Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης
Πολυτεχνική Σχολή

Τμήμα Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών

Μάθημα:Τεχνικές Βελτιστοποίησης

3η Εργαστηριακή Άσκηση

Ελεάνα Ζέρη   ΑΕΜ:10811

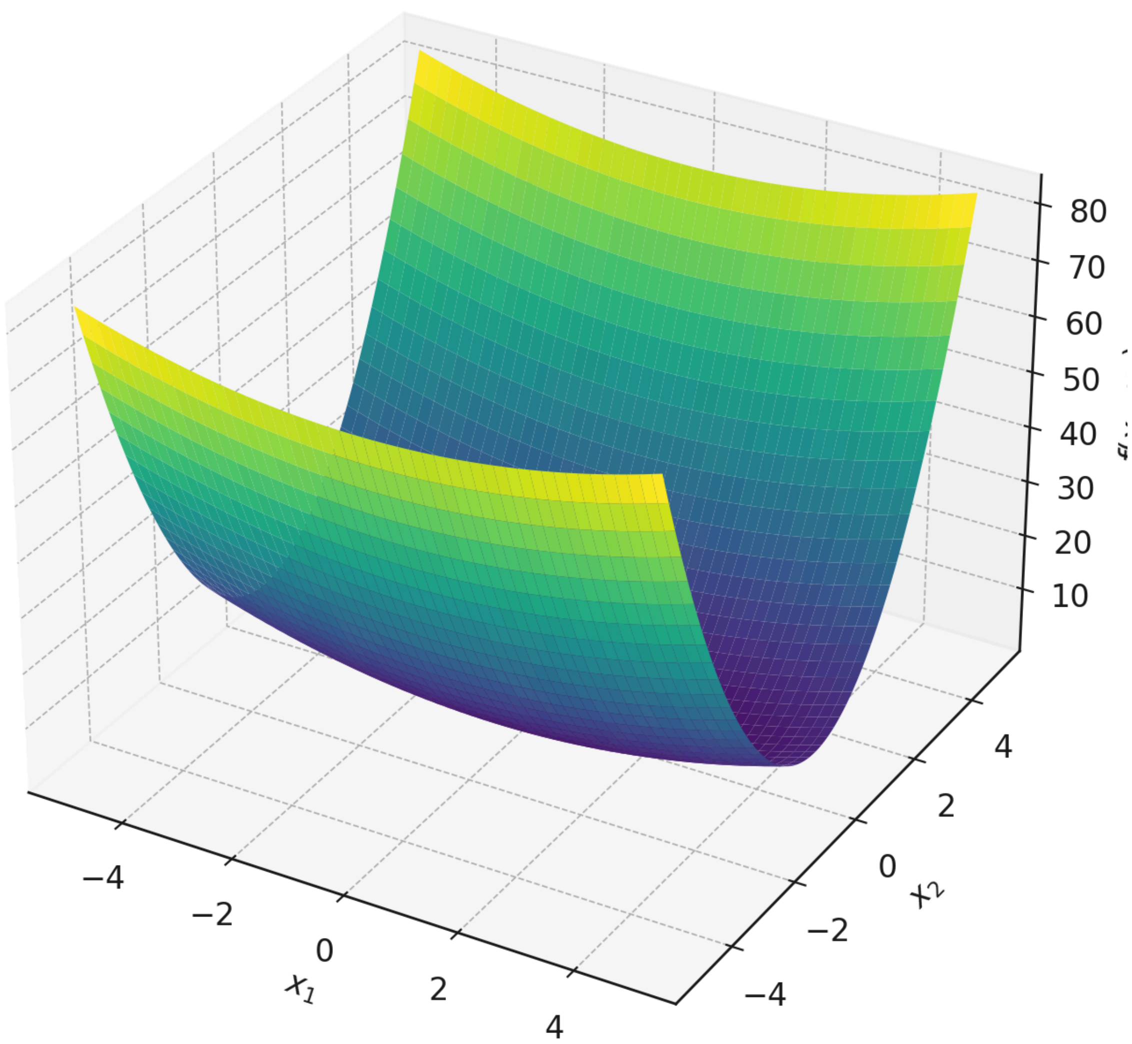
ΣΥΝΤΟΜΗ ΠΕΡΙΓΡΑΦΗ

Η μέθοδος *Μέγιστης Καθόδου με Προβολή* είναι μια παραλλαγή της κλασικής μεθόδου Μέγιστης Καθόδου (Steepest Descent), που χρησιμοποιείται για την ελαχιστοποίηση συναρτήσεων πολλών μεταβλητών. Η διαφορά είναι ότι εδώ υπάρχουν περιορισμοί στη λύση, δηλαδή η λύση πρέπει να ανήκει σε ένα συγκεκριμένο επιτρεπτό σύνολο.

Στην κλασική Μέγιστη Κάθοδο: Ξεκινάμε από ένα αρχικό σημείο και κινούμαστε προς την κατεύθυνση του αρνητικού gradient (που είναι η κατεύθυνση μέγιστης καθόδου) μέχρι να ελαχιστοποιήσουμε τη συνάρτηση.

Στη Μέγιστη Κάθοδο με Προβολή: Αν το νέο σημείο που υπολογίζουμε δεν ανήκει στο επιτρεπτό σύνολο (δηλαδή παραβιάζει τους περιορισμούς), τότε προβάλλουμε αυτό το σημείο πίσω στο επιτρεπτό σύνολο.

$$f(x) = \frac{1}{3}x_1^2 + 3x_2^2$$



