappa - 0.1x_\kappa$ οδηγούμαστε σε χαμηλότερες τιμές της συνάρτησης, καθώς κινούμαστε προς μικρότερες ισοβαρείς και ο αλγόριθμος συγκλίνει στο (0, 0). Στο παρόν πείραμα, η χρήση της προβολής πριν από την πρώτη επανάληψη οδηγεί τον αλγόριθμο εντός του X (για την ακρίβεια στο όριό του). Σε κάθε επανάληψη επίσης, έχουμε όπως είπαμε $\bar{x}$ = $\Pr_X\{x_\kappa - s_\kappa \nabla f(x_\kappa)\} = \Pr_X\{x_\kappa - s_\kappa x_\kappa\}$ με s_k = 0.1. Η τιμή όμως της παράστασης $x_\kappa - 0.1x_\kappa$ θα είναι σίγουρα εντός του X αφού αν βγει εκτός, θα οδηγηθεί σε μεγαλύτερη ισοβαρή, που είναι άτοπο, όπως εξηγήσαμε. Άρα θα έχουμε πάντα $\bar{x} = \Pr_X\{x_\kappa - s_\kappa x_\kappa\} = x_\kappa - s_\kappa x_\kappa$, οπότε $x_{\kappa+1} = x_\kappa - \gamma_\kappa(x_\kappa - s_\kappa x_\kappa - x_\kappa) = x_\kappa - \gamma_\kappa s_\kappa x_\kappa$, δηλαδή προκύπτει απλή μέθοδος μέγιστης καθόδου με βήμα $\gamma_\kappa s_k$=0.001. Όπως αποδείξαμε στο Θέμα 1, θα έχουμε πολύ αργή σύγκλιση στο (0, 0), χωρίς καμία αλλαγή διεύθυνσης και με πολύ μεγάλο αριθμό επαναλήψεων, λόγω του πολύ μικρού βήματος. Ο αλγόριθμος λοιπόν, αναμένεται να αργήσει αρκετά να φτάσει στο επιθυμητό αποτέλεσμα.

Παρουσιάζουμε τα διαγράμματα που προκύπτουν από την εκτέλεση του αλγορίθμου σε Matlab:









Παρατηρούμε πως επαληθεύονται οι θεωρητικές εκτιμήσεις μας, καθώς ο αλγόριθμος είναι ευσταθής και συγκλίνει με πολύ αργό ρυθμό στο (0, 0). Η μέθοδος έχει όντως μετατραπεί σε απλή μέθοδο μέγιστης καθόδου με βήμα 0.001 και το πολύ μικρό αυτό βήμα οδηγεί σε πολύ μεγάλο αριθμό επαναλήψεων. Επίσης, βλέπουμε πως το πρώτο σημείο του αλγορίθμου δεν είναι το (11, 3) που δίνουμε ως είσοδο, αλλά το (10, 3), όπως αναφέραμε νωρίτερα.

Σύντομη περιγραφή του κώδικα

Στον φάκελο της εργασίας περιλαμβάνεται ο υποφάκελος *matlab_code*, όπου βρίσκεται ο κώδικας της εργασίας. Στον υποφάκελο *gradient_descent* βρίσκεται ο κώδικας για την απλή Μέθοδο Μέγιστης Καθόδου που χρησιμοποιήθηκε στο Θέμα 1, ενώ η υλοποίηση της Μεθόδου Μέγιστης Καθόδου με προβολή βρίσκεται στον υποφάκελο *gradient_descent_with_projection*. Και στις δύο περιπτώσεις, υπάρχει ένα αρχείο *main.m* το οποίο τρέχουμε για την εκτέλεση του κώδικα. Επίσης και στις δύο περιπτώσεις υπάρχει το αρχείο *grad_des.m* που περιλαμβάνει την εκάστοτε συνάρτηση ελαχιστοποίησης. Όταν κληθεί στη main με τα κατάλληλα ορίσματα, επιστρέφει το τελικό διάνυσμα που ελαχιστοποιεί την συνάρτηση με την επιθυμητή ακρίβεια και τον αριθμό *κ* των διανυσμάτων που μεσολάβησαν (*κ*=αριθμός επαναλήψεων+1). Ταυτόχρονα, η κλήση της συνάρτησης παράγει και τα επιθυμητά διαγράμματα που εκφράζουν τη συμπεριφορά του αλγορίθμου. Σημειώνεται πως στο εσωτερικό της επαναληπτικής μεθόδου περιέχεται μια συνθήκη, η οποία τερματίζει τον αλγόριθμο αν ο αριθμός των επαναλήψεων ξεπεράσει κάποιο όριο, για να ελεγχθεί στο διάγραμμα αν ο αλγόριθμος αποκλίνει σε απομακρυσμένες ισοβαρείς, ή αν υπάρχει κάποιο άλλο πρόβλημα (για παράδειγμα στο Θέμα 1(iii) χρησιμοποιήθηκε η συνθήκη για *κ=20* να τερματίσει ο αλγόριθμος, διότι διαφορετικά δε θα τερμάτιζε ποτέ). Η συνάρτηση που βρίσκεται στο αρχείο *plot_dir.m* αποτελεί ένα script που προσθέτει βέλη κατεύθυνσης όταν κληθεί στο figure. Τέλος, η μέθοδος με προβολή περιλαμβάνει μια συνάρτηση *projection* (*projection.m*), η οποία υλοποιεί τον τύπο της προβολής διανύσματος δύο διαστάσεων σε υποσύνολο του $\mathbb{R}^2$ και καλείται σε κάθε επανάληψη της Μεθόδου Μέγιστης Καθόδου με προβολή.