

Τεχνικές Βελτιστοποίησης - Εργασία 3

Ρουσομάνης Γεώργιος (ΑΕΜ: 10703)

Δεκέμβριος 2024

Εισαγωγή

Στόχος της συγκεκριμένης εργασίας είναι η ελαχιστοποίηση μίας μη γραμμικής συνάρτησης $f(x)$, η οποία υπόκειται σε περιορισμούς της μορφής $\alpha_i \leq x_i \leq \beta_i$, με τη χρήση της μεθόδου μέγιστης καθόδου με προβολή. Η αντικειμενική συνάρτηση που θα μελετηθεί είναι η:

$$f(x) = \frac{1}{3}x_1^2 + 3x_2^2, \quad x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}. \quad (1)$$

Η γραφική παράσταση της συνάρτησης παρουσιάζεται στο Σχ. 1.

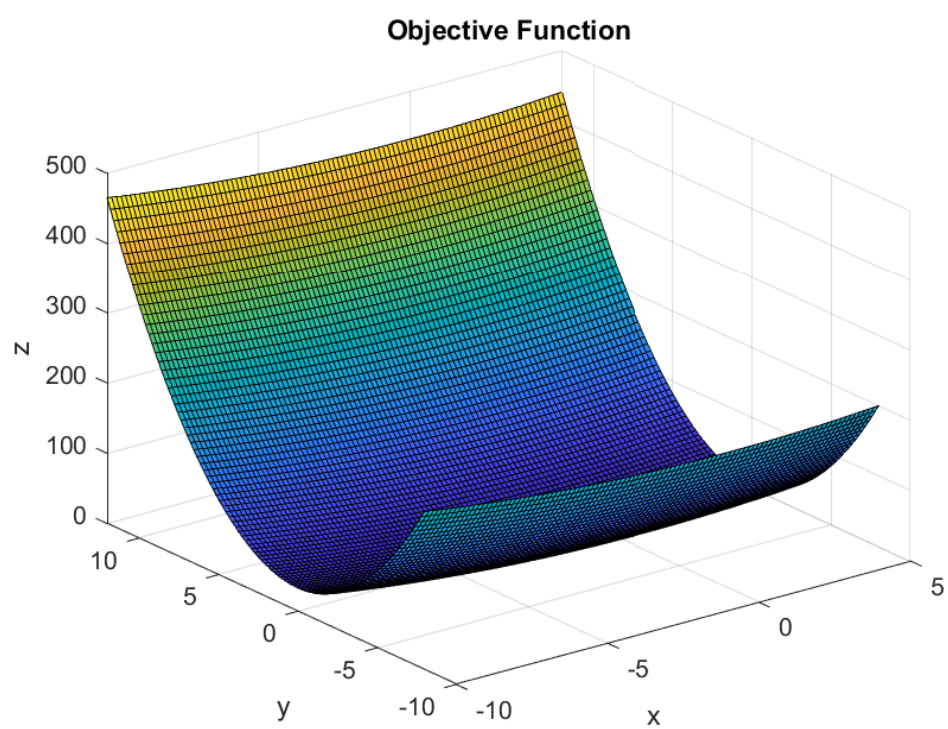
Η ανάλυση θα περιλαμβάνει την επίδραση των γ_k, s_k καθώς και του σημείου αφετηρίας x_0 στη σύγκλιση της μεθόδου, ενώ τα δεδομένα της προσομοίωσης στον υπολογιστή θα επαληθευτούν με μαθηματική ανάλυση.

Τα αποτελέσματα της μεθόδου μέγιστης καθόδου με προβολή θα συγκριθούν με εκείνα της κλασικής μεθόδου μέγιστης καθόδου χωρίς περιορισμούς. Τέλος θα επισημανθούν τα πλεονεκτήματα και τα μειονεκτήματα της κάθε μεθόδου καθώς και το πότε χρησιμοποιείται η κάθε μία.

Θέμα 1

Όπως είναι ήδη γνωστό, η μέθοδος μέγιστης καθόδου αποτελεί μια τεχνική βελτιστοποίησης για προβλήματα απουσία περιορισμών. Ο αλγόριθμος υπολογίζει διαδοχικά σημεία μέσω της αναδρομικής σχέσης:

$$x_{k+1} = x_k - \gamma_k \nabla f(x_k) \implies \begin{bmatrix} x_{1(k+1)} \\ x_{2(k+2)} \end{bmatrix} = \begin{bmatrix} x_{1k} \\ x_{2k} \end{bmatrix} - \gamma_k \begin{bmatrix} \frac{2}{3}x_{1k} \\ 6x_{2k} \end{bmatrix} \quad (2)$$



Σχήμα 1: Αντικειμενική Συνάρτηση

Προφανώς, η f παρουσιάζει ολικό ελάχιστο στο $x^* = (0, 0)$. Για να συγκλίνει η μέθοδος για οποιοδήποτε αρχικό σημείο $x_0 \neq x^*$ θα πρέπει να ισχύει:

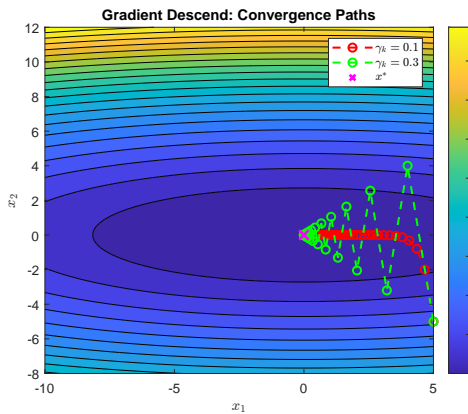
$$\left| \frac{x_{1(k+1)}}{x_{1k}} \right| < 1 \implies \left| 1 - \frac{2}{3}\gamma_k \right| < 1 \implies 0 < \gamma_k < 3 \quad (3)$$

και

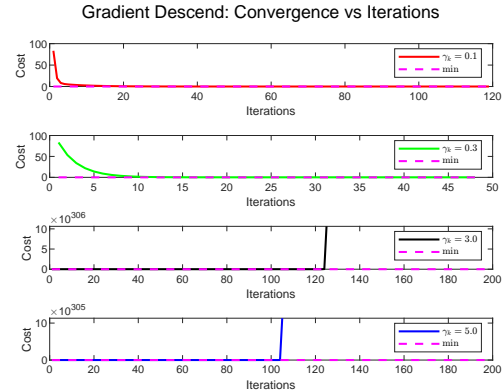
$$\left| \frac{x_{2(k+1)}}{x_{2k}} \right| < 1 \implies |1 - 6\gamma_k| < 1 \implies 0 < \gamma_k < \frac{1}{3} \quad (4)$$

Από τα παραπάνω προκύπτει ότι για να έχουμε σύγκλιση στο $x^* = (0, 0)$ θα πρέπει να ισχύει $0 < \gamma_k < \frac{1}{3}$.

Στο Σχ. 3 παρουσιάζεται το γράφημα σύγκλισης της αντικειμενικής συνάρτησης ως προς τον αριθμό των επαναλήψεων με αρχικό σημείο $x_0 = (5, -5)$, για διάφορες τιμές του γ_k . Παρατηρούμε ότι για $\gamma_k = 0.1$ και $\gamma_k = 0.3$ έχουμε επιτυχή σύγκλιση (δεδομένου ότι $\gamma_k < 1/3$), ενώ για $\gamma_k = 3$ και $\gamma_k = 5$ η μέθοδος αποκλίνει (δεδομένου ότι $\gamma_k > 1/3$). Το αποτέλεσμα αυτό συμφωνεί με τα συμπεράσματα της θεωρητικής ανάλυσης. Επίσης για $\gamma_k = 0.3$ έχουμε ταχύτερη σύγκλιση από ότι για $\gamma_k = 0.1$.



Σχήμα 2: Διαδοχικά σημεία υπολογισμού για τις περιπτώσεις σύγκλισης για το Θέμα 1



Σχήμα 3: Σύγκλιση της αντικειμενικής συνάρτησης ως προς τον αριθμό των επαναλήψεων για την μέθοδο μέγιστης καθόδου

Θέμα 2

Η μέθοδος μέγιστης καθόδου με προβολή αποτελεί μία τεχνική βελτιστοποίησης για προβλήματα με περιορισμούς. Ο αλγόριθμος υπολογίζει διαδοχικά εφικτά σημεία της μορφής

$$x_{k+1} = x_k + \gamma_k(\bar{x}_k - x_k), \quad \gamma_k \in (0, 1] \quad (5)$$

όπου

$$\bar{x}_k = Pr_X\{x_k - s_k \nabla f(x_k)\}, \quad s_k > 0. \quad (6)$$

Επομένως, το \bar{x}_k προσδιορίζεται από το x_k αν κινηθούμε στην κατεύθυνση της αρνητικής κλήσης με βήμα s_k και στην συνέχεια βρούμε την προβολή του $x_k - s_k \nabla f(x_k)$ στο X , καταλήγοντας με τον τρόπο αυτό σε εφικτό σημείο. Έχοντας βρει την εφικτή κατεύθυνση αναζήτησης $\bar{x}_k - x_k$, κινούμαστε με βήμα $\gamma_k \in (0, 1]$ σ' αυτήν και προσδιορίζουμε το νέο σημείο x_{k+1} . Το x_{k+1} θα είναι επίσης εφικτό, καθώς αποτελεί γραμμικό συνδυασμό των $x_k, \bar{x}_k \in X$ με $\gamma_k \in (0, 1]$, και X κυρτό σύνολο.

Θα βρούμε τις συνθήκες που θα πρέπει να ικανοποιούν τα γ_k, s_k ώστε η μέθοδος να συγκλίνει στο $x^* = (0, 0)$. Θέτουμε $\hat{x}_k = x_k - s_k \nabla f(x_k)$ οπότε η σχέση (6) μπορεί να γραφεί:

$$\bar{x}_k = Pr_X\{\hat{x}_k\} = Pr_X\left\{\begin{bmatrix} x_{1k} \\ x_{2k} \end{bmatrix} - s_k \begin{bmatrix} \frac{2}{3}x_{1k} \\ 6x_{2k} \end{bmatrix}\right\} = Pr_X\left\{\begin{bmatrix} x_{1k}(1 - \frac{2}{3}s_k) \\ x_{2k}(1 - 6s_k) \end{bmatrix}\right\} \quad (7)$$

Διακρίνουμε τις εξής δύο περιπτώσεις:

- Αν $\hat{x}_k \in X$, τότε ισχύει $\bar{x}_k = \hat{x}_k$ οπότε, μετά από λίγες πράξεις η (5) γράφεται:

$$\begin{bmatrix} x_{1(k+1)} \\ x_{2(k+1)} \end{bmatrix} = \begin{bmatrix} x_{1k}(1 - \frac{2}{3}\gamma_k s_k) \\ x_{2k}(1 - 6\gamma_k s_k) \end{bmatrix} \quad (8)$$

Εργαζόμενοι όπως στο Θέμα 1 βρίσκουμε ότι για να συγκλίνει η μέθοδος θα πρέπει να ισχύει $0 < \gamma_k s_k < \frac{1}{3}$.

- Αν $\hat{x}_k \notin X$, θα εξετάσουμε την περίπτωση όπου $\hat{x}_{1k} < \alpha_1$. Στο ίδιο συμπέρασμα καταλήγουμε και για την περίπτωση όπου $\hat{x}_{1k} > \beta_1$. Άρα έχουμε ότι $\bar{x}_k = [\alpha_1 \quad x_{2k}(1 - 6x_k)]^T$ και η (5) μετά από μερικές πράξεις γράφεται:

$$\begin{bmatrix} x_{1(k+1)} \\ x_{2(k+1)} \end{bmatrix} = \begin{bmatrix} x_{1k}(1 - \gamma_k) + \gamma_k \alpha_1 \\ x_{2k}(1 - 6\gamma_k s_k) \end{bmatrix} \quad (9)$$

Παρατηρούμε ότι το $x_{1(k+1)}$ γράφεται σαν γραμμικός συνδυασμός των x_{1k}, α_1 με $\gamma_k \in (0, 1]$ και εφόσον το x_k είναι εφικτό σημείο, ισχύει ότι $x_{1k} > \alpha_1$ άρα $x_{1(k+1)} \in (\alpha_1, x_{1k}]$. Συνεπώς η συνιστώσα $x_{1(k+1)}$ δεν μας εισάγει κάποιον περιορισμό για τα γ_k, s_k . Για το $x_{2(k+1)}$ βρίσκουμε ότι $0 < \gamma_k s_k < 3$.

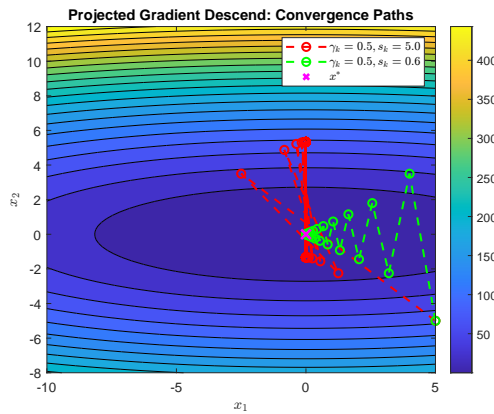
Εργαζόμενοι κατ' αντίστοιχο τρόπο για την περίπτωση όπου $\hat{x}_{2k} < \alpha_2$ ή $\hat{x}_{2k} > \beta_2$ βρίσκουμε ότι $0 < \gamma_k s_k < \frac{1}{3}$.

Από την παραπάνω ανάλυση προκύπτει ότι για να συγκλίνει η μέθοδος μέγιστης καθόδου με προβολή, θα πρέπει να ισχύει $0 < \gamma_k s_k < \frac{1}{3}$.

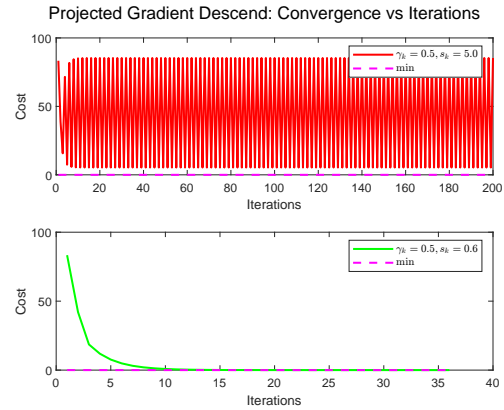
Στο Σχ. 4 φαίνονται τα διαδοχικά σημεία που υπολογίζει ο αλγόριθμος μέγιστης καθόδου με προβολή, με σημείο αφετηρίας $x_0 = (5, -5)$, $s_k = 5$, $\gamma_k = 0.5$ και ακρίβεια $\epsilon = 0.01$. Παρατηρούμε ότι εμφανίζεται ταλάντωση γύρω από το x^* . Αυτό συμβαίνει διότι $\gamma_k s_k > \frac{1}{3}$. Επίσης, καθώς ο αλγόριθμος κινείται προς το x^* , η κατεύθυνση κίνησης που ορίζεται από το $\bar{x}_k - x_k$ έχει την τάση να γίνει κάθετη προς την κατεύθυνση που οδηγεί στο ελάχιστο, κάνοντας τον αλγόριθμο ιδιαίτερα αναποτελεσματικό.

Στο ίδιο σχήμα βλέπουμε την τροχιά των σημείων που υπολογίζει ο αλγόριθμος για $s_k = 0.6$, $\gamma_k = 0.5$ και $x_0 = (5, -5)$. Σε αυτήν την περίπτωση έχουμε σύγκλιση της μεθόδου καθώς ισχύει $\gamma_k s_k < \frac{1}{3}$.

Στο Σχ. 5 παρουσιάζεται το γράφημα σύγκλισης της αντικειμενικής συνάρτησης συναρτήσει του αριθμού επαναλήψεων. Στην πρώτη περίπτωση (όπου $\gamma_k s_k > \frac{1}{3}$) βλέπουμε ότι η αδυναμία σύγκλισης της μεθόδου μέγιστης καθόδου με προβολή οδηγεί σε ταλάντωση της τιμής της αντικειμενικής συνάρτησης καθώς αυξάνεται το πλήθος των επαναλήψεων. Αυτό είναι αντίθετο με τη συμπεριφορά της κλασικής μεθόδου μέγιστης καθόδου, όπου η αδυναμία σύγκλισης οδηγεί σε απειρισμό της αντικειμενικής συνάρτησης. Η παραπάνω διαφορά εξηγείται από το γεγονός ότι η μέθοδος μέγιστης καθόδου με προβολή υπολογίζει σημεία μόνο εντός του κυρτού συνόλου X , αποτρέποντας τον απειρισμό της συνάρτησης.



Σχήμα 4: Διαδοχικά σημεία υπολογισμού της μεθόδου μέγιστης καθόδου με προβολή για το Θέμα 2



Σχήμα 5: Σύγκλιση της αντικειμενικής συνάρτησης ως προς τον αριθμό των επαναλήψεων για την μέθοδο μέγιστης καθόδου με προβολή για το Θέμα 2

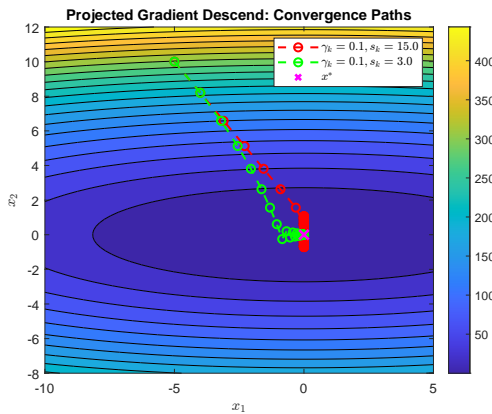
Θέμα 3

Σε αυτή την περίπτωση θα χρησιμοποιήσουμε την μέθοδο μέγιστης καθόδου με προβολή με $s_k = 15$, $\gamma_k = 0.1$ με σημείο εκκίνησης το $x_0 = (-5, 10)$ και ακρίβεια $\epsilon = 0.01$.

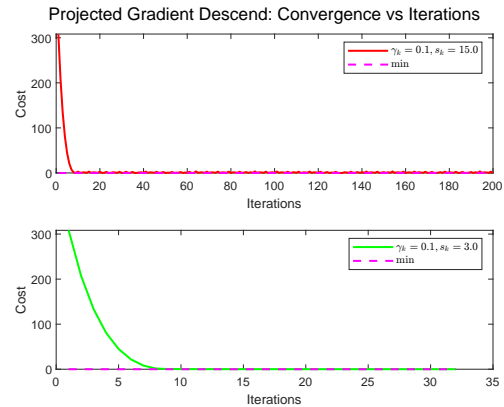
Στο Σχ. 6 φαίνεται η τροχιά των σημείων που υπολογίζει ο αλγόριθμος. Όπως και στο Θέμα 2, παρατηρείται αδυναμία σύγκλισης, καθώς $\gamma_k s_k > \frac{1}{3}$. Ωστόσο, η ταλάντωση που παρουσιάζεται έχει πολύ μικρότερο πλάτος, λόγω της μικρότερης τιμής του γ_k .

Επιπλέον, από το Σχ. 7 βλέπουμε ότι δεν εμφανίζεται απειρισμός της αντικειμενικής συνάρτησης, σε αντίθεση με το Θέμα 1. Όπως έχει ήδη επισημανθεί, αυτό συμβαίνει επειδή ο αλγόριθμος μέγιστης καθόδου με προβολή υπολογίζει σημεία εντός του κυρτού συνόλου X .

Ένας πρακτικός τρόπος ώστε η μέθοδος να συγκλίνει στο ελάχιστο, είναι η μείωση του s_k ή γ_k έτσι ώστε να ισχύει $\gamma_k s_k < \frac{1}{3}$. Στα Σχ. 6 και Σχ. 7 βλέπουμε την σύγκλιση της μεθόδου για $\gamma_k = 0.1$ και $s_k = 3$.



Σχήμα 6: Διαδοχικά σημεία υπολογισμού της μεθόδου μέγιστης καθόδου με προβολή για το Θέμα 3



Σχήμα 7: Σύγκλιση της αντικειμενικής συνάρτησης ως προς τον αριθμό των επαναλήψεων για την μέθοδο μέγιστης καθόδου με προβολή για το Θέμα 3

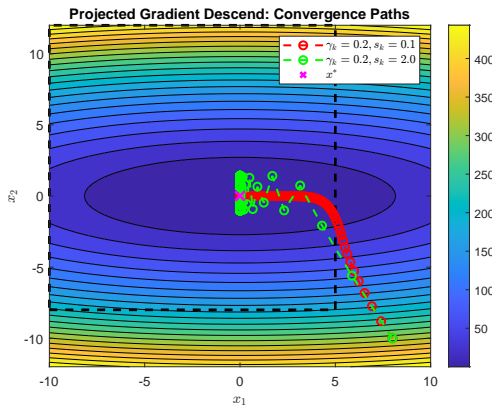
Θέμα 4

Θα χρησιμοποιήσουμε την μέθοδο μέγιστης καθόδου με προβολή με $s_k = 0.1$, $\gamma_k = 0.2$ με σημείο εκκίνησης το $x_0 = (8, -10)$ και ακρίβεια $\epsilon = 0.01$. Σε αυτήν την περίπτωση το σημείο αφετηρίας x_0 δεν είναι εφικτό, οπότε θα αναμέναμε ότι ο αλγόριθμος δεν θα συγκλίνει στο ελάχιστο. Ωστόσο, από τα Σχ. 8 και Σχ. 9 διαπιστώνουμε ότι κάτι τέτοιο δεν ισχύει.

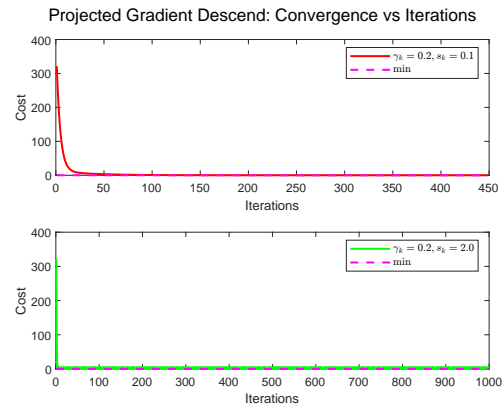
Αυτό μπορεί να εξηγηθεί αν ανατρέξουμε στον τρόπο λειτουργίας της μεθόδου. Σε κάθε επανάληψη, ο αλγόριθμος προβάλλει το διάνυσμα μέγιστης καθόδου $x_k - s_k \nabla f(x_k)$ στο εφικτό και κυρτό σύνολο X . Με αυτόν τον τρόπο διορθώνεται η κατεύθυνση κίνησης του σημείου, οδηγώντας το κάθε φορά προς το κοντινότερο σημείο εντός του X .

Μόλις το σημείο εισέλθει εντός του X , εφόσον ισχύει $\gamma_k s_k < \frac{1}{3}$, ο αλγόριθμος θα συγκλίνει στο ελάχιστο. Στα Σχ. 8 και Σχ. 9 φαίνεται η αδυναμία σύγκλισης της μεθόδου όταν θέσουμε $s_k = 2$ έτσι ώστε $\gamma_k s_k > \frac{1}{3}$. Ωστόσο, και σε αυτή την περίπτωση, το σημείο εισέρχεται εντός του X .

Συνεπώς, η χρήση της προβολής επιτρέπει στον αλγόριθμο να ξεπεράσει το αρχικό πρόβλημα του μη εφικτού σημείου εκκίνησης και να συγκλίνει κανονικά στο βέλτιστο σημείο εντός των περιορισμών, με την προϋπόθεση ότι $\gamma_k s_k < \frac{1}{3}$.



Σχήμα 8: Διαδοχικά σημεία υπολογισμού της μεθόδου μέγιστης καθόδου με προβολή για το Θέμα 4



Σχήμα 9: Σύγκλιση της αντικειμενικής συνάρτησης ως προς τον αριθμό των επαναλήψεων για την μέθοδο μέγιστης καθόδου με προβολή για το Θέμα 4

Συμπεράσματα

Από την ανάλυση των αποτελεσμάτων για τη μέθοδο μέγιστης καθόδου και τη μέθοδο μέγιστης καθόδου με προβολή, προκύπτουν τα εξής συμπεράσματα:

- **Μέθοδος Μέγιστης Καθόδου χωρίς Περιορισμούς**

1. Η σύγκλιση της μεθόδου εξαρτάται από την επιλογή του βήματος γ_k και όχι από το σημείο αφετηρίας.
2. Για να εξασφαλιστεί σύγκλιση προς το ελάχιστο της αντικειμενικής συνάρτησης, πρέπει να ισχύει $0 < \gamma_k < \frac{1}{3}$. Εάν γ_k είναι μεγαλύτερο, η μέθοδος αποκλίνει.
3. Τα μεγαλύτερα αποδεκτά γ_k εντός του παραπάνω διαστήματος ($\gamma_k \rightarrow \frac{1}{3}$) εξασφαλίζουν ταχύτερη σύγκλιση.

- **Μέθοδος Μέγιστης Καθόδου με Προβολή**

1. Η μέθοδος είναι κατάλληλη για προβλήματα με περιορισμούς και υπολογίζει διαδοχικά εφικτά σημεία εντός του κυρτού συνόλου X .
2. Η σύγκλιση εξαρτάται από τις παραμέτρους s_k και γ_k , με την υπόθεση ότι $0 < \gamma_k s_k < \frac{1}{3}$ για να επιτευχθεί σύγκλιση.
3. Σε περιπτώσεις που η συνθήκη $\gamma_k s_k < \frac{1}{3}$ δεν ισχύει, παρατηρείται ταλάντωση γύρω από το βέλτιστο σημείο, χωρίς όμως να εμφανίζεται απειρισμός της αντικειμενικής συνάρτησης, καθώς τα σημεία παραμένουν εντός του κυρτού συνόλου X .
4. Η μέθοδος μέγιστης καθόδου με προβολή μπορεί να ξεκινήσει από μη εφικτά σημεία, καθώς η προβολή διασφαλίζει ότι τα επόμενα σημεία υπολογίζονται εντός του κυρτού συνόλου X .
5. Αφού το σημείο εκκίνησης εισέλθει στο X , η μέθοδος μπορεί να συγκλίνει στο ελάχιστο εφόσον $\gamma_k s_k < \frac{1}{3}$.

- **Γενικά Συμπεράσματα**

1. Η μέθοδος μέγιστης καθόδου με προβολή είναι πιο ευέλικτη σε σύγκριση με την κλασική μέθοδο μέγιστης καθόδου, καθώς μπορεί να διαχειριστεί προβλήματα με περιορισμούς και μη εφικτά σημεία εκκίνησης.
2. Η μέθοδος μέγιστης καθόδου με προβολή εγγυάται ότι δεν θα αποκλίνει, ακόμα και αν οι παράμετροι γ_k και s_k δεν ικανοποιούν τη σχέση $0 < \gamma_k s_k < \frac{1}{3}$. Η λύση παραμένει 'παγιδευμένη' εντός του

κυρτού συνόλου X , υπό την προϋπόθεση ότι η αντικειμενική συνάρτηση παρουσιάζει ελάχιστο εντός του X . Αντίθετα, στην κλασική μέθοδο μέγιστης καθόδου, όταν $\gamma_k > \frac{1}{3}$, παρατηρείται απόκλιση.