(*Μας δίνεται η παραπάνω συνάρτηση πάνω στην οποία θα δουλέψουμε.Είναι μια κυρτή συνάρτηση που έχει μόνο ένα ελάχιστο, το οποίο βρίσκεται στο σημείο x1=0,x2=0)*

ΑΠΟΤΕΛΕΣΜΑΤΑ ΣΥΝΟΛΙΚΑ ΤΩΝ ΘΕΜΑΤΩΝ 1,2,3,4:

Θέμα	Αρχικό Σημείο	γ_k	s_k	Τελικό Σημείο	Αριθμός Επαναλήψεων
1	(8, -10)	0.1, 0.3, 3.0, 5.0	-	(βλέπε κάτω από το γράφημα του θέματος 1)	93, 45, 100000, 100000
2	(5, -5)	0.5	5	(0, -1.3333)	100000
3	(-5, 10)	0.1	15	(0, 0.00091599)	1215
4	(8, -10)	0.2	0.1	(0.014885, 0)	448

ΘΕΜΑ 1: Να χρησιμοποιηθεί η Μέθοδος Μέγιστης Καθόδου με ακρίβεια ε = 0.001 και βήμα i) γκ = 0.1, ii) γκ = 0.3, iii) γκ = 3, iv) γκ = 5 και οποιοδήποτε αρχικό σημείο εκκίνησης διαφορετικό του (0,0). Τι παρατηρείτε; Να αποδειχθούν τα αποτελέσματα αυτά με μαθηματική αυστηρότητα.

Εφαρμόζοντας την μέθοδο της Μέγιστης Καθόδου με τις παραμέτρους της εκφώνησης και αρχικό σημείο επιλογής (8, −10) παίρνουμε το εξης figure και αποτελέσματα:

Μέγιστη Κάθοδος: Αρχικό σημείο: (8.000000, -10.000000), Σύγκλιση σε (0.014011, -0.000000) με γάμμα 0.100000 σε 93 επαναλήψεις.

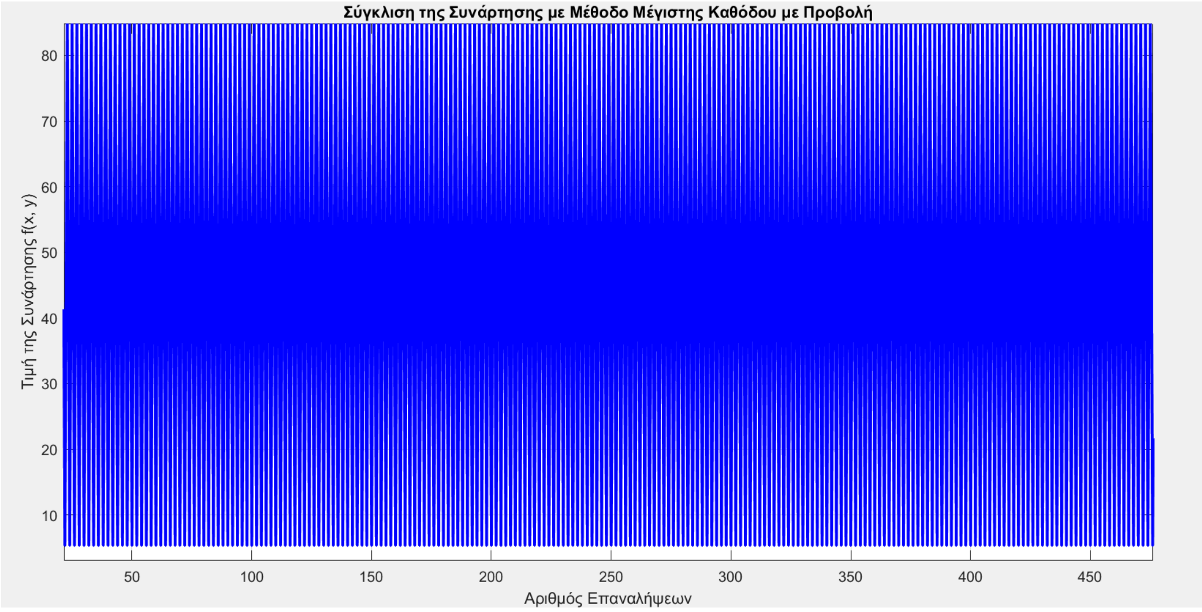
Μέγιστη Κάθοδος: Αρχικό σημείο: (8.000000, -10.000000), Σύγκλιση σε (0.000436, -0.000544) με γάμμα 0.300000 σε 45 επαναλήψεις.

Μέγιστη Κάθοδος: Αρχικό σημείο: (8.000000, -10.000000), Σύγκλιση σε (8.000000, NaN) με γάμμα 3.000000 σε 100000 επαναλήψεις.
Μέγιστη Κάθοδος: Αρχικό σημείο: (8.000000, -10.000000), Σύγκλιση σε (NaN, NaN) με γάμμα 5.000000 σε 100000 επαναλήψεις.

Σχολιασμός:

Για γ=0.1 η μέθοδος συγκλίνει αργά, σε 93 επαναλήψεις, στο σημείο (0.014011,0).Αυτό δείχνει ότι η τιμή του γ είναι πολύ μικρή, με αποτέλεσμα μικρά βήματα και αργή σύγκλιση.
Για γ=0.3 η μέθοδος συγκλίνει πιο γρήγορα, σε 45 επαναλήψεις, στο σημείο (0.000436,−0.000544).Η τιμή αυτή του γ,λοιπόν,είναι πιο κατάλληλη, προσφέροντας ισορροπία μεταξύ βήματος και σταθερότητας.Για γ=0.1 και γ=0.3, ο αλγόριθμος ακολουθεί μια πιο ελεγχόμενη πορεία, με σταθερότερες τιμές στη συνάρτηση, και τελικά καταλήγει σε ένα σημείο κοντά στο ελάχιστο.
Για γ=3 και γ=5 παρατηρείται αποτυχία σύγκλισης.Πιο συγκεκριμένα,η συνάρτηση εκτοξεύεται προς τα πάνω, αυξάνεται συνεχώς χωρίς περιορισμούς και δεν συγκλίνει στο ελάχιστο, καθώς οι τιμές των μεταβλητών γίνονται NaN (Not A Number,όχι αριθμός). Αυτό οφείλεται στο ότι το βήμα είναι υπερβολικά μεγάλο, οδηγώντας τη μέθοδο εκτός της περιοχής ελαχίστου.Για γ=5 η αποτυχία είναι ακόμα πιο έντονη.
Αυτή η εκτροπή του βήματος μπορεί να προκύψει αν το μέγεθος του βήματος γk είναι μεγαλύτερο από το όριο της σταθερότητας, το οποίο μπορεί να υπολογιστεί ως:
{όπου H είναι η Hessian (δεύτερη παράγωγος) της συνάρτησης στο σημείο τοπικού ελάχιστου}
Για τη συνάρτηση μας ισχύει H=6
γk<2/ // H // =2/6=0.3333
Συνεπώς, η μέγιστη αποδεκτή τιμή για το βήμα είναι: γk<0.3333
Αν το γk υπερβαίνει αυτό το όριο (όπως για γk=3 και γk=5), η μέθοδος δεν θα συγκλίνει.
Συμπερασματικά,η τιμή γ=0.3,τελικά,είναι η πιο κατάλληλη σε αυτό το παράδειγμα.

*ΘΕΜΑ 2:*Να χρησιμοποιηθεί η Μέθοδος Μέγιστης Καθόδου με Προβολή, με sk = 5, γκ= 0.5, σημείο εκκίνησης το (5, −5) και ακρίβεια ε = 0.01. Τι παρατηρείτε σε σχέση με το Θέμα 1;



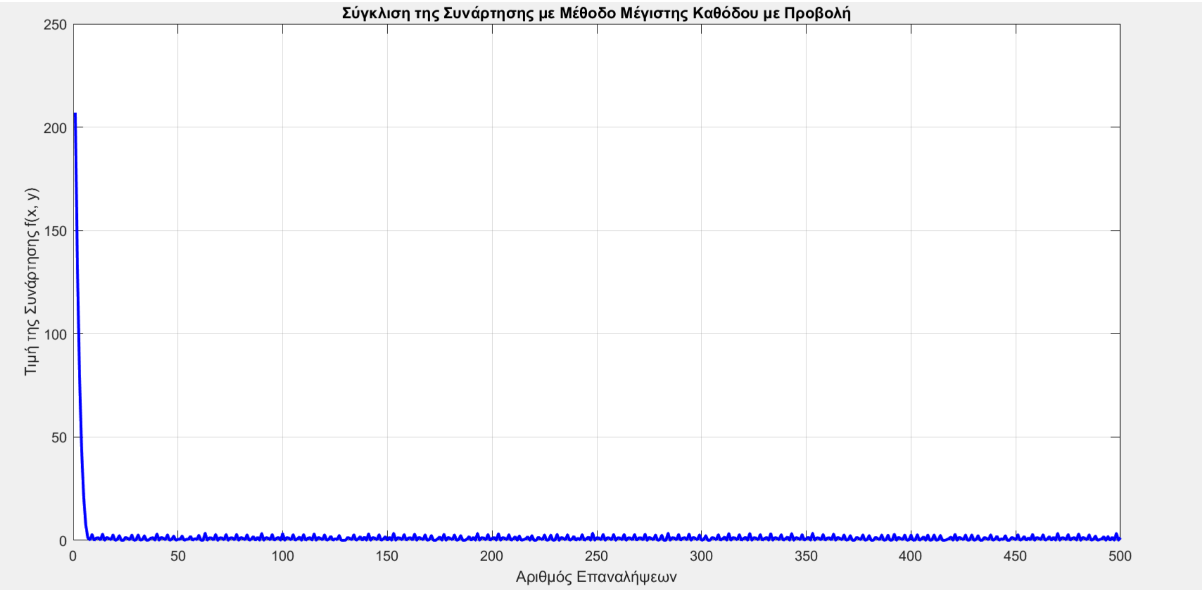
Αποτελέσματα Μεθόδου Μέγιστης Καθόδου με Προβολή:
Τελικό σημείο: (2.532e-88, -1.3333)
Αριθμός επαναλήψεων: 500

Αποτελέσματα Μεθόδου Μέγιστης Καθόδου με Προβολή:
Τελικό σημείο: (1.9763e-323, -1.3333)
Αριθμός επαναλήψεων: 100000

Σχολιασμός:

Παρατηρούμε απο τη παραπάνω γραφικη παράσταση ότι η τιμή της συνάρτησης αρχικά μειώνεται γρήγορα καθώς το αρχικό μας σημειο (5, −5) με τιμή συναρτησης f(5, −5)~83 είναι αρκετά μακριά από το ολικό ελάχιστο(=0).Με την πάροδο των επαναλήψεων και καθώς η τιμή της συνάρτησης μειώνεται προς την κατεύθυνση του ολικού ελαχίστου, ο αλγόριθμος κάνει αναπηδήσεις και αντί να καταλήξει στο ελάχιστο τελικά εκτροχιάζεται και απομακρύνεται τελείως από αυτό καταλήγοντας σε πολύ μεγάλες τιμές συνάρτησης.Η μέθοδος Μέγιστης Καθόδου κάνει μεγάλες μετατοπίσεις στην κατεύθυνση του gradient,μεταξύ των τιμών της συνάρτησης 5,3333 και 85.3333.Οι μεγάλες αναπηδήσεις προκαλούν ασταθή σύγκλιση και αναδρομικές ταλαντώσεις όπως βλέπουμε και αυτές οι εκτροπές οφείλονται στο γεγονός ότι το βήμα γk ξεπερνά την βέλτιστη τιμή 0.3 που υπολογίστηκε παραπάνω και το sk είναι αρκετά μεγάλο, οπότε το επόμενο σημείο που υπολογίζεται μπορεί να είναι πολύ μακριά από το ελάχιστο, με αποτέλεσμα ο αλγόριθμος να μην καταλήγει ποτέ στην ακριβή λύση.Τα αριθμητικά αποτελέσματα (2.532⋅10−88,−1.3333) ή (1.9763⋅10−323,−1.3333),δηλαδή το τελικό σημείο δεν πλησιάζει ιδιαίτερα το ελάχιστο και ο μέγιστος αριθμός επαναλήψεων υποδηλώνουν ότι ο αλγόριθμος παγιδεύεται χωρίς να επιτύχει την επιθυμητή ακρίβεια.

ΘΕΜΑ 3:Να χρησιμοποιηθεί η Μέθοδος Μέγιστης Καθόδου με Προβολή, με sk = 15, γk = 0.1, σημείο εκκίνησης το (−5,10) και ακρίβεια ε = 0.01. Τι παρατηρείτε σε σχέση με τα Θέματα 1 και 2; Προτείνετε έναν απλό πρακτικό τρόπο ώστε η μέθοδος να συγκλίνει στο ελάχιστο.



Αποτελέσματα Μεθόδου Μέγιστης Καθόδου με Προβολή:
Τελικό σημείο: (0, -0.66212)
Αριθμός επαναλήψεων: 500

Αποτελέσματα Μεθόδου Μέγιστης Καθόδου με Προβολή:
Τελικό σημείο: (0, 0.00091599)
Αριθμός επαναλήψεων: 1215

Σχολιασμός:

Η τιμή της συνάρτησης μειώνεται σταδιακά με την αύξηση των επαναλήψεων. Η μείωση στην αρχή είναι απότομη γιατί το αρχικό σημείο (−5,10) είναι μακριά από το ολικό ελάχιστο (0,0){f(−5,10)~308 που απέχει πολύ από το ολικό ελάχιστο =0}.Στη συνέχεια παρατηρούμε,μικρές ταλαντώσεις άρα ασταθή σύγκλιση πάλι.

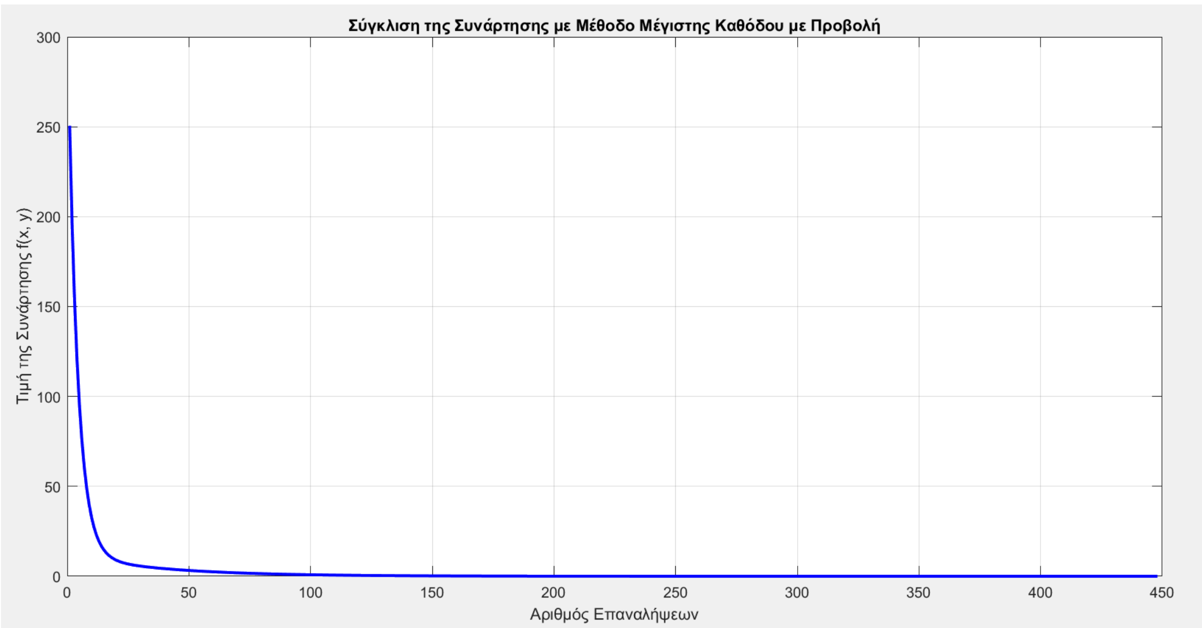
Με βάση τα αποτελέσματα του Θέματος 1 για γk=0.1,παρατηρήθηκε αργή σύγκλιση λόγω μικρών βημάτων. Συγκεκριμένα, στο Θέμα 1 η μέθοδος συγκλίνει σε 93 επαναλήψεις στο σημείο (0.014011,0), ενώ στο Θέμα 3 η σύγκλιση είναι πιο γρήγορη και πιο σταθερή παρά το ίδιο γk καθώς η μέθοδος απαιτεί λιγότερες επαναλήψεις για να φτάσει κοντά στο ελάχιστο στην αρχική φάση{πιο συγκεκριμένα μόλις στις 10 επαναλήψεις φτάνει για 1η φορα στην τιμή ~0.017}, λόγω του μεγαλύτερου sk που επιτρέπει μεγαλύτερα βήματα.Βέβαια,στο Θέμα 1, η τελική τιμή της συνάρτησης στο σημείο (0.014011,0) στις 93 επαναλήψεις είναι πολύ κοντά στο μηδέν ενώ η μέθοδος μέγιστης καθόδου με προβολή(Θέμα 3) καταλήγει στις συνολικά 1215 επαναλήψεις στο σημείο (0, 0.00091599) που είναι μεν πιο κοντινό του (0,0) αλλά μετα από πάρα πολλές επαναλήψεις.

Η μέθοδος δεν παρουσιάζει τόσο ασταθή σύγκλιση όπως το Θέμα 2.Η τελική τιμή της συνάρτησης είναι 1.3152, η οποία είναι αρκετά κοντά στο ελάχιστο, που είναι το 0 και συνεπώς προσεγγίζεται καλύτερα το ελάχιστο απ’οτι στο Θέμα 2(έλαχιστο σημείο ταλάντωσης 5.3333).Το μεγαλύτερο sk=15 προκαλεί μεγαλύτερες αρχικές "αναπηδήσεις", αλλά η μικρή τιμή γk=0.1 διατηρεί τη σύγκλιση σταθερή.Στο Θέμα 3,λοιπόν,σε σύγκριση με το Θέμα 2 που αυξήθηκε κατά πολύ το sk αλλά ταυτόχρονα μειώθηκε το γk είδαμε βελτιωμένη συμπεριφορά και καλύτερη σύγκλιση.

Πρόταση για Βελτίωση
μπορούμε να:

- **μειώσουμε το sk**: Ένα sk≤5 θα μείωνε τις απότομες κινήσεις, αποτρέποντας την υπέρβαση των ορίων και διευκολύνοντας την ομαλή σύγκλιση.
- **αυξήσουμε το γk**: Ένα γk≈0.3 θα επέτρεπε καλύτερη προσαρμογή των βημάτων, επιτυγχάνοντας πιο γρήγορη και σταθερή σύγκλιση.
- **προσαρμοστικό βήμα**: Χρήση μιας δυναμικής προσέγγισης όπου το sk μειώνεται καθώς πλησιάζουμε το ελάχιστο(όπως κάναμε στη 2η εργαστηριακή άσκηση στο (β) υποερώτημα των Θέματων 2,3,4)

ΘΕΜΑ 4:Να χρησιμοποιηθεί η Μέθοδος Μέγιστης Καθόδου με Προβολή, με sk = 0.1, γk = 0.2, σημείο εκκίνησης το (8, −10) και ακρίβεια ε = 0.01. Σε αυτή την περίπτωση, έχουμε εκ των προτέρων κάποια πληροφορία σχετικά με την σύγκλιση του αλγορίθμου; Να γίνει η εκτέλεση του αλγορίθμου. Τι παρατηρείτε;



Αποτελέσματα Μεθόδου Μέγιστης Καθόδου με Προβολή:
Τελικό σημείο: (0.014885, -1.3435e-24)
Αριθμός επαναλήψεων: 448

Σχολιασμός:

Στο Θέμα 4, το μικρότερο sk=0.1 βοηθά στην αργή αλλά σταθερή σύγκλιση, χωρίς υπερβολικές μετακινήσεις κοντά στο ελάχιστο.Παρόλα αυτά, το μικρότερο sk απαιτεί περισσότερες επαναλήψεις για να πλησιάσει το ελάχιστο απ’ότι το sk=15(Θέμα

3),ενώ με το sk=0.1 ερχόμαστε ακόμη πιο κοντά στο ολικό ελάχιστο:{sk=0.1:Τελικό σημείο: (0.014885, -1.3435e-24),με sk=15:Τελικό σημείο: (0, -0.66212)} για σχεδόν τον ίδιο αριθμό επαναλήψεων:{448 έναντι των 500 επαναλήψεων αντίστοιχα}

Τέλος,προς απάντηση του ερωτήματος αν σε αυτή την περίπτωση έχουμε εκ των προτέρων κάποια πληροφορία σχετικά με την σύγκλιση του αλγορίθμου,θα απαντήσω πως έχοντας δει τη συμπεριφορά της συνάρτησης στις προηγούμενες δοκιμές(Θέμα 1,Θέμα 2,Θέμα 3) και βλέποντας τις παραμέτρους που χρησιμοποιούνται σε αυτή τη δοκιμή(Θέμα 4:sk=0.1,γk=0.2) υποψιαζόμουν ότι θα έχουμε μια πιο ομαλή πορεία προς το ολικό ελάχιστο και θα έχουμε τη καλύτερη σύγκλιση από ολές τις άλλες εκδοχές της μεθόδου μέγιστης καθόδου με προβολή.Αυτό γιατί όπως έχει αναλυθεί και παραπάνω το γk είναι πολύ κοντά στο 0.3 και ικανοποιεί και προσεγγίζει τη σχέση που αποδείξαμε στο Θέμα 1: γk<0.333 ενώ το sk είναι πολύ μικρό οπότε θα έχουμε πολλά μικρά και σταθερά βήματα άρα και πολλές επαναλήψεις αλλά θα έχουμε καλύτερη τελική σύγκλιση.