Βελτιστοποίηση ΙΙ

5. Το βήμα 2.β μπορεί να αναδιατυπωθεί:

"β. Διαφορετικά υπολογίζουμε τον Hessian Matrix $\nabla^2 f(x_k)$ και την κατεύθυνση ανάμησης p_k ως λύση του συστήματος:

$$[\nabla^2 f(x_k)]\, p_k = -\nabla f(x_k)$$ (Υπολογιστικά ευκολότερο)

Παράδειγμα 4.4

Εφαρμόστε Newton για $\alpha_k = 1$, $x_0 = [1 \; 1]^T$

minimize $f(x_1, x_2) = 5x_1^4 + 6x_2^4 - 6x_1^2 + 2x_1x_2 + 5x_2^2 + 15x_1 - 7x_2 + 13$

Λύση

$f_{x_1} = 20x_1^3 - 12x_1 + 2x_2 + 15$

$f_{x_2} = 24x_2^3 + 2x_1 + 10x_2 - 7$

$f_{x_1x_1} = 60x_1^2 - 12$

$f_{x_1x_2} = 2$

$f_{x_2x_1} = 2$

$f_{x_2x_2} = 72x_2^2 + 10$

Τελικά λαμβάνουμε $x^* = [-1.1421 \;\; 0.5434]^T$

με $f(x^*) = -6.4961$

$\nabla f(x^*) = 5.4026 \cdot 10^{-15}$

και $\nabla^2 f(X^*) = \begin{bmatrix} 66.2635 & 2 \\ 2 & 31.2604 \end{bmatrix}$

$\Delta_1 > 0$, $\Delta_2 > 0$ άρα $\nabla^2 f(X^*)$ θετικά ορισμένος

και άρα X^* → τοπικός ελαχιστοποιητής.

$\boxed{4.3}$

$\boxed{4.3}$ Υλοποίηση της μεθόδου Newton

Υπολογίζω p_k από $(\nabla^2 f(x_k))\, p_k = -\nabla f(x_k)$
για δοθέν x_k.

Θυμόμαστε $Ax = b$.

$A = LU$ * → άνω (upper) triangle. γρήγορο
↓ κάτω τριγωνικός

$\Big\} \Rightarrow LUx = b \Rightarrow \Big\{ \begin{array}{l} Ly = b \\ Ux = y. \end{array}$

* αν A θετικά ορισμένος
↳ $A = LL^T$ // προσδιορίζω τον πίνακα L στήλη - στήλη
↓ lower triangle.

!!! Αν $\nabla^2 f(x_k)$ δεν είναι θετικά ορισμένος κάνουμε $\Big\{ \nabla^2 f(x_k) + E = LL^T$
↓ κατάλληλος διαγώνιος πίνακας

λύω $(\nabla^2 f(x_k) + E)\, p_k = -\nabla f(x_k)$

Παράδειγμα 4.5

Βρείτε την κατεύθυνση p της μεθόδου Newton στο $x_0 = [0\ 1]^T$ με Cholesky

ελαχιστοποίησε $f(x_1, x_2) = x_1^4 + 2x_1^3 + 2x_1^2 + x_2^2 - 2x_1 x_2$.

$f_{x_1} = 4x_1^3 + 6x_1^2 + 4x_1 - 2x_2$ $\qquad f_{x_2} = 2x_2 - 2x_1$

$f_{x_1x_1} = 12x_1^2 + 12x_1 + 4$ $\qquad f_{x_2x_2} = 2$

$f_{x_1x_2} = -2$ $\qquad f_{x_2x_1} = -2$

$$\nabla f(x_0) = [-2 \; 2]^T$$

$$\nabla^2 f(x_0) = \begin{bmatrix} 4 & -2 \\ -2 & 2 \end{bmatrix} = LL^T = \begin{bmatrix} 2 & 0 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ 0 & 1 \end{bmatrix}$$

$$LL^T p = -\nabla f(x_0)$$

$$Ly = -f(x_0) \Rightarrow y = [1 \; -1]^T$$

$$L^T p = y \Rightarrow p = [0 \; -1]^T$$

$\boxed{4.4}$ Ταχύτητες σύγκλισης Μεθόδων Steepest Descent και Newton

St. Des. για την τετραγ. συνάρτηση

$$f(x) = \frac{1}{2} x^T Q x - b^T x + c$$

$$\nabla f(x) = Qx - b \quad , \text{ τοπ. βελτιστοποίηση } \quad Qx = b$$

με βέλτιστο $a_k = -\dfrac{p_k^T \nabla f(x_k)}{p_k^T Q p_k}$, για $p_k = -\nabla f(x_k)$

για $x_{k+1} = x_k - \cancel{\;} a_k p_k \Rightarrow x_{k+1} - \dfrac{\nabla f(x_k)^T \nabla f(x_k)}{\nabla f(x_k)^T Q \nabla f(x_k)} \nabla f(x_k)$?

με $\nabla f(x) = Qx - b$

Θεώρημα 4.1

~~[illegible]~~ 2) ~~[illegible]~~ Steepest Descent στο τετραγωνική συνάρτηση συγκλίνει γραμμικά στο μοναδικό τοπικό ελαχιστοποιητή x^* της f για κάθε αρχικό σημείο x_0.

Θεώρημα 4.2 (τοπική σύγκλιση)

Τετραγωνική σύγκλιση της μεθόδου Newton

για $x_{k+1} = x_k - [\nabla^2 f(x_k)]^{-1} \nabla f(x_k)$

για $\nabla^2 f(x^*)$ θετικά ορισμένο, για κάθε αρχικό σημείο

x_0 "αρκετά κοντά" στο x^*, η ακολουθία x_k συγκλίνει στο κατάλληλα

x^* με τετραγωνικό ρυθμό. (αν τα βήματα επιλεγεί πάντα ίσο με 1)

Βελτιστοποίηση με Δεσμεύσεις / Περιορισμούς

Περιορισμοί $\begin{cases} \text{ισοτικοί} \\ \text{-//- + ανισοτικοί} \end{cases}$

Παράδειγμα (δύο μεταβλητές)

Γινόμενο $x \cdot y$ δύο θετικών πραγματικών αριθμών με σταθερό άθροισμα $x+y = \alpha > 0$ γίνεται μέγιστο όταν $x = y = \frac{\alpha}{2}$

Λύση

Λύνουμε την $g(\vec{x}) = g(x,y) = x+y-\alpha = 0$ ως προς y.

άρα $y = \alpha - x$ και αφού $y > 0$ $\alpha - x > 0 \Rightarrow 0 < x < \alpha$

και αναζητούμε το (αδέσμευτο) μέγιστο της συνάρτησης $f(x) = x(\alpha - x)$, $0 < x < \alpha$.

$$f'(x) = \alpha - 2x = 2\left(\frac{\alpha}{2} - x\right)$$

με $f'\left(\frac{\alpha}{2}\right) = 0$ $\begin{cases} f'(x) > 0 & \text{για } 0 < x < \alpha/2 \quad \text{// } f \text{ γν. αύξουσα} \\ f'(x) < 0 & \text{για } \alpha/2 < x < \alpha \quad \text{// } f \text{ γν. φθίνουσα} \end{cases} \Rightarrow$

$\Rightarrow$ { άρα στο $x = \alpha/2$ η f λαμβάνει ολικό μέγιστο

άρα $x^* = [x \;\; y] = \left[\frac{\alpha}{2} \;\; \frac{\alpha}{2}\right]$ η συνάρτηση $f(x,y) = x \cdot y$ λαμβάνει ολ. max υπό την ισοτική δέσμευση $g(x,y) = x+y-\alpha = 0$ }

Παράδειγμα (3 μεταβλητές)

Το γινόμενο xyz τριών θετικών πραγματικών αριθμών x, y, z με σταθερό άθροισμα $x+y+z = \alpha > 0$ γίνεται μέγιστο όταν $x = y = z = \alpha/3$

Έχω σαν δύο περιορισμούς, βάζω τον ισοτ. περιορισμό στην συνάρτηση. Δεν τις περιορ. χωρίς είναι εφικτό.

Λύση

$$g(\vec{x}) = g(x, y, z) = x+y+z-\alpha = 0$$

Λύνω ως προς z : $z = \alpha - x - y$

και αναζητώ (ελεύθερο) μέγιστο της συνάρτησης $f(x,y) = xy(\alpha - x - y) \Rightarrow$

$\Rightarrow f(x,y) = \alpha xy - x^2 y - xy^2 \quad , x > 0 , y > 0.$

$f_x = \alpha y - 2xy - y^2 \qquad f_y = \alpha x - x^2 - 2xy$

$f_{xx} = -2y \quad , f_{xy} = \alpha - 2x - 2y \quad , f_{yy} = -2x$

$f_{yx} = \alpha - 2x - 2y.$

είναι συνεχείς.

Το σύστημα $(\nabla f(x,y) = \vec{0})$ $[f_x \ \ f_y] = [0 \ \ 0]$ έχει τη μοναδική λύση

$$[x \ \ y] = \left[\frac{\alpha}{3} \ \ \frac{\alpha}{3}\right] = \vec{x}^*$$

Στο x^* υπολογίζω τις ορίζουσες του $\nabla^2 f(x,y)$

$\Delta_1 = -\dfrac{2\alpha}{3} < 0$ και $\Delta_2 = \dfrac{\alpha^2}{3} > 0$

άρα $x^* = \left[\frac{\alpha}{3} \ \ \frac{\alpha}{3}\right] \rightarrow$ τοπ. μεγιστοποιητής της f.

άρα τελικά το $[x \;\; y \;\; z] = [\frac{a}{3} \;\; \frac{a}{3} \;\; \frac{a}{3}]$ είναι ο δεσμευμένος τοπ. μέγιστ. της $f(x,y,z)$ υπό την ισοτική δέσμευση $g(x,y,z)=0$.

→ Μπορούμε να δείξουμε ότι το τοπικό αυτό μέγιστο, είναι και ολικό.

→ Γενικά δηλαδή επιλύουμε την g ως προς x_n και κάνουμε αντικατάσταση στη f.
Τις περισσότερες φορές αυτό δεν είναι εφικτό.

Παράδειγμα.

minimize $x_1 + x_2$

subject to $x_1^2 + x_2^2 - 2 = 0$.

Λύση

$f(x) = x_1 + x_2$, $I = \emptyset$, $E = \{1\}$, $g_1(\vec{x}) = x_1^2 + x_2^2 - 2 = 0$.

Ω του προβλήματος: κύκλος με κέντρο το $(0,0)$ και ακτίνα $\sqrt{2}$

λύση $x^* = [-1 \;\; -1]^T$ βλ. σχήμα...

Θ Θεώρημα: Αναγκαία Συνθήκη Πρώτης Τάξης.
Έστω δύο C^1 πραγμ. συναρτήσεις $f = f(\vec{x})$, $g = g(\vec{x})$
$U \subseteq R^n \to R$ με U ανοικτό.

Υποθέτουμε ότι η f έχει στο $x^* \in U$ δεσμευμένο τοπικό βελτιστοποιητή, υπό τη δέσμευση $g(\vec{x}) = 0$, και $\nabla g(\vec{x}^*) \neq 0$.
Τότε υπάρχει $\lambda^* \in R$ τ.ω. $\nabla f(\vec{x}^*) = \lambda^* \nabla g(\vec{x}^*)$ και $g(\vec{x}^*) = 0$.
ο λ^* ονομάζεται πολλαπλασιαστής Lagrange.

Ορισμός : Δεσμευμένο κρίσιμο σημείο.

Έστω δύο C^1 πραγμ. συναρτ. $f=f(\vec{x})$ και $g=g(\vec{x})$: $U\subseteq \mathbb{R}^n \to \mathbb{R}$

Ένα σημείο x^* ονομάζεται δεσμευμένο κρίσιμο σημείο (U ανοικτό) της f υπό τη δέσμευση $g(\vec{x})=0$ όταν υπάρχει λ^* τ.ω. το σημείο $(\vec{x}^*, \lambda^*)$ να είναι λύση του συστήματος.

$$\boxed{\nabla f(\vec{x}) = \lambda \nabla g(\vec{x}) \text{ και } g(\vec{x}) = 0}$$

Άρα (βάσει προηγ. θεωρήματος) τα σημεία δεσμευμένων τοπικών βελτιστοποιητών βρίσκονται μεταξύ των δεσμευμένων κρίσιμων σημείων και των σημείων και των σημείων $\vec{x}\in U$ για τα οποία η f και/η g δεν είναι C^1 ή $\nabla g(\vec{x}) = \vec{0}$.

// Επαναδιατύπωση θεωρήματος //

Θεώρημα : Αναγκαία Συνθήκη Πρώτης Τάξης.

f, g C^1 συναρτήσεις στο $U\subseteq\mathbb{R}^n\to\mathbb{R}$, (U ανοικτό)

Αν η f έχει στο $\vec{x}^*$ δεσμ. τοπ. βελτιστοπ. υπό $g(\vec{x})=0$ και ισχύει $\nabla g(\vec{x}^*)\neq 0$ τότε το $\vec{x}^*$ είναι δεσμευμένο κρίσιμο σημείο.

Θεώρημα.

Έστω δύο C^1 διαφορίσιμες συναρτήσεις,

$f = f(\vec{x})$ και $g = g(\vec{x}) : U \subseteq R^n \to R$ (U ανοιχτό)

Υποθέτουμε ότι το κλειστό υποσύνολο $\Omega = \{\vec{x} \in U : g(\vec{x}) = 0\}$ του U είναι φραγμένο και ικανοποιεί $\nabla g(\vec{x}) \neq 0$, $\forall x \in \Omega$.

Τότε η f έχει δεσμευμένη μέγιστη και ελάχιστη τιμή, υπό τη δέσμευση $g(\vec{x}) = 0$

Δηλαδή υπάρχουν δύο δεσμευμένα κρίσιμα σημεία

$\vec{x}_\varepsilon^*$ και $\vec{x}_\mu^*$ με την ιδιότητα:

$$f(\vec{x}_\varepsilon^*) \le f(\vec{x}) \le f(\vec{x}_\mu^*), \quad \forall x \in \Omega$$

Οι τιμές $f(\vec{x}_\varepsilon^*)$ και $f(\vec{x}_\mu^*)$ ευρίσκονται με σύγκριση των τιμών $f(\vec{x}^*)$ στα δεσμευμένα κρίσιμα σημεία x^*.

→ η δεσμευμένη μέγιστη/ελάχιστη τιμή, είναι η μεγαλύτερη/μικρότερη από τις τιμές $f(\vec{x}^*)$ στα δ.κ.σ. $\vec{x}^*$ της f, υπό τη δέσμευση $g(\vec{x}) = 0$.

Παράδειγμα.

Βρείτε τη δεσμευμένη μέγιστη και ελάχιστη τιμή της συνάρτησης $f(x_1, x_2) = x_1^3 + x_2^2$ υπό τη δέσμευση

$$g(x_1, x_2) = x_1^2 + x_2^2 - 4 = 0.$$

Λύση →

Λύση

f, g C^2 στο R^2

$\nabla f(x_1, x_2) = 3[\, x_1^2 \quad x_2^2 \,]$

$\nabla g(x_1, x_2) = 2[\, x_1 \quad x_2 \,]$

$$\boxed{\nabla f(x_1, x_2) = \lambda \nabla g(x_1, x_2) \text{ και } g(x_1, x_2) = 0} \Rightarrow$$

$$\Rightarrow \begin{cases} 3x_1^2 = 2\lambda x_1 \\ 3x_2^2 = 2\lambda x_2 \\ x_1^2 + x_2^2 = 4 \end{cases} \Rightarrow \begin{cases} x_1(3x_1 - 2\lambda) = 0 \\ x_2(3x_2 - 2\lambda) = 0 \\ x_1^2 + x_2^2 = 4 \end{cases} \Rightarrow$$

$$\Rightarrow \begin{cases} \cdot\ x_1 = 0,\ x_2 = \frac{2\lambda}{3} \text{ και } \lambda = \pm 3 \\ \cdot\ x_2 = 0,\ x_1 = 2\lambda/3 \text{ και } \lambda = \pm 3 \end{cases}$$

$\cdot\ x_1 = 2\lambda/3,\ x_2 = \frac{2\lambda}{3}$ και $\lambda = \pm\sqrt{3}/2$

Άρα έχουμε τις λύσεις

$(x_1, x_2, \lambda) = (0, 2, 3)$

$(0, -2, -3)$

$(2, 0, 3)$

$(-2, 0, -3)$

$(\sqrt{2}, \sqrt{2}, 3/\sqrt{2})$

$(-\sqrt{2}, -\sqrt{2}, -3/\sqrt{2})$

Άρα

$\vec{x_1}^* = [0, 2]^T$

$\vec{x_2}^* = [0\ -2]^T$

$\vec{x_3}^* = [2\ 0]^T$

$\vec{x_4}^* = [-2\ 0]^T$

$\vec{x_5}^* = [\sqrt{2}\ \sqrt{2}]^T$

$\vec{x_6}^* [-\sqrt{2}\ -\sqrt{2}]^T.$

με τιμές $f(\vec{x}_1^*) = 8$ $\quad f(\vec{x}_5^*) = 4\sqrt{2}$

$f(\vec{x}_2^*) = -8$ $\quad f(\vec{x}_6^*) = -4\sqrt{2}$

$f(\vec{x}_3^*) = 8$

$f(\vec{x}_4^*) = -8$ με εφικτό σύνολο

$$\Omega = \{(x_1, x_2) \in \mathbb{R} : x_1^2 + x_2^2 = 4\}$$

Άρα σύμφωνα με το θεώρημα, η δεσμευμένα μέγιστη τιμή της f είναι 8, και η δεσμ. ελάχιστη είναι -8 με δεσμ. τοπ. ελαχιστοποιητές τους $\vec{x}_2^*$ και $\vec{x}_4^*$ και δεσμ. τοπ. μεγιστοποιητές τους $\vec{x}_1^*$ και $\vec{x}_3^*$

\# Ορισμός: Επιφάνεια του R^n

Έστω $g : X \subseteq R^n \to R$ το υποσύνολο του R^n

$$S = \{\vec{x} \in X : g(\vec{x}) = 0\}$$

ονομάζεται επιφάνεια του R^n με εξίσωση $g(\vec{x}) = 0$.

\# Ορισμός: Εφαπτόμενος κώνος - Εφαπτόμενος χώρος

Έστω $g : X \subseteq \mathbb{R}^n \to \mathbb{R}$ C^1 πραγμ. συνάρτηση

και $\vec{\alpha} \in S = \{\vec{x} \in X : g(\vec{x}) = 0\}$ με $\nabla g(\vec{\alpha}) \neq \vec{0}$ το σύνολο

$$C(\vec{\alpha}) = \{y \in \mathbb{R}^n : \nabla g(\vec{\alpha}) \cdot \vec{y} = 0\}$$

ονομάζεται εφαπτόμενος κώνος/χώρος της επιφάνειας S στο $\vec{\alpha}$.

10) Συνάρτηση Lagrange

Θεωρούμε δύο πραγματικές συναρτήσεις $f=f(\vec{x})$

$g=g(\vec{x}) \quad : X\subseteq \mathbb{R}^n \to \mathbb{R}$

Η πραγματική συνάρτηση:

$$L=L(\vec{x},\lambda) \quad : X\times \mathbb{R}\subseteq \mathbb{R}^{n+1}\to \mathbb{R}$$

η οποία ορίζεται από τον τύπο: $\boxed{L(\vec{x},\lambda)=f(\vec{x})-\lambda g(\vec{x})}$

ονομάζεται συνάρτηση Lagrange των συναρτήσεων f και g.

Αν f και g παραγωγίσιμες.

$$\boxed{\begin{cases} L_{x_i}(\vec{x},\lambda)=f_{x_i}(\vec{x})-\lambda g_{x_i}(\vec{x}) \\ L_{\lambda}(\vec{x},\lambda)=-g(\vec{x}) \end{cases}}$$

άρα $\nabla_{\vec{x}} L(\vec{x},\lambda)=\nabla f(\vec{x})-\lambda \nabla g(\vec{x})$.

Άρα, ένα σημείο $\vec{x}^*\in X$ είναι δεσμευμένο κρίσιμο σημείο της f υπό τη δέσμευση $g(\vec{x})=0$ τότε και μόνο τότε όταν υπάρχει $\lambda^*\in \mathbb{R}$ τ.ω.

$$\boxed{\nabla L(\vec{x}^*,\lambda^*)=\vec{0}}$$

(δηλαδή υπάρχει $\lambda^*\in\mathbb{R}$ για τον οποίο το $(\vec{x}^*,\lambda^*)$ είναι κρ. σημείο της Λαγκρανζιανής L.

για δεσμ. κρίσιμο σημείο..

Θεώρημα: Αναγκαίες Συνθήκες Πρώτης Τάξης

Έστω $f, g : X \subseteq \mathbb{R}^n \to \mathbb{R}$ (X ανοικτό) C^1 συναρτήσεις.

Υποθέτουμε ότι $\vec{x}^* \in X$ είναι δεσμευμένος τοπικός βελτιστοποιητής υπό τη δέσμευση $g(\vec{x}) = 0$, για τον οποίο ισχύει:

$$\nabla g(\vec{x}^*) \neq \vec{0}.$$

Τότε υπάρχει $\lambda^* \in \mathbb{R}$ έτσι ώστε να ισχύει

$$\nabla L(\vec{x}^*, \lambda^*) = \vec{0}.$$

δηλαδή ισοδύναμα να ισχύει $\nabla_{\vec{x}} L(\vec{x}^*, \lambda^*) = \vec{0}$ και
$g(\vec{x}^*) = 0$.

Θεώρημα: Αναγκαίες Συνθήκες Δεύτερης Τάξης

Έστω $f, g : X \subseteq \mathbb{R}^n \to \mathbb{R}$ (X ανοικτό) C^2 συναρτήσεις

Υποθέτουμε $\vec{x}^* \in X$ δεσμ. τοπικός ελαχιστοποιητής υπό $g(\vec{x}) = 0$ με $\nabla g(\vec{x}^*) \neq \vec{0}$

Τότε υπάρχει ένα $\lambda^* \in \mathbb{R}$ για το οποίο ισχύουν:

α) $\nabla_{\vec{x}} L(\vec{x}^*, \lambda^*) = \vec{0}$ και $g(\vec{x}^*) = 0$

β) ο πίνακας $\nabla^2_{\vec{x}\vec{x}} L(\vec{x}^*, \lambda^*)$ είναι θετικά ημιορισμένος στον εφαπτόμενο χώρο $C(\vec{x}^*)$

δηλαδή ισχύει:

$$\vec{u}^T \nabla^2_{\vec{x}\vec{x}} L(\vec{x}^*, \lambda^*) \vec{u} \geq 0, \text{ για κάθε } \vec{u} \in C(\vec{x}^*) = \{\vec{y} \in \mathbb{R}^n : \nabla g(\vec{x}^*) \cdot \vec{y} = 0\}$$

Θεώρημα Ικανής Συνθήκης Δεύτερης Τάξης

Έστω $f, g : X \subseteq \mathbb{R}^n \to \mathbb{R}$ (X ανοικτό) C^2 συναρτήσεις

και $x^* \in X$ με τις ιδιότητες:

1) $g(\vec{x}^*) = 0$ και $\nabla g(\vec{x}^*) \neq 0$

2) υπάρχει $\lambda^* \in \mathbb{R}$ τ.ω. $\nabla_{\vec{x}} L(\vec{x}^*, \lambda^*) = \vec{0}$

3) ο πίνακας $\nabla^2_{\vec{x}\vec{x}} L(\vec{x}^*, \lambda^*)$ είναι θετικά ορισμένος στον εφαπτόμενο χώρο $C(\vec{x}^*)$

δηλαδή $\vec{u} \cdot \nabla^2_{\vec{x}\vec{x}} L(\vec{x}^*, \lambda^*) > 0$, για κάθε $\vec{u} \in C(\vec{x}^*)$, $\vec{u} \neq 0$

Τότε το $\vec{x}^*$ είναι (γνήσιος) δεσμευμένος τοπικός ~~ελ~~ ελάχιστος της f υπό τον δεσμό $g(\vec{x}) = 0$.

Εσσιανός Πίνακας της Λαγκρανζιανής Συνάρτησης

$$\nabla^2 L(\vec{x}, \vec{\lambda}) = \begin{array}{c} \\ \left[\begin{array}{ccccc} L_{x_1x_1} & L_{x_1x_2} & \cdots & L_{x_1x_n} & L_{x_1\lambda} \\ L_{x_2x_1} & L_{x_2x_2} & \cdots & L_{x_2x_n} & L_{x_2\lambda} \\ \vdots & \vdots & \ddots & \vdots & \\ L_{x_nx_1} & L_{x_nx_2} & \cdots & L_{x_nx_n} & L_{x_n\lambda} \end{array}\right] \end{array}_{(\vec{x}, \lambda)}$$

με $L_{x_nx_n} = f_{x_nx_n}(\vec{x}) - \lambda g_{x_nx_n}(\vec{x})$

$\longrightarrow$

Θεώρημα: (Εναλλακτική) Ικανή Συνθήκη Δεύτερης Τάξης

Έστω $f = f(\vec{x})$ και $g = g(\vec{x}) : X \subseteq \mathbb{R}^n \to \mathbb{R}$ (C^2 συναρτήσεις)
(X ανοικτό)

- $\vec{X}^*$ είναι δεσμευμένο κρίσιμο σημείο της f υπό τη δέσμευση $g(\vec{x}) = 0$
- L η συνάρτηση Langrange των f και g και $\nabla^2 L(\vec{X}^*, \lambda^*)$ ο Εσσιανός της L στο κρίσιμο σημείο $(\vec{X}^*, \lambda^*)$

Υποθέτουμε ότι $|\nabla^2 L(\vec{X}^*, \lambda^*)| \neq 0$

Τότε ισχύουν οι ισχυρισμοί: **SOS**

1) Αν $\Delta_3 < 0$, $\Delta_4 < 0$, $\dots$, $\Delta_{n+1} < 0$ τότε η f έχει στο $\vec{X}^*$ δεσμευμένο τοπικό ελάχιστο υπό $g(\vec{x}) = 0$.

2) Αν $\Delta_3 < 0$, $\Delta_4 > 0$, $\dots$, $(-1)^{n+1}\Delta_{n+1} < 0$ (εναλλάξ με το πρώτο < 0.) τότε η f έχει στο $\vec{X}^*$ δεσμ. τοπικό μέγιστο υπό $g(\vec{x}) = 0$.

3) Αν δεν ισχύει καμία από τις περιπτώσεις 1) και 2) τότε το $\vec{X}^*$ δεν είναι δεσμ. τοπικός βελτιστοποιητής της f.

Παράδειγμα sos

Βρείτε τα σημεία δεσμευμένου τοπικού ακροτάτου της

$f(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2$ υπό τη δέσμευση $g(x_1,x_2,x_3) =$

$= x_1 + x_2 + x_3 - 1 = 0.$

Λύση

f, g C^2 συναρτήσεις στο $\mathbb{R}^3$

$f_{x_1} = 2x_1$, $f_{x_2} = 2x_2$, $f_{x_3} = 2x_3$

$f_{x_1x_1} = 2$, $f_{x_2x_2} = 2$, $f_{x_3x_3} = 2$

$f_{x_1x_2} = f_{x_2x_1} = f_{x_1x_3} = f_{x_2x_3} = f_{x_3x_1} = f_{x_3x_2} = 0.$

$g_{x_1} = g_{x_2} = g_{x_3} = 1.$

Τα δεσμευμένα σημεία προκύπτουν ως λύσεις του συστήματος.

$\nabla f(\vec{x}) - \lambda \nabla g(x) = 0$ και $g(x) = 0. \Rightarrow$

$$\Rightarrow \begin{cases} (2x_1 - \lambda, 2x_2 - \lambda, 2x_3 - \lambda) = 0 \\ x_1 + x_2 + x_3 - 1 = 0 \end{cases} \Rightarrow \begin{matrix} x_1 = 1/3 \\ x_2 = 1/3 \\ x_3 = 1/3 \end{matrix} \quad \lambda = 2/3.$$

άρα $(\vec{x}^*, \lambda^*) = \frac{1}{3}(1, 1, 1, 2).$

$$\nabla^2 L(\vec{x}^*, \lambda^*) = \begin{bmatrix} 2 & 0 & 0 & -1 \\ 0 & 2 & 0 & -1 \\ 0 & 0 & 2 & -1 \\ -1 & -1 & -1 & 0 \end{bmatrix}$$

$\Delta_3 = 8 > 0$, $\Delta_4 = -12 < 0$

Άρα το $\vec{x}^*$ είναι δεσμ. τοπ. μέγιστο της f.

Παράδειγμα. 5ος

Βρείτε τους δεσμευμένους τοπ. βελτιστοποιητές της συνάρτησης $f(x_1, x_2, x_3) = x_1 x_2 x_3$ υπό τη δέσμευση $g(x_1, x_2, x_3) = x_1 x_2 + x_1 x_3 + x_2 x_3 - 6 = 0$.

Λύση

f, g C^2 συναρτήσεις στο $\mathbb{R}^3$ με:

$f_{x_1} = x_2 x_3$, $f_{x_2} = x_1 x_3$, $f_{x_3} = x_1 x_2$

$f_{x_1 x_1} = 0$, $f_{x_2 x_2} = 0$, $f_{x_3 x_3} = 0$

$f_{x_1 x_2} = x_3$, $f_{x_2 x_1} = x_3$, $f_{x_3 x_1} = x_2$

$f_{x_1 x_3} = x_2$, $f_{x_2 x_3} = x_1$, $f_{x_3 x_2} = x_1$

$g_{x_1} = x_2 + x_3$, $g_{x_2} = x_1 + x_3$, $g_{x_3} = x_1 + x_2$

$g_{x_1 x_1} = g_{x_2 x_2} = g_{x_3 x_3} = 0$.

$g_{x_1x_2} = g_{x_2x_1} = 1$, $g_{x_3x_1} = g_{x_1x_3} = 1$

$g_{x_2x_3} = g_{x_3x_2} = 1$

Τα δεσμευμένα κρίσιμα σημεία βρίσκονται ως λύσεις του συστήματος:

$$\begin{cases} \nabla f(x_1,x_2,x_3) - \lambda \nabla g(x_1,x_2,x_3) = 0 \\ g(x_1,x_2,x_3) = 0 \end{cases} \Rightarrow$$

$$\Rightarrow \begin{cases} \dots \end{cases} \Rightarrow \begin{aligned} (\vec{x}_a^*, \lambda^*) &= (\sqrt{2}, \sqrt{2}, \sqrt{2}, \sqrt{2}/2) \\ (\vec{x}_b^*, \lambda^*) &= (-\sqrt{2}, -\sqrt{2}, -\sqrt{2}, -\sqrt{2}/2) \end{aligned}$$

Υπολογίζω Hessian Matrix.

$$\nabla^2 L(\vec{x}, \lambda) = \begin{bmatrix} 0 & x_3 - \lambda & x_2 - \lambda & -(x_2+x_3) \\ x_3 - \lambda & 0 & x_1 - \lambda & -(x_1+x_3) \\ x_2 - \lambda & x_1 - \lambda & 0 & -(x_1+x_2) \\ -(x_2+x_3) & -(x_1+x_3) & -(x_1+x_2) & 0 \end{bmatrix}$$

Για το (x_a^*, λ^*): $\begin{bmatrix} \dots \end{bmatrix}$

$\Delta_3 = \sqrt{2}/2 > 0$, $\Delta_4 = -12 < 0$

→ δεσμ. τοπ. μεγιστοποιητής

Για το $(\vec{x}_b^*, \lambda^*)$: $\begin{bmatrix} \dots \end{bmatrix}$

$\Delta_3 = -\sqrt{2}/2 < 0$, $\Delta_4 = -12 < 0$

→ δεσμ. τοπ. ελαχιστοποιητής.

Βελτιστοποίηση α.

Newton for Nonlinear Equations in $\mathbb{R}$

ΒΑΣΙΚΑ ΣΤΟΙΧΕΙΑ ΤΗΣ ΜΕΘΟΔΟΥ NEWTON ΣΤΟ $\mathbb{R}$.

Μέθοδο αναζήτησης ευθείας που βασίζεται στη μέθοδο Newton για προσδιορισμό κατεύθυνσης αναζήτησης $\vec{p_k}$.

Βασικά στοιχεία μεθόδου Newton στο $\mathbb{R}$.

Αναζητάμε προσέγγιση λύσης

$$f(x) = 0.$$

→ Η μέθοδος Newton για την προσέγγιση της x^* της παραγωγίσιμης συνάρτησης f βασίζεται στην ιδέα ότι η εφαπτομένη του γραφήματος της f βρίσκεται κοντά στο γράφημα της f στην περιοχή του σημείου επαφής.

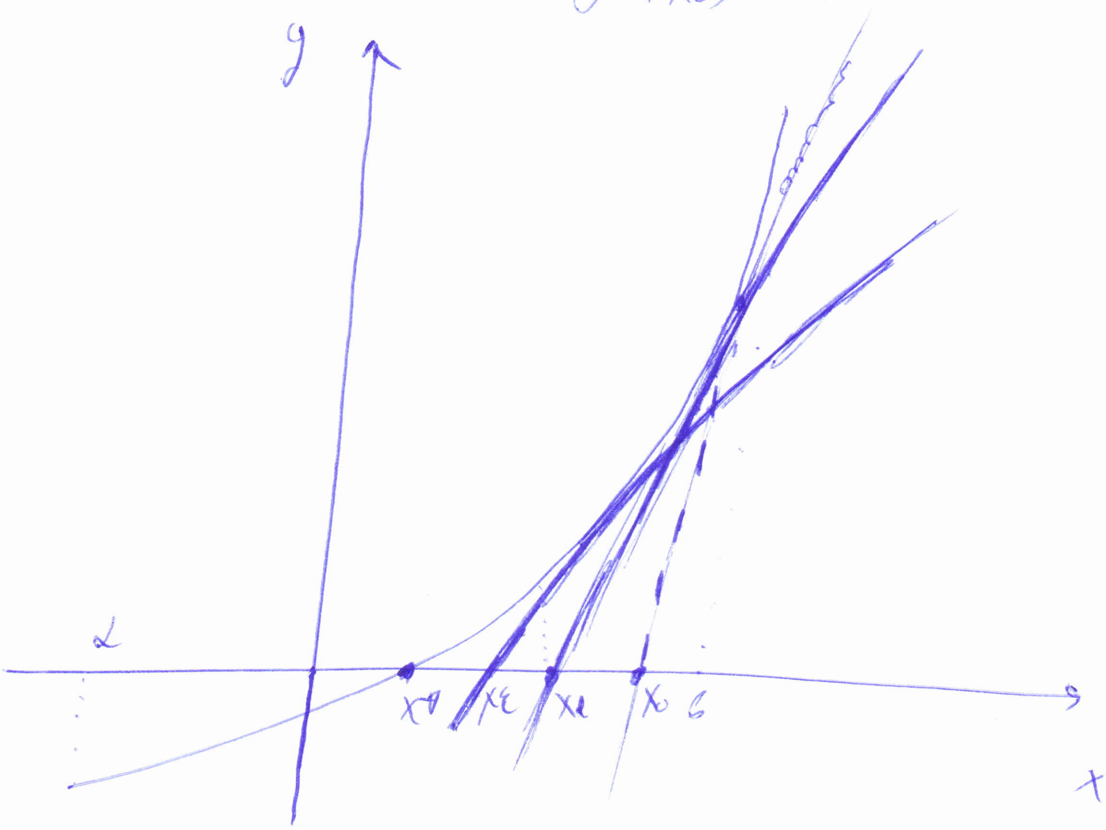
→ Αρχική προσέγγιση x_0 της x^* και θεωρούμε την εφαπτομένη (T) του γραφήματος της $y = f(x)$ στο $(x_0, f(x_0))$, η T και το γράφημα της f τέμνουν του x-άξονα σε κοντινό σημείο διότι η T βρίσκεται κοντά στο γράφημα της f.

→ Αν x_0 κατάλληλα κοντά στο x^*, η x_1 τετμημένη του σημείου τομής του x-άξονα θα μπορούσε να είναι μια καλύτερη προσέγγιση της x^*.

$$(T): \; y = f(x_0) + f'(x_0)(x - x_0)$$

για $y = 0$, $x = x_1$ αν $f'(x_0) \neq 0$.

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}$$



λαμβάνοντας την x_1 ως δεύτερη προσέγγιση της x^* επαναλαμβάνουμε από την προσέγγιση αν $f'(x_1) \neq 0$

έτσι $$x_2 = x_1 - \frac{f(x_1)}{f'(x_1)}$$

συνεχίζω μέχρι να προσεγγίσω όσο επιθυμώ.

Βελτιστοποίηση 6.

Μέθοδος Newton

$f:[\alpha,b] \to \mathbb{R}$ $(\alpha < b)$ συνεχής με $f(\alpha)\cdot f(b) < 0$, παραγωγίσιμη στο (α,b) με $f'(x) \neq 0$ για κάθε $x \in (\alpha,b)$.

Τότε από το θεώρημα της ενδιάμεσης τιμής υπάρχει μια τουλάχιστον ρίζα x^* της f στο (α,b).

Επαναληπτική διαδικασία για προσέγγιση της x^* της f στο (α,b)

1. Θεωρούμε αρχ. προσέγγιση x_0 της ρίζας x^*

2. Οι επόμενες προσεγγίσεις λαμβάνονται από

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}, \quad n=0,1,2,\dots$$

3. Ως τελική προσέγγιση x_{n+1} λαμβάνουμε εκείνη για την οποία $|x_{n+1} - x_n| \leq \epsilon$ (ανάλογα με την ακρίβεια που θέλουμε) ϵ κατάλληλο επιθυμητό σφάλμα.

Παράδειγμα 1

Υπολογίστε μια προσεγγιστική τιμή της $\sqrt{2}$ χρησιμοποιώντας τις προσεγγίσεις $x_0 = 1$, x_1, x_2 της Μεθόδου Newton.

Λύση

Θεωρούμε τη συνάρτηση $f(x) = x^2 - 2$

$(x^2 - 2 = 0 \Rightarrow x = \sqrt{2})$

$$f'(x) = 2x \quad \text{άρα} \quad x_{n+1} = x_n - \frac{x_n^2 - 2}{2x_n}, \quad n = 0, 1, 2, \ldots$$

$x_0 = 1$, $x_1 = 1.416667$, $x_2 = 1.414216$

άρα $\sqrt{2} \approx 1.414216$.

Παράδειγμα 2 ίδιο

Χρησιμοποιείστε τη μέθοδο Newton για την προσέγγιση της πραγματικής ρίζας της συνάρτησης:

$$f(x) = 2x^3 + x^2 - x + 1$$

συνεχίζοντας μέχρι δύο διαδοχικές προσεγγίσεις να απέχουν λιγότερο από $\epsilon = 10^{-4}$. Επιλέξτε $x_0 = -1.2$

$$f'(x) = 6x^2 + 2x - 1 \quad \text{άρα} \quad x_{n+1} = x_n - \frac{2x_n^3 + x_n^2 - x_n + 1}{6x_n^2 + 2x_n - 1}$$

1

Μέθοδος Newton

1 → x_0

2 → $$\boxed{x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}}$$

3 → x_{n+1} αποδεκτό αν $|x_{n+1} - x_n| \leq \epsilon$

Απόδειξη #1 (τύπος μεθόδου Newton από Taylor)

$$f(x_n + p) \simeq f(x_n) + p f'(x_n) \quad \text{(Taylor)}$$

αν $f'(x_n) \neq 0$

$$f(x^*) \simeq f(x_n) + p f'(x_n) = 0$$

άρα $p = -\dfrac{f(x_n)}{f'(x_n)} \Rightarrow$ $(x_{n+1} = x_n + p)$ $$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}.$$

Ταχύτητα/Ρυθμός Σύγκλισης

$$\boxed{\lim_{n \to +\infty} \frac{\|\vec{e}_{n+1}\|}{\|\vec{e}_n\|^r} = c}.$$

- $r = 1$: γραμμική σύγκλιση.
  - $0 < c < 1$: μειώνεται με σταθερό ρυθμό
  - $c > 1$: $\vec{e}_n$: αυξάνεται

~~[illegible]~~

- $r > 1$: superlinear
- $r = 2$: τετραγωνική

* $e_n = x_n - x^*$

Απόδειξη #2 (Newton έχει τετραγωνική σύγκλιση).

$$e_n = x_n - x^*$$

$f(x^*) = 0$ και $f'(x_n) \neq 0$

$\lim_{n\to\infty} x_n = x^*$

Από Taylor: $f(x^*) = f(x_n - e_n) =$

$$= f(x_n) - e_n f'(x_n) + \frac{1}{2} e_n^2 f''(\xi) \Rightarrow$$

$$\Rightarrow e_n - \frac{f(x_n)}{f'(x_n)} = \frac{1}{2} e_n^2 \frac{f''(\xi)}{f'(x_n)} \Rightarrow$$

$$\Rightarrow x_n - x^* - \frac{f(x_n)}{f'(x_n)} = \frac{1}{2} (x_n - x^*)^2 \frac{f''(\xi)}{f'(x_n)} \Rightarrow$$

$$\Rightarrow x_{n+1} - x^* \simeq \left(\frac{1}{2} \frac{f''(x^*)}{f'(x^*)}\right) (x_n - x^*)^2$$

άρα $r = 2$

$$c = \frac{1}{2} \frac{f''(x^*)}{f'(x^*)}$$

→ Απόδειξη ότι αποκλίνει αν $\left|\frac{x_{n+1}}{x_n}\right| > 1$ → αποκλίνει...

Σ Η διαφορά απ' τον κλασσικό τύπο είναι: για το πρόβλημα ελαχιστοποίησης συναρτήσεως n μεταβλητών, εφαρμόζουμε μια επέκταση του κλασ. τύπου που αφορά την επίλυση συνεχ. εξίσωσης $\nabla f(x) = 0$.

Μέθοδος Newton. (Newton's Method for Optimization)

- $x_{k+1} = x_k - \alpha_k [\nabla^2 f(x_k)]^{-1} \nabla f(x_k)$

Αλγόριθμος Newton

προσδιορίζω x_0, ε, $k = 0$

έχω κατεύθυνση καθόδου όταν $\nabla f(x_k)^T p_k < 0$.

(α) while (NOT $\|\nabla f(x_k)\| < \varepsilon$) {

υπολογίζω

(β) • Hessian Matrix $\nabla^2 f(x_k)$

• $p_k = -[\nabla^2 f(x_k)]^{-1} \nabla f(x_k)$

Υπολογίζω Hessian Matrix και η p_k από (σύστημα) $[\nabla^2 f(x_k)] p_k = -\nabla f(x_k)$

(γ) • προσδιορίζω το α_k τ.ω. $\varphi(\alpha_k) = f(x_k + \alpha_k p_k)$ minimize ως προς $\alpha_k > 0$.

(δ) • προσδιορίζω $x_{k+1} = x_k + \alpha_k p_k$

συνήθως θα έχουμε βήμα $\alpha_k = 1$ (αν προσφέρει ικανοποιητική μείωση)

(ε) • θέτω $k = k+1$ }

return x_k ;

Η τετρ. συνάρτηση $f(x) = \frac{1}{2} x^T Q x - b^T x + c$

συγκλίνει σε ένα μόνο βήμα για οποιοδήποτε x_0. ($x_1 = x^*$)

Ισχύουν $x^* = Q^{-1} b$. (παίρνω $\alpha_0 = 1$)

Η.Μ 2x2

$$H = \begin{bmatrix} xx & xy \\ yx & yy \end{bmatrix}$$

Η.Μ 3x3

$$H = \begin{bmatrix} xx & xy & xz \\ yx & yy & yz \\ zx & zy & zz \end{bmatrix}$$

Από Η.Μ του X^* έχουμε:

$\Delta_1, \Delta_2 < 0$ → μη ελαχ.

$\Delta_1 < 0, \Delta_2 > 0$ → το μέγιστο ?

Υπολογισμός κατεύθυνσης p με Cholesky..

$$LL^T p = -\nabla f(x_0) \quad \text{όπου} \quad LL^T = \nabla^2 f(x_0)$$

άνω τριγ.

κάτω τριγ.

και βρίσκω το p.

Η St. Descent στο τετρ. συνάρτηση συγκλίνει γραμμικά.

Η Newton (κλασική) συγκλίνει τετραγωνικά.

Η Newton στο τετρ. συνάρτηση συγκλίνει σε ένα βήμα.

3

Για δεσμευμένα ακρότατα ισχύει: (δεσμ. μέγιστα & ελάχιστα τ.ρ.τ.)

$$\boxed{\nabla f(\vec{x}) = \lambda \nabla g(\vec{x}) \quad \text{και} \quad g(\vec{x}) = 0.}$$

Λύνω το σύστημα τα ~~μικρότερα ελαχιστοποιητές~~

τα ~~μεγαλύτερα μεγιστοποιητές~~

ελάχιστη τιμή = δεσμ. ελάχιστη τιμή

μέγιστη τιμή = δεσμ. μέγιστη τιμή

Lagrange

$$L(\vec{x},\lambda) = f(\vec{x}) - \lambda g(\vec{x})$$

$$\begin{cases} L_{x_i}(\vec{x},\lambda) = f_