Εφαρμοσμένη Βελτιστοποίηση

Ενότητα 6: Αριθμητικές μέθοδοι βελτιστοποίησης για προβλήματα πολλών μεταβλητών ($2^{ο}$ μέρος)

Δημήτρης Παπαδασκαλόπουλος (Επικ. Καθηγητής)
Πολυτεχνική Σχολή
Τμήμα Ηλεκτρολόγων Μηχανικών και Τεχνολογίας Υπολογιστών

Περίληψη προηγούμενων διαλέξεων

- **Ενότητα 1: Εισαγωγή στο μάθημα**
- **Ενότητα 2: Απαιτούμενες μαθηματικές γνώσεις**
- **Ενότητα 3: Αναλυτικές μέθοδοι βελτιστοποίησης για προβλήματα μιας μεταβλητής**
- **Ενότητα 4: Αναλυτικές μέθοδοι βελτιστοποίησης για προβλήματα πολλών μεταβλητών (4 μέρη)**
- **Ενότητα 5: Αριθμητικές μέθοδοι απευθείας έρευνας για προβλήματα μιας μεταβλητής (μέθοδοι απαλοιφής διαστημάτων)**
- **Ενότητα 6 (1^{o} μέρος): Αριθμητικές μέθοδοι απευθείας έρευνας για προβλήματα πολλών μεταβλητών (γεωμετρική έρευνα)**

Σύνδεση προηγούμενων με τη σημερινή διάλεξη (1)

- Οι μέθοδοι επίλυσης προβλημάτων βελτιστοποίησης ταξινομούνται σε δύο βασικές κατηγορίες:

- **Αναλυτικές μέθοδοι (analytical methods)**: η βέλτιστη λύση βρίσκεται με ακρίβεια και μπορεί να εκφραστεί σε “κλειστή μορφή (closed form)”, διατυπώνοντας τις *αναγκαίες και ικανές συνθήκες* για την εύρεσή της > συνήθως για προβλήματα μικρότερης πολυπλοκότητας

Με αυτές ασχοληθήκαμε στις Ενότητες 3 (1 μεταβλητής) και 4 (πολλών μεταβλητών)

- **Αριθμητικές μέθοδοι (numerical methods)**: η βέλτιστη λύση βρίσκεται με κάποια προσέγγιση, συνήθως μέσω κάποιου *επαναληπτικού αλγορίθμου* > συνήθως για προβλήματα μεγαλύτερης πολυπλοκότητας (και όχι μόνο*)

Με αυτές ασχοληθήκαμε στην Ενότητα 5 (1 μεταβλητής) και στην Ενότητα 6 και σήμερα (πολλών μεταβλητών)

Σύνδεση προηγούμενων με τη σημερινή διάλεξη (2)

- Οι αριθμητικές μέθοδοι επίλυσης προβλημάτων βελτιστοποίησης ταξινομούνται σε δύο βασικές κατηγορίες:

➢ **Μέθοδοι απευθείας έρευνας**: επαναληπτική αναζήτηση βέλτιστης λύσης με βάση ΜΟΝΟ την τιμή της αντικειμενικής συνάρτησης για διαφορετικές τιμές των μεταβλητών

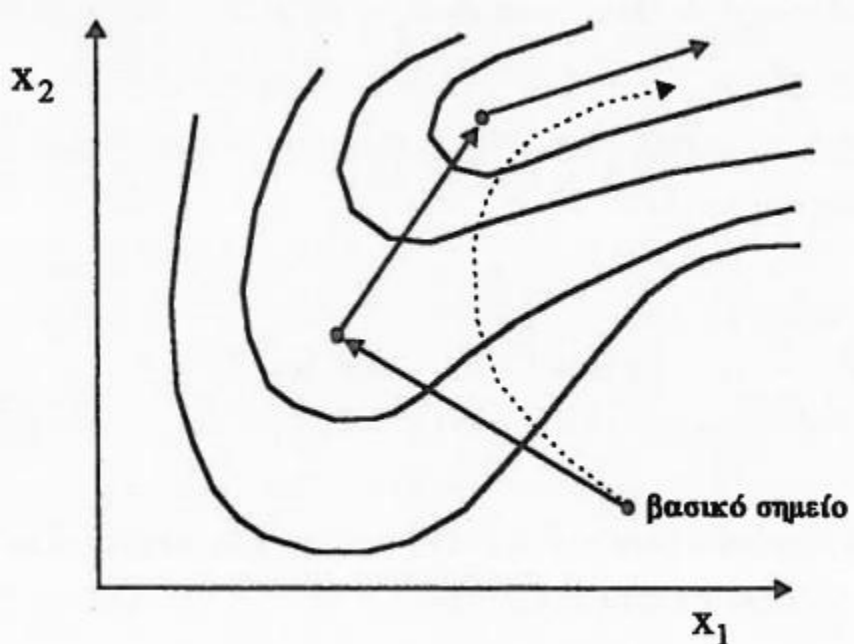
Με αυτές ασχοληθήκαμε στην Ενότητα 5 (1 μεταβλητής)…και 6 (πολλών μεταβλητών)

➢ **Μέθοδοι βάθμωσης (ή μέθοδοι χρήσης παραγώγων)**: επαναληπτική αναζήτηση βέλτιστης λύσης με βάση ΚΑΙ την τιμή παραγώγων της αντικειμενικής συνάρτησης για διαφορετικές τιμές των μεταβλητών

Με αυτές θα ασχοληθούμε σήμερα (κοινές για προβλήματα 1 μεταβλητής και πολλών μεταβλητών)

Μέθοδοι χρήσης παραγώγων (ή μέθοδοι βάθμωσης) (1)

- Επαναληπτική αναζήτηση βέλτιστης λύσης **με βάση και την τιμή παραγώγων** της αντικειμενικής συνάρτησης…
- …επειδή οι παράγωγοι υποδεικνύουν το μεγαλύτερο ρυθμό βελτίωσης της τιμής της συνάρτησης (**κατεύθυνση απότομης ανόδου / καθόδου** αν αναζητώ μέγιστο / ελάχιστο)…
- …δυστυχώς αυτή η ιδιότητα είναι **τοπική και όχι ολική**…



Στην πραγματικότητα, η κατεύθυνση απότομης ανόδου / καθόδου διαφέρει από σημείο σε σημείο…
…ο γεωμετρικός τόπος των άπειρων μικρών κινήσεων κατά μήκος της τοπικής κατεύθυνσης απότομης ανόδου αντιστοιχεί στην διακεκομμένη γραμμή…
…για αυτό και οι μέθοδοι βάθμωσης είναι επαναληπτικές

Μέθοδοι χρήσης παραγώγων (ή μέθοδοι βάθμωσης) (2)

• Οι υπάρχουσες μέθοδοι βάθμωσης διαφοροποιούνται ανάλογα με τον τρόπο που εφαρμόζεται η παραπάνω γενική αρχή….

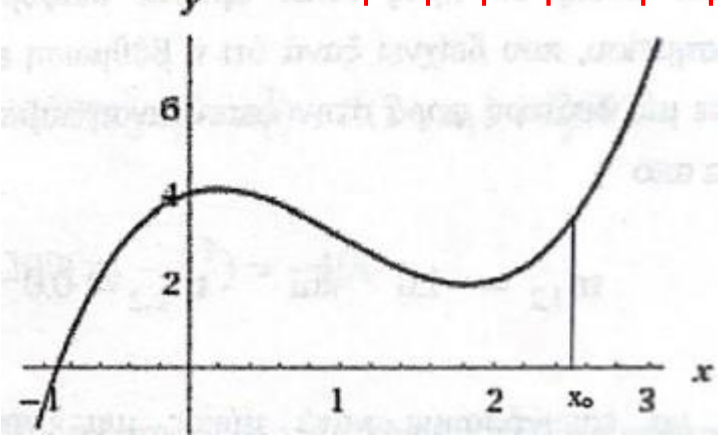
•…θα εξετάσουμε 5 διαφορετικές μεθόδους βάθμωσης:

- **Απλή μέθοδος βάθμωσης**
- **Μέθοδος βάθμωσης με επαρκή συντελεστή**
- **Μέθοδος βάθμωσης με βέλτιστο συντελεστή**
- **Μέθοδος Newton**
- **Τροποποιημένη μέθοδος Newton**

Απλή μέθοδος βάθμωσης (1)

- ΑΠΛΗ ΙΔΕΑ: Για να εντοπίσω τοπικό ελάχιστο (μέγιστο), κάνω βήματα ανάλογα με την αρνητική (θετική) βάθμωση της συνάρτησης στο σημείο που βρίσκομαι

Συνάρτηση 1 μεταβλητής

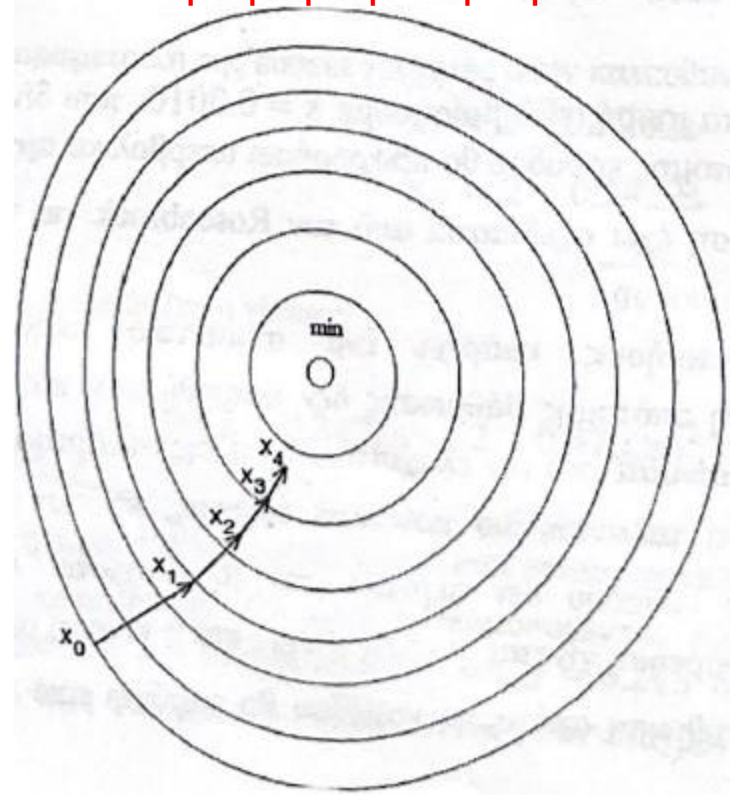


Αν ξεκινήσω από το σημείο $x_0 = 2.5$, ισχύει $f'(x_0) > 0\ldots$

…αν ψάχνω ελάχιστο θα κινηθώ ανάλογα με την $-f'(x_0)$, δηλαδή προς τα αριστερά…

…τελικά θα προσεγγίσω το τοπικό ελάχιστο $x_{min} = 1.8$

Συνάρτηση 2 μεταβλητών



Απλή μέθοδος βάθμωσης (2)

- ΑΠΟΔΕΙΞΗ: Αν Δx_i το βήμα του επαναληπτικού αλγόριθμου, το κάθε επόμενο σημείο είναι $x_{i+1} = x_i + \Delta x_i$…
- …εάν όντως κινούμαστε προς τοπικό ελάχιστο ισχύει $f(x_{i+1}) \leq f(x_i) \rightarrow f(x_i) + df(x_i, \Delta x_i) \leq f(x_i) \rightarrow df(x_i, \Delta x_i) \leq 0 \rightarrow \nabla^T f(x_i) \Delta x_i \leq 0$
- …το οποίο ισχύει αν $\Delta x_i = -\nabla f(x_i)$

- Άρα ο επαναληπτικός αλγόριθμος της απλής μεθόδου βάθμωσης για εύρεση τοπικού ελαχίστου είναι $x_{i+1} = x_i - \nabla f(x_i)$

Παράδειγμα απλής μεθόδου βάθμωσης

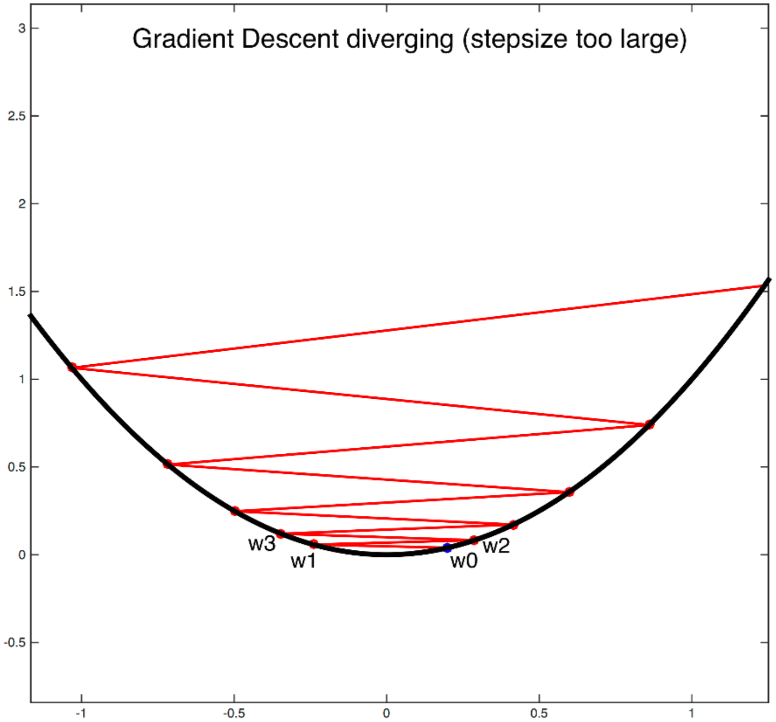
- Τοπικό ελάχιστο συνάρτησης $f(x_1, x_2) = 0.1x_1^2 + 0.25x_2^2 + 4$ με αρχικό σημείο $x_0 = (-12{,}14)$

Βήμα i	x_1	x_2	f	$\frac{\partial f}{\partial x_1}$	$\frac{\partial f}{\partial x_2}$
0	-12	14	67.4	-2.4	7
1	-9.6	7	25.466	-1.92	3.5
2	-7.68	3.5	12.9607	-1.536	1.75
3	-6.144	1.75	8.5405	-1.2288	0.875
…					
13	-0.6597	0.0017	4.0435	-0.1319	0.0009
…					
28	-0.0232	$5.2e^{-08}$	4.0001	-0.0046	$2.6e^{-08}$
29	-0.0186	$2.6e^{-08}$	4.0000	-0.0037	$1.3e^{-08}$
30	-0.0149	$1.3e^{-08}$	4.0000	-0.0030	$6.5e^{-09}$

Σύγκλιση με ακρίβεια 4 δεκαδικών ψηφίων (δηλαδή $f_{i-1} - f_i < 0.0001$) μετά από 30 επαναλήψεις (πίνακας βιβλίου)

Απλή μέθοδος βάθμωσης (3)

- ΠΡΟΒΛΗΜΑ ΑΠΛΗΣ ΜΕΘΟΔΟΥ: Τα βήματα του επαναληπτικού αλγορίθμου είναι πολύ μεγάλα…
- …με αποτέλεσμα φαινόμενα “overshooting” ή ακόμα και απόκλισης (“divergence”)



Μέθοδος βάθμωσης με επαρκή συντελεστή (1)

• ΙΔΕΑ: Εισάγω έναν συντελεστή α_i για να "ελέγξω" το βήμα του επαναληπτικού αλγορίθμου…

•…δηλαδή ο επαναληπτικός αλγόριθμος γίνεται $x_{i+1} = x_i - \alpha_i \nabla f(x_i)$…

•…όπου $0 < \alpha_i \leq 1$…θετικός ώστε να ακολουθώ την κατεύθυνση απότομης καθόδου (για ελάχιστο) και μικρότερος ή ίσος του 1 για να περιορίσω τα μεγάλα βήματα της απλής μεθόδου βάθμωσης…

•…πώς όμως θα "επιλέξω κατάλληλα" το α_i ?

Μέθοδος βάθμωσης με επαρκή συντελεστή (2)

- Χρησιμοποιώντας το ανάπτυγμα Taylor της f μέχρι τον όρο 2ης τάξης (Ενότητα 2)…
- …$f(x_{i+1}) = f\left(x_i - \alpha_i \nabla f(x_i)\right) \cong f(x_i) + \nabla^T f(x_i) \cdot \left(-\alpha_i \nabla f(x_i)\right) + \frac{1}{2}\left(-\alpha_i \nabla^T f(x_i)\right) \cdot \nabla^2 f(x_i) \cdot \left(-\alpha_i \nabla f(x_i)\right) \rightarrow$
$f(x_{i+1}) - f(x_i) \cong -\alpha_i \nabla^T f(x_i) \cdot \nabla f(x_i) + \frac{1}{2}\alpha_i^2 \nabla^T f(x_i) \cdot \nabla^2 f(x_i) \cdot \nabla f(x_i)$…
- …αφού σκοπός είναι να κινηθούμε προς τοπικό ελάχιστο θα πρέπει να ισχύει $f(x_{i+1}) \leq f(x_i) \rightarrow \alpha_i \cdot \left[-\nabla^T f(x_i) \cdot \nabla f(x_i) + \frac{\alpha_i}{2} \nabla^T f(x_i) \cdot \nabla^2 f(x_i) \cdot \nabla f(x_i)\right] \leq 0 \rightarrow \alpha_i \leq \boxed{\frac{2 \cdot \nabla^T f(x_i) \cdot \nabla f(x_i)}{\nabla^T f(x_i) \cdot \nabla^2 f(x_i) \cdot \nabla f(x_i)}}$ Έστω $g(x_i)$

Αν ο (θετικός) συντελεστής α_i πληροί αυτή τη συνθήκη τότε σίγουρα ο αλγόριθμος θα μας οδηγήσει σε τοπικό ελάχιστο…
…και εφόσον βέβαια η βάθμωση της f ($\nabla f(x_i)$) δεν είναι μηδενική και η Hessian της f ($\nabla^2 f(x_i)$) είναι θετικά ορισμένη

Παράδειγμα μεθόδου βάθμωσης με επαρκή συντελεστή (1)

- Τοπικό ελάχιστο συνάρτησης $f(x_1, x_2) = \sin(x_1^2) + 5x_2^2 + 9$ με αρχικό σημείο $x_0 = (3,2)$

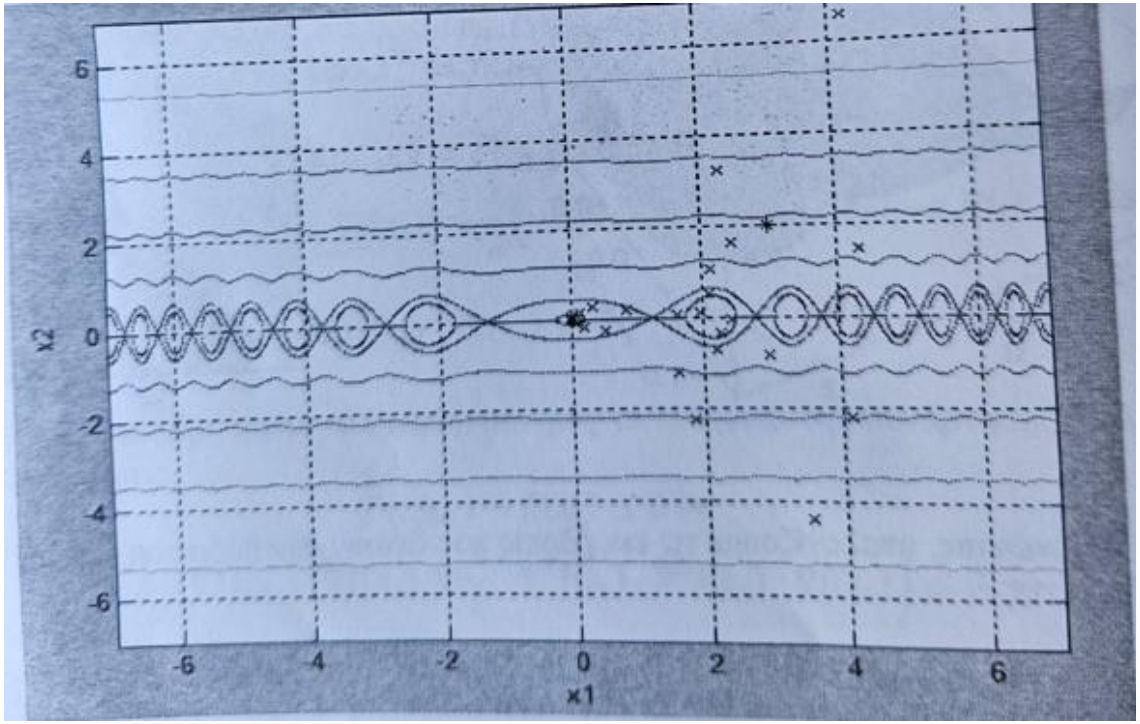
➢ Υπολογίζω τη βάθμωση: $\nabla f(x) = \begin{bmatrix} 2x_1 \cos(x_1^2) \\ 10x_2 \end{bmatrix}$

➢ …και τη Hessian $\nabla^2 f(x) = \begin{bmatrix} 2\cos(x_1^2) - 4x_1^2 \sin(x_1^2) & 0 \\ 0 & 10 \end{bmatrix}$

➢ …επομένως μπορώ να καθορίσω το όριο του επαρκούς συντελεστή $\alpha_i \leq \frac{2 \cdot \nabla^T f(x_i) \cdot \nabla f(x_i)}{\nabla^T f(x_i) \cdot \nabla^2 f(x_i) \cdot \nabla f(x_i)} = g(x_i)$…

Παράδειγμα μεθόδου βάθμωσης με επαρκή συντελεστή (2)

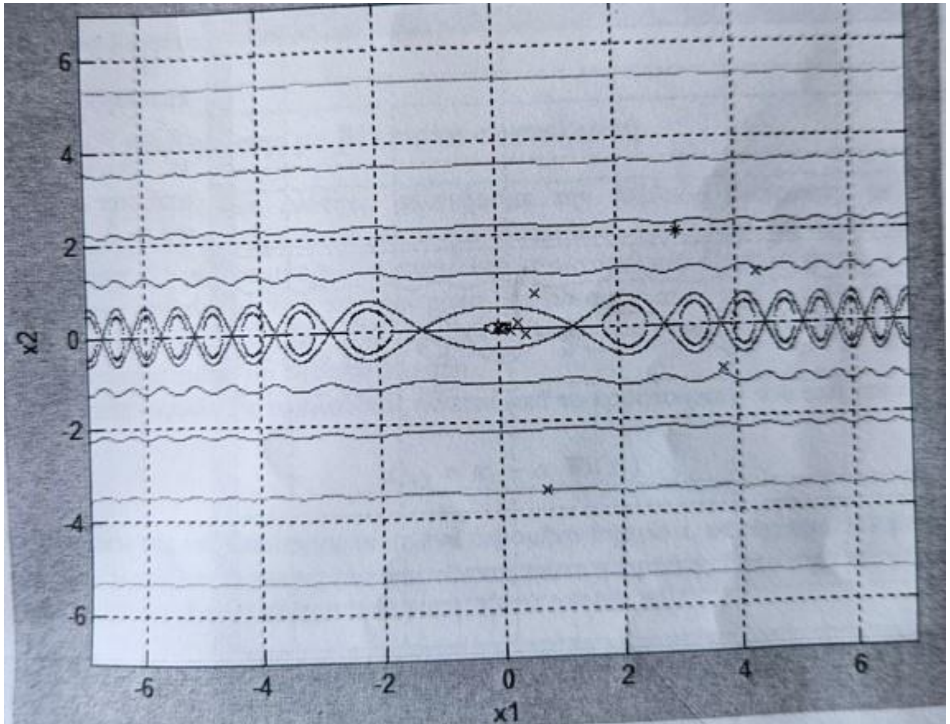
- Τοπικό ελάχιστο συνάρτησης $f(x_1, x_2) = \sin(x_1^2) + 5x_2^2 + 9$ με αρχικό σημείο $x_0 = (3,2)$
  - Με $\alpha_i = 0.85 g(x_i)$, η μέθοδος συγκλίνει στο τοπικό ελάχιστο (0,0) με ακρίβεια 4 δεκαδικών ψηφίων μετά από 35 επαναλήψεις (λάθος στο βιβλίο)



Παράδειγμα μεθόδου βάθμωσης με επαρκή συντελεστή (3)

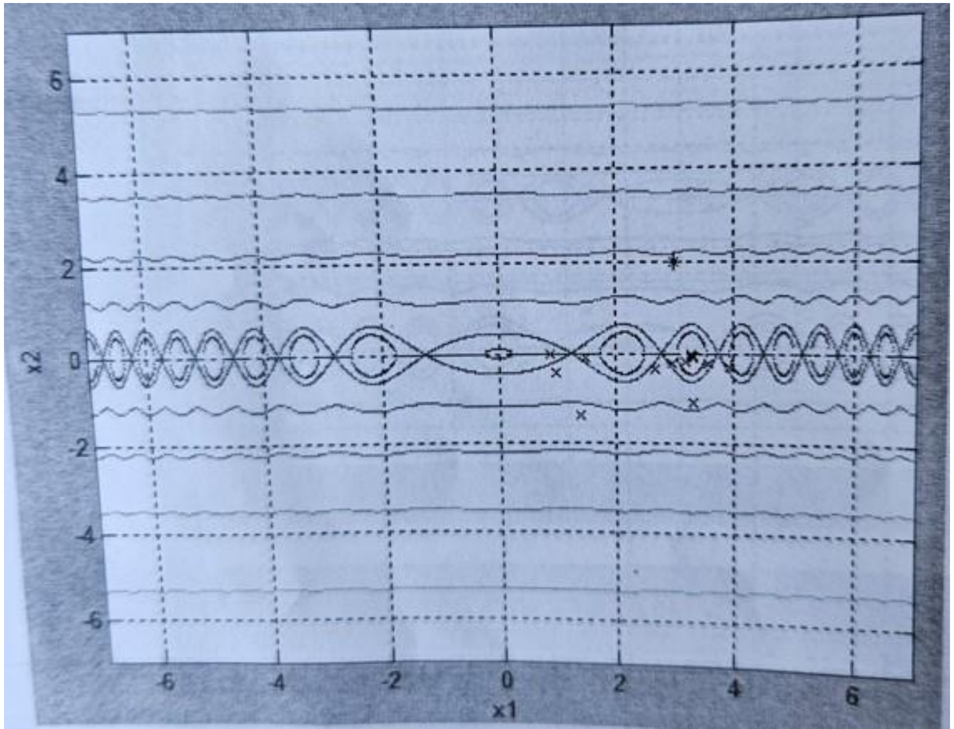
- Τοπικό ελάχιστο συνάρτησης $f(x_1, x_2) = \sin(x_1^2) + 5x_2^2 + 9$ με αρχικό σημείο $x_0 = (3,2)$

➢ Με $\alpha_i = 0.6 g(x_i)$, η μέθοδος συγκλίνει στο τοπικό ελάχιστο (0,0) με ακρίβεια 4 δεκαδικών ψηφίων μετά από 16 επαναλήψεις (λάθος στο βιβλίο)



Παράδειγμα μεθόδου βάθμωσης με επαρκή συντελεστή (4)

- Τοπικό ελάχιστο συνάρτησης $f(x_1, x_2) = \sin(x_1^2) + 5x_2^2 + 9$ με αρχικό σημείο $x_0 = (3{,}2)$
  - Με $\alpha_i = 0.45 g(x_i)$, η μέθοδος συγκλίνει στο τοπικό ελάχιστο (3.32,0) με ακρίβεια 4 δεκαδικών ψηφίων μετά από 19 επαναλήψεις (λάθος στο βιβλίο)



…δηλαδή ο συντελεστής α_i επηρεάζει και την ταχύτητα σύγκλισης αλλά και το συγκεκριμένο τοπικό ελάχιστο που θα βρούμε (αν η συνάρτηση έχει πολλαπλά τοπικά ελάχιστα) !

Μέθοδος βάθμωσης με βέλτιστο συντελεστή (1)

• ΕΡΩΤΗΣΗ: Υπάρχει κάποια “βέλτιστη” τιμή του συντελεστή α_i για την οποία η ταχύτητα σύγκλισης είναι η γρηγορότερη δυνατή?

•…ας θεωρήσουμε πως βρισκόμαστε σε ένα δεδομένο σημείο x_i και αναζητούμε τη βέλτιστη τιμή της (τώρα μεταβλητής!) α_i ώστε η $f(x_{i+1})$ να είναι η ελάχιστη δυνατή…

•…όπου (διαφάνεια 12) $f(x_{i+1}) = f\big(x_i - \alpha_i \nabla f(x_i)\big) \cong f(x_i) - \alpha_i \nabla^T f(x_i) \cdot \nabla f(x_i) + \frac{1}{2}\alpha_i^2 \nabla^T f(x_i) \cdot \nabla^2 f(x_i) \cdot \nabla f(x_i) \ldots$

•…οπότε για να βρω το ελάχιστο της $f(x_{i+1})$ ως προς το α_i

•…$\frac{\partial f(x_{i+1})}{\partial \alpha_i} = 0 \rightarrow \frac{\partial f\left(x_i - \alpha_i \nabla f(x_i)\right)}{\partial \alpha_i} = 0 \rightarrow -\nabla^T f(x_i) \cdot \nabla f(x_i) + \alpha_i \nabla^T f(x_i) \cdot \nabla^2 f(x_i) \cdot \nabla f(x_i) = 0 \rightarrow \boxed{\alpha_i = \frac{\nabla^T f(x_i) \cdot \nabla f(x_i)}{\nabla^T f(x_i) \cdot \nabla^2 f(x_i) \cdot \nabla f(x_i)}}$ $\frac{1}{2} g(x_i)$

…ο βέλτιστος συντελεστής είναι ίσος του μισού του άνω ορίου του επαρκή !

Παράδειγμα μεθόδου βάθμωσης με βέλτιστο συντελεστή (1)

- Τοπικό ελάχιστο συνάρτησης $f(x_1, x_2) = x_1^3 + 10x_1^2 + 4x_2^2 + 7x_1 + 20$ με αρχικό σημείο $x_0 = (3,3)$

➢ Υπολογίζω τη βάθμωση: $\nabla f(x) = \begin{bmatrix} 3x_1^2 + 20x_1 + 7 \\ 8x_2 \end{bmatrix}$

➢ …και τη Hessian $\nabla^2 f(x) = \begin{bmatrix} 6x_1 + 20 & 0 \\ 0 & 8 \end{bmatrix}$

➢ …επομένως μπορώ να καθορίσω το βέλτιστο συντελεστή $\alpha_i = \frac{\nabla^T f(x_i) \cdot \nabla f(x_i)}{\nabla^T f(x_i) \cdot \nabla^2 f(x_i) \cdot \nabla f(x_i)}$…

Παράδειγμα μεθόδου βάθμωσης με βέλτιστο συντελεστή (2)

- Τοπικό ελάχιστο συνάρτησης $f(x_1, x_2) = x_1^3 + 10x_1^2 + 4x_2^2 + 7x_1 + 20$ με αρχικό σημείο $x_0 = (3, -3)$

Βήμα i	x_1	x_2	f	α_i
0	3	-3	194	0.0277
1	0.4007	-2.3364	46.3096	0.0721
2	-0.7167	-0.9886	23.6607	0.0936
3	-0.1747	-0.2487	19.3242	0.0610
4	-0.3942	-0.1273	18.7981	0.1065
5	-0.3497	-0.0189	18.7337	0.0606
6	-0.3723	-0.0097	18.7287	0.1080
7	-0.3691	-0.0013	18.7284	0.0605
8	-0.3707	-6.8e^{-04}	18.7283	0.1081

Σύγκλιση με ακρίβεια 4 δεκαδικών ψηφίων μετά από 8 επαναλήψεις

Μέθοδος Newton (1)

- Επαναληπτικός αλγόριθμος που χρησιμοποιείται στην αριθμητική ανάλυση για την εύρεση των ριζών μιας διαφορίσιμης συνάρτησης f (δηλαδή των λύσεων της εξίσωσης $f(x) = 0$)…
- …αφού γνωρίζουμε πως αναγκαία και ικανή συνθήκη για την εύρεση τοπικού ακροτάτου είναι η $f'(x) = 0$ (για προβλήματα μιας μεταβλητής, Ενότητα 3) ή $\nabla f(x) = 0$ (για προβλήματα πολλών μεταβλητών, Ενότητα 4)…
- …μπορούμε να χρησιμοποιήσουμε τη μέθοδο Newton για την εύρεση ακροτάτων, εφαρμόζοντάς την στην παράγωγο f' μιας συνάρτησης μιας μεταβλητής (ή γενικά στη βάθμωση μιας συνάρτησης πολλών μεταβλητών) για να βρούμε αριθμητικά τις λύσεις της $f'(x) = 0$…
- …και άρα τα ακρότατα της f !

Μέθοδος Newton (2)

- Όπως και στην απλή μέθοδο βάθμωσης (διαφάνεια 10), αν Δx_i το βήμα του επαναληπτικού αλγόριθμου, το κάθε επόμενο σημείο είναι $x_{i+1} = x_i + \Delta x_i \ldots$
- …χρησιμοποιώντας το ανάπτυγμα Taylor της f μέχρι τον όρο 2ης τάξης (Ενότητα 2)…
- $\ldots f(x_{i+1}) = f(x_i + \Delta x_i) \cong f(x_i) + \nabla^T f(x_i) \cdot \Delta x_i + \frac{1}{2} \Delta {x_i}^T \cdot \nabla^2 f(x_i) \cdot \Delta x_i \ldots$
- …θεωρώντας πως βρισκόμαστε σε ένα δεδομένο σημείο x_i και θέλουμε να καθορίσουμε το βέλτιστο βήμα Δx_i (ώστε η $f(x_{i+1})$ να είναι η ελάχιστη δυνατή), η ικανή συνθήκη είναι…
- $\ldots \frac{df(x_{i+1})}{d\Delta x_i} = 0 \rightarrow \nabla f(x_i) + \boxed{\nabla^2 f(x_i)} \cdot \Delta x_i = 0 \ldots$ η Hessian $H(x_i)$ της f
- $\ldots H(x_i) \cdot \Delta x_i = -\nabla f(x_i) \rightarrow \Delta x_i = -H^{-1}(x_i) \cdot \nabla f(x_i)$

Μέθοδος Newton (3)

- Άρα ο επαναληπτικός αλγόριθμος της μεθόδου Newton για εύρεση τοπικού ελαχίστου είναι $x_{i+1} = x_i - H^{-1}(x_i) \cdot \nabla f(x_i)$
- …η μέθοδος είναι εφαρμόσιμη μόνο αν η Hessian είναι αντιστρέψιμη…
- …και μας οδηγεί στο τοπικό ελάχιστο μόνο αν η Hessian είναι θετικά ορισμένη…

Ένας πίνακας είναι αντιστρέψιμος αν και μόνο αν η (τελευταία) ορίζουσά του δεν είναι ίση με 0…
…όλες οι ορίζουσες (και η τελευταία) ένας θετικά ορισμένου πίνακα είναι θετικές…
…άρα οι παραπάνω 2 συνθήκες εφαρμογής της μεθόδου Newton απλοποιούνται σε 1 μόνο συνθήκη: η Hessian να είναι θετικά ορισμένη

Παράδειγμα μεθόδου Newton (1)

- Τοπικό ελάχιστο συνάρτησης $f(x_1, x_2) = 3x_1^2 + 9x_2^2 + 5\sin x_1 + 7.5\cos x_2 + 17$ με αρχικό σημείο $x_0 = (3, -2)$
  - Υπολογίζω τη βάθμωση: $\nabla f(x) = \begin{bmatrix} 6x_1 + 5\cos x_1 \\ 18x_2 - 7.5\sin x_2 \end{bmatrix}$
  - …και τη Hessian $\nabla^2 f(x) = \begin{bmatrix} 6 - 5\sin x_1 & 0 \\ 0 & 18 - 7.5\cos x_2 \end{bmatrix}$
  - …με ορίζουσες $D_1 = 6 - 5\sin x_1 > 0$ και $D_2 = (6 -$

Παρένθεση: Αντιστροφή πινάκων (1)

- Ορισμός αντίστροφου πίνακα P^{-1} ενός τετραγωνικού πίνακα P: είναι εκείνος για τον οποίο ισχύει η σχέση $P \cdot P^{-1} = \boxed{I}$ όπου I τετραγωνικός μοναδιαίος πίνακας διαστάσεων ίσων με του P
- Πώς αντιστρέφω έναν πίνακα 2x2 $P = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$?
- Μέσω του τύπου $P^{-1} = \frac{1}{\boxed{ad-bc}} \boxed{\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}}$ ο συζυγής πίνακας του P

η τελευταία ορίζουσα του P

- Πώς αντιστρέφω ένα μεγαλύτερο πίνακα 3x3, 4x4 κλπ. ?

…πιο δύσκολος τύπος…

- …αλλά εάν ο πίνακας είναι της μορφής $P = \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$ τότε μπορώ να χρησιμοποιήσω τον τύπο $P^{-1} = \begin{bmatrix} A^{-1} & 0 \\ 0 & B^{-1} \end{bmatrix}$

Παρένθεση: Αντιστροφή πινάκων (2)

- Παράδειγμα: Να βρεθεί ο αντίστροφος του πίνακα $P = \begin{bmatrix} 22 & -20 & 0 \\ -20 & 20 & 0 \\ 0 & 0 & 2 \end{bmatrix}$

➢ Είναι πίνακας 3x3 αλλά ευτυχώς είναι της μορφής $P = \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$ με $A = \begin{bmatrix} 22 & -20 \\ -20 & 20 \end{bmatrix}$ και $B = [2]$

➢ …άρα μπορώ να εφαρμόσω τον τύπο $P^{-1} = \begin{bmatrix} A^{-1} & 0 \\ 0 & B^{-1} \end{bmatrix}$…

➢ …όπου $A^{-1} = \frac{1}{22 \cdot 20 - (-20) \cdot (-20)} \begin{bmatrix} 20 & 20 \\ 20 & 22 \end{bmatrix} = \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 11/20 \end{bmatrix}$…

➢ …και $B^{-1} = 1/2 \ldots$

➢ …τελικά ο αντίστροφος είναι $P^{-1} = \begin{bmatrix} 1/2 & 1/2 & 0 \\ 1/2 & 11/20 & 0 \\ 0 & 0 & 1/2 \end{bmatrix}$

Παράδειγμα μεθόδου Newton (2)

- Τοπικό ελάχιστο συνάρτησης $f(x_1, x_2) = 3x_1^2 + 9x_2^2 + 5\sin x_1 + 7.5\cos x_2 + 17$ με αρχικό σημείο $x_0 = (3, -2)$

➢ Η Hessian υπολογίστηκε (διαφάνεια 23) ως $\mathrm{H}(x) = \begin{bmatrix} 6 - 5\sin x_1 & 0 \\ 0 & 18 - 7.5\cos x_2 \end{bmatrix}$

➢ …εφαρμόζοντας τον τύπο για τον αντίστροφο πίνακα 2x2 λαμβάνω $H^{-1}(x) = \frac{1}{(6-5\sin x_1)\cdot(18-7.5\cos x_2)-0\cdot 0}\begin{bmatrix} 18 - 7.5\cos x_2 & 0 \\ 0 & 6 - 5\sin x_1 \end{bmatrix} = \begin{bmatrix} \frac{1}{6-5\sin x_1} & 0 \\ 0 & \frac{1}{18-7.5\cos x_2} \end{bmatrix}$

Παράδειγμα μεθόδου Newton (3)

- Τοπικό ελάχιστο συνάρτησης $f(x_1, x_2) = 3x_1^2 + 9x_2^2 + 5\sin x_1 + 7.5\cos x_2 + 17$ με αρχικό σημείο $x_0 = (3, -2)$
  - Άρα εφαρμόζω τον επαναληπτικό αλγόριθμο της μεθόδου Newton $x_{i+1} = x_i - H^{-1}(x_i) \cdot \nabla f(x_i) \rightarrow x_{i+1} = x_i - \begin{bmatrix} \frac{1}{6-5\sin x_1} & 0 \\ 0 & \frac{1}{18-7.5\cos x_2} \end{bmatrix} \cdot \begin{bmatrix} 6x_1 + 5\cos x_1 \\ 18x_2 - 7.5\sin x_2 \end{bmatrix}$

Βήμα *i*	x_1	x_2	f
0	3	-2	77.5845
1	0.5351	-0.6184	29.9618
2	-1.6420	-0.0479	27.6136
3	-0.7130	-2.6e^{-05}	22.7546
4	-0.6595	-4.2e^{-05}	22.7412
5	-0.6589	0	22.7412

Σύγκλιση με ακρίβεια 4 δεκαδικών ψηφίων μετά από 5 επαναλήψεις

Μέθοδος Newton (4)

- ΠΛΕΟΝΕΚΤΗΜΑΤΑ: Είναι η ταχύτερη από όλες τις μεθόδους βάθμωσης που εξετάσαμε σήμερα…
- …και ειδικά για τετραγωνικές μονοτροπικές συναρτήσεις βρίσκει το (μοναδικό) ελάχιστο με μηδενικό σφάλμα σε μία μόνο επανάληψη και ανεξάρτητα από το αρχικό σημείο…
- …επειδή η Hessian είναι σταθερός πίνακας…δηλαδή ανεξάρτητα από το αρχικό σημείο ενσωματώνει την απαραίτητη πληροφορία για τη θέση του ελαχίστου

Παράδειγμα μεθόδου Newton (4)

- Τοπικό ελάχιστο συνάρτησης $f(x_1, x_2) = ax_1^2 + bx_2^2 + c$ με $a, b, c > 0$ παραμέτρους και (γενικό) αρχικό σημείο $x_0 = (x_{0,1}, x_{0,2})$

➢ Υπολογίζω τη βάθμωση: $\nabla f(x) = \begin{bmatrix} 2ax_1 \\ 2bx_2 \end{bmatrix}$

➢ …και τη Hessian $\nabla^2 f(x) = \begin{bmatrix} 2a & 0 \\ 0 & 2b \end{bmatrix}$

➢ …με ορίζουσες $D_1 = 2a > 0$ και $D_2 = 4ab > 0 \ldots$

➢ …άρα η Hessian είναι πάντα θετικά ορισμένη και επομένως η μέθοδος Newton είναι εφαρμόσιμη…

➢ …και ο αντίστροφος της Hessian είναι $H^{-1}(x) = \begin{bmatrix} \frac{1}{2a} & 0 \\ 0 & \frac{1}{2b} \end{bmatrix}$

Παράδειγμα μεθόδου Newton (5)

- Τοπικό ελάχιστο συνάρτησης $f(x_1, x_2) = ax_1^2 + bx_2^2 + c$ με $a, b, c > 0$ παραμέτρους και (γενικό) αρχικό σημείο $x_0 = (x_{0,1}, x_{0,2})$

➢ Άρα εφαρμόζω τον επαναληπτικό αλγόριθμο της μεθόδου Newton $x_{i+1} = x_i - H^{-1}(x_i) \cdot \nabla f(x_i) \rightarrow \begin{bmatrix} x_{1,1} \\ x_{1,2} \end{bmatrix} = \begin{bmatrix} x_{0,1} \\ x_{0,2} \end{bmatrix} - \begin{bmatrix} \frac{1}{2a} & 0 \\ 0 & \frac{1}{2b} \end{bmatrix} \cdot \begin{bmatrix} 2ax_{0,1} \\ 2bx_{0,2} \end{bmatrix} = \begin{bmatrix} x_{0,1} \\ x_{0,2} \end{bmatrix} - \begin{bmatrix} x_{0,1} \\ x_{0,2} \end{bmatrix} = \boxed{\begin{bmatrix} 0 \\ 0 \end{bmatrix}}$

Το οποίο είναι όντως το (μοναδικό) τοπικό ελάχιστο…

…δηλαδή, όπως προβλέπει η θεωρία, επειδή η συνάρτηση είναι τετραγωνική μονοτροπική η μέθοδος Newton βρίσκει το (μοναδικό) τοπικό ελάχιστο με μηδενικό σφάλμα σε μία μόνο επανάληψη και ανεξάρτητα από το αρχικό σημείο

Μέθοδος Newton (5)

• ΜΕΙΝΟΝΕΚΤΗΜΑ: Απαιτεί την αντιστροφή πινάκων…

•…η οποία εμπεριέχει μεγάλο υπολογιστικό φορτίο…

•…και η οποία σε κάποιες περιπτώσεις δεν είναι δυνατή (αν η Hessian δεν είναι αντιστρέψιμη, διαφάνεια 22)

•…για να αντιμετωπιστεί το τελευταίο πρόβλημα, έχει αναπτυχθεί η **τροποποιημένη μέθοδος Newton**…

•…όπου η (μη-αντιστρέψιμη) Hessian αντικαθίσταται με τον (αντιστρέψιμο) πίνακα $M(x_i) = H(x_i) + \mu_i \cdot I_n$, όπου μ_i κατάλληλη παράμετρος ώστε η $M(x_i)$ να είναι αντιστρέψιμη

Τέλος Ενότητας