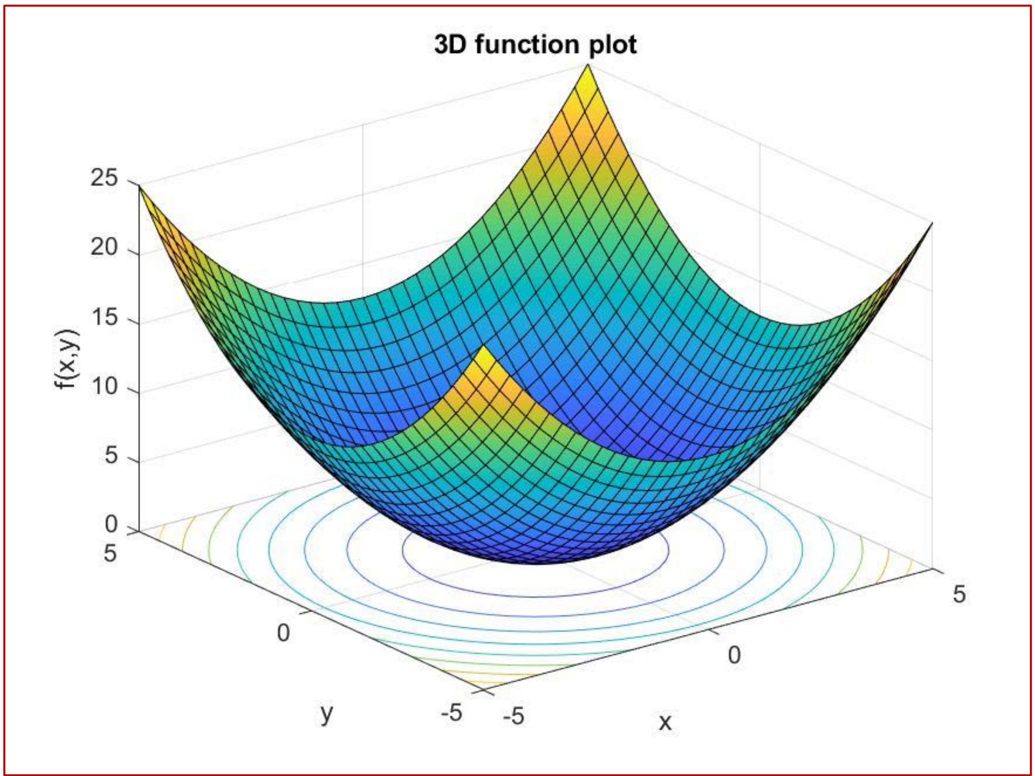
Δοϊνάκης Μιχαήλ
ΑΕΜ:9292
e-mail: doinakis@ece.auth.gr

ΕΙΣΑΓΩΓΗ

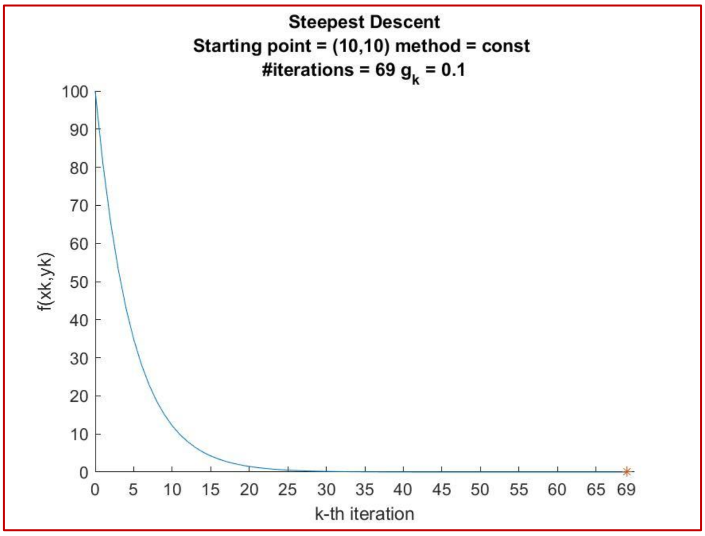
Στόχος της τρίτης εργασίας είναι ελαχιστοποίηση μιας συνάρτησης 2 μεταβλητών με περιορισμούς. Πιο συγκεκριμένα, συνάρτηση προς ελαχιστοποίηση είναι η $f(x_1,x_2) = \frac{1}{2} * x_1^2 + \frac{1}{2} * x_2^2$ η οποία απεικονίζεται και στο σχήμα 1. Αρχικά, μελετάται η συμπεριφορά της μεθόδου της μέγιστης καθόδου, της προηγούμενης εργασίας, με διάφορα σταθερά γ και το ίδιο αρχικό σημείο για να βρεθεί το διάστημα για το οποίο η μέθοδος συγκλίνει. Έπειτα, παρουσιάζονται τα αποτελέσματα της μεθόδου μέγιστης καθόδου με προβολή για διάφορες αρχικές τιμές $s_k, \gamma_\kappa, \varepsilon$ και σημείων εκκίνησης. Οι περιορισμοί που τίθενται είναι: $-20 \le \chi_1 \le 10 \; \kappa\alpha\iota -12 \le \chi_2 \le 15$.

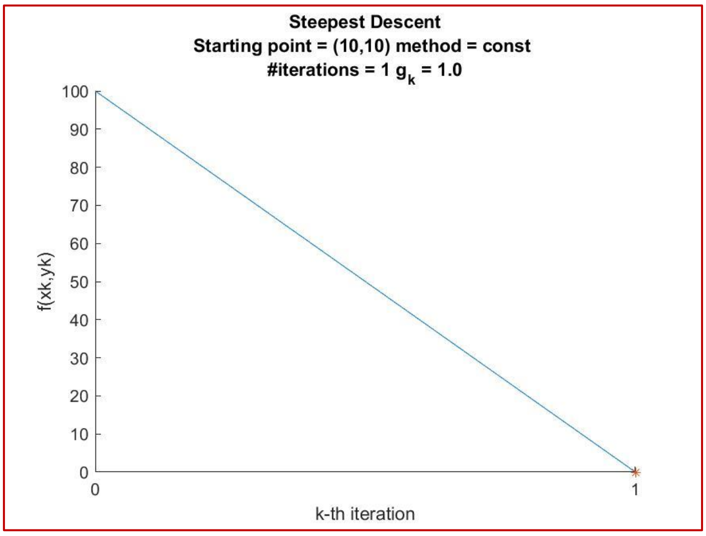
Επιπλέον, υλοποιήθηκαν 5 matlab scripts **Subject0, Subject1st**, **Subject2nd**, **Subject3rd**, **Subject4th** (το Subject0 είναι για την απεικόνιση της συνάρτησης), και κάθε ένα υλοποιεί το αντίστοιχο θέμα της εργασίας.

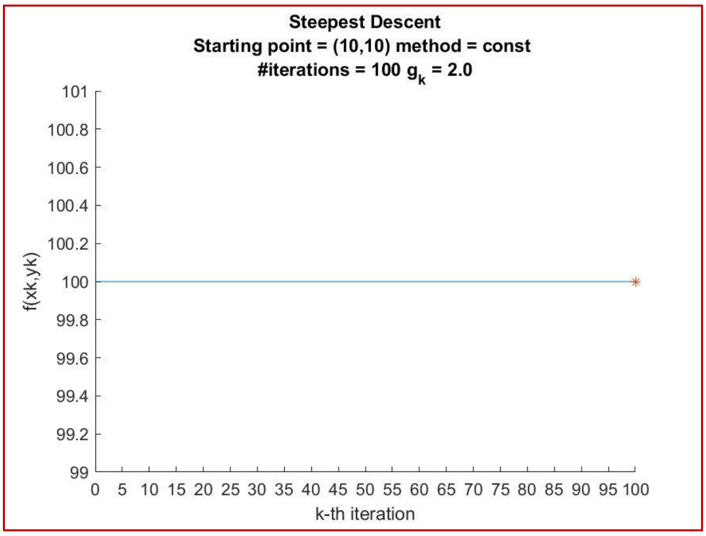


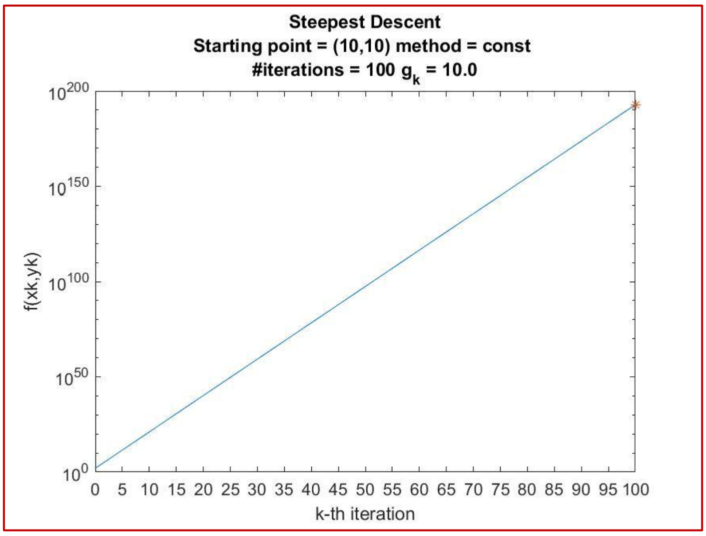
ΘΕΜΑ 1: ΜΕΘΟΔΟΣ ΜΕΓΙΣΤΗΣ ΚΑΘΟΔΟΥ

Στο πρώτο θέμα χρησιμοποιήθηκε η μέθοδος της μέγιστης καθόδου της προηγούμενης εργασίας για $\varepsilon = 0.01$και βήματα i) $\gamma_\kappa = 0.1$ ii) $\gamma_\kappa = 1$ iii) $\gamma_\kappa = 2$ iv) $\gamma_\kappa = 10$ και ως αρχικό σημείο επιλέχθηκε το $(-10,-10)$. Στις περιπτώσεις (i) και (ii) η μέθοδος κατάφερε να βρει το ελάχιστο με πολύ καλή ακρίβεια, σε αντίθεση με τις περιπτώσεις (iii) και (iv) όπου η μέθοδος δεν τερμάτιζε. Στη περίπτωση που το $\gamma_\kappa = 2$,ταλαντεύεται μεταξύ των σημείων $(-10,-10)$ και $(10,10)$, ενώ όταν $\gamma_\kappa = 10$ ο αλγόριθμος αποκλίνει κατά πολύ από το ελάχιστο και οι τιμές τις συνάρτησης αυξάνονται εκθετικά οδηγώντας σε αστάθεια. Παρακάτω παρουσιάζονται διαγράμματα που δείχνουν πως αλλάζει η τιμή της συνάρτησης συναρτήσει του αριθμού των επαναλήψεων. Σημειώνεται πως για τις δύο τελευταίες περιπτώσεις ο αλγόριθμος σταμάτησε εξαναγκασμένα στις 100 επαναλήψεις και ο άξονας των y για $\gamma_\kappa = 10$ είναι σε λογαριθμική κλίμακα.









Για ποιο λόγω συμβαίνει όμως αυτό; Η επιλογή του γ_κ 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 $f(x_1,x_2) = \frac{1}{2} * x_1^2 + \frac{1}{2} * x_2^2$.

<u>Απόδειξη</u>

Αρχικά υπολογίζουμε το gradient της συνάρτησης ως εξής:

$\nabla f(x_1,x_2) = \begin{bmatrix} \frac{\partial f(x_1,x_2)}{\partial x_2} \\ \frac{\partial f(x_1,x_2)}{\partial x_2} \end{bmatrix} = \begin{bmatrix} \frac{2}{2} * x_1 + 0 \\ 0 + \frac{2}{2} * x_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ άρα: $\nabla f\left(x_{1_k}, x_{2_k}\right) = \begin{bmatrix} x_{1_k} \\ x_{2_k} \end{bmatrix}$. Επομένως το επόμενο σημείο $\begin{bmatrix} x_{1_{k+1}} \\ x_{2_{k+1}} \end{bmatrix}$ είναι $\begin{bmatrix} x_{1_{k+1}} \\ x_{2_{k+1}} \end{bmatrix} = \begin{bmatrix} x_{1_k} \\ x_{2_k} \end{bmatrix} - \gamma * \begin{bmatrix} x_{1_k} \\ x_{2_k} \end{bmatrix} \Leftrightarrow \begin{cases} x_{1_{k+1}} = x_{1_k} - \gamma * x_{1_k} \\ x_{2_{\kappa+1}} = x_{2_k} - \gamma * x_{2_k} \end{cases} \Leftrightarrow \begin{cases} x_{1_{k+1}} = (1-\gamma) * x_{1_k} \\ x_{2_{\kappa+1}} = (1-\gamma) * x_{2_k} \end{cases} (1).$

Από τη μορφή που έχει η συνάρτηση είναι προφανές ότι το ελάχιστο βρίσκεται στο σημείο $\begin{bmatrix} x_1^* \\ x_2^* \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$. Για να συγκλίνει η μέθοδος λοιπόν από την $(1) \Leftrightarrow |1-\gamma| < 1 \Leftrightarrow -1 < 1-\gamma < 1 \Leftrightarrow -2 < -\gamma < 0 \Leftrightarrow 0 < \gamma < 2$ ■

Επομένως, η τιμή του γ πρέπει να είναι μεταξύ του μηδενός και μικρότερη του δύο για να συγκλίνει η μέθοδος, επαληθεύοντας έτσι και τα πειραματικά αποτελέσματα.

ΜΕΘΟΔΟΣ ΜΕΓΙΣΤΗΣ ΚΑΘΟΔΟΥ ΜΕ ΠΡΟΒΟΛΗ

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:

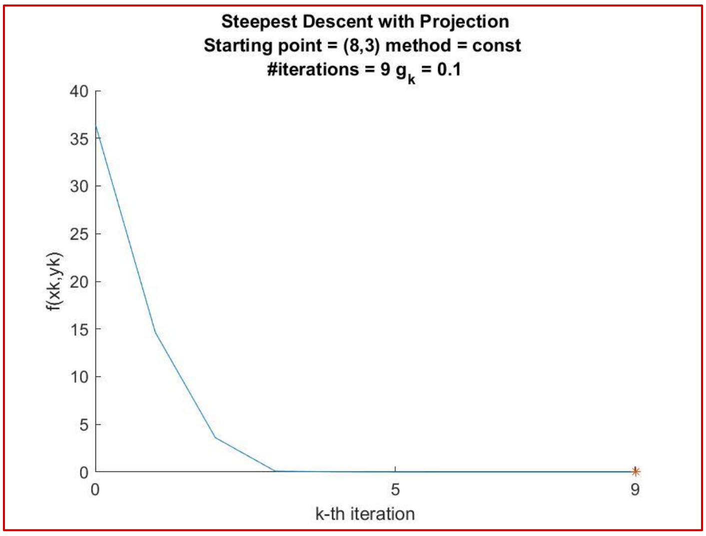
$x_{k+1} = x_k + \gamma_k * (\bar{x} - x_k)(2)$,όπου $\bar{x} = \Pr_X\{x_k - s_k * \nabla f(x_k)\}$.

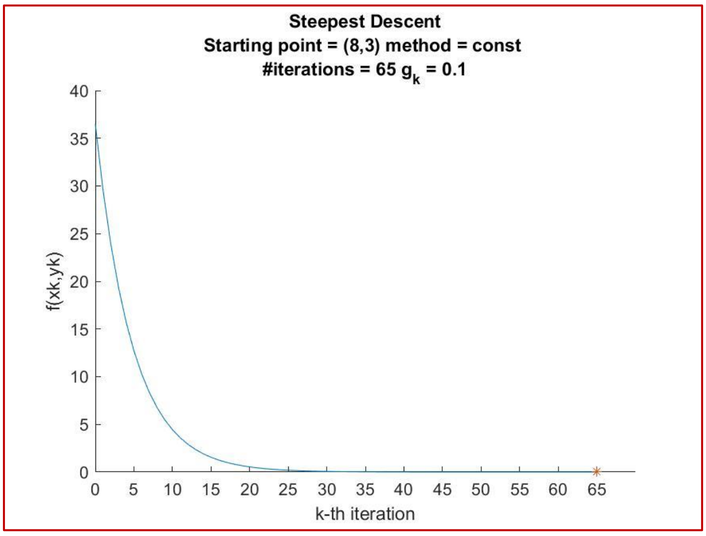
Όμως αν το x_k είναι εφικτό σημείο τότε $\bar{x} = x_k - s_k * \nabla f(x_k)(3)$

$(2) \overset{(3)}{\Leftrightarrow} x_{k+1} = x_k - \gamma_k * s_k * \nabla f(x_k)$. Βγαίνει λοιπόν το συμπέρασμα πως το x_{k+1} υπολογίζεται από τον αλγόριθμο της μέγιστης καθόδου με $\gamma_k' = \gamma_k * s_k$. Επομένως, ο ίδιος περιορισμός που αποδείχθηκε για το γ πρέπει να ισχύει και για το $\gamma_k{}'$. Με βάση αυτό το συμπέρασμα αναλύονται τα υπόλοιπα θέματα της εργασίας.

ΘΕΜΑ 2

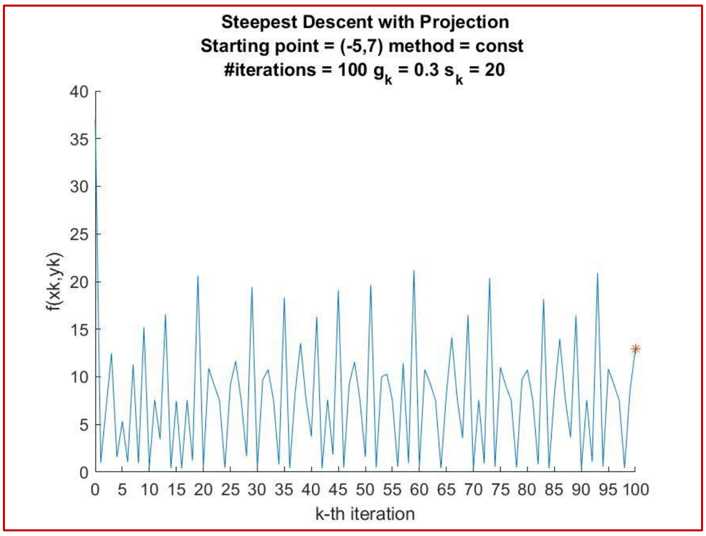
Έχοντας υπόψιν τα παραπάνω, $\gamma_k' = \gamma_k * s_k = 0.1 * 15 = 1.5 < 2$ επομένως ο αλγόριθμος αναμένεται να τερματίσει και να συγκλίνει προς το ελάχιστο. Επιπλέον, όπως φαίνεται και από τα διαγράμματα, η μέθοδος με την προβολή καταλήγει σε ελάχιστο σε λιγότερα βήματα.

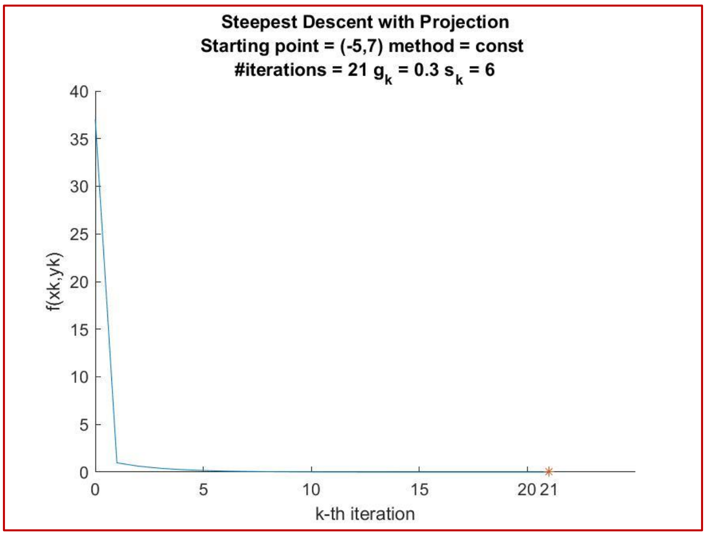




ΘΕΜΑ 3

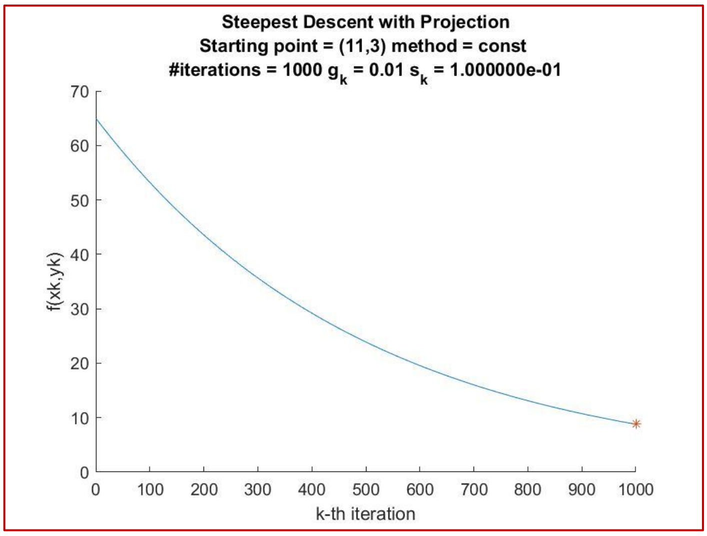
Όμοια, όπως και για το θέμα 2,για το θέμα 3 έχουμε $\gamma_k' = \gamma_k * s_k = 0.3 * 20 = 6 > 2$ επομένως, ο αλγόριθμος αναμένουμε να είναι ασταθής. Ο αλγόριθμος της μέγιστης καθόδου, όπως αποδείχθηκε στο θέμα 1, όταν το γ_k είναι αρκετά μεγαλύτερο από το 2 αποκλίνει συνεχώς από το ελάχιστο (Θέμα 1,iv),και μάλιστα οι τιμές της συνάρτησης αυξάνονται εκθετικά. Παρόμοια συμπεριφορά θα αναμέναμε και από τον αλγόριθμο με προβολή που πράγματι και αυτός είναι ασταθής, όμως ταλαντεύεται εξαιτίας της «διόρθωσης» που επιβάλλει ο αλγόριθμος. Ένας απλός τρόπος για να συγκλίνει η μέθοδος θα ήταν να θέσουμε $s_k = 6$ ώστε $\gamma_k' = \gamma_k * s_k = 6 * 0.3 = 1.8 < 2$, και πράγματι ο αλγόριθμος συγκλίνει όπως φαίνεται και στο δεξιά σχήμα.





ΘΕΜΑ 4

Για το θέμα 4, γνωρίζουμε ότι $\gamma_k' = \gamma_k * s_k = 0.01 * 0{,}1 = 0.001 \ll 2$, άρα ικανοποιεί τους περιορισμούς που θέτει η συνάρτηση που μελετάται. Ωστόσο, η τιμή είναι μόλις λίγο πιο πάνω από το 0 γεγονός που προϊδεάζει πως ο αλγόριθμος θα είναι ευσταθείς, όμως η σύγκλιση του αλγορίθμου θα είναι εξαιρετικά αργή, γεγονός που επιβεβαιώνεται και από τα πειραματικά αποτελέσματα, φαίνεται να τείνει ασυμπωτικά στο ελάχιστο . Η εκτέλεση σταμάτησε εξαναγκασμένα στις 600 επαναλήψεις.



ΣΥΜΠΕΡΑΣΜΑΤΑ/ΣΧΟΛΙΑΣΜΟΣ ΑΠΟΤΕΛΕΣΜΑΤΩΝ

Από την παραπάνω μελέτη 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 γ_k *και* s_k, διότι σε αντίθετη περίπτωση (όπως στο θέμα 4) μπορεί να γίνει απελπιστικά αργός. 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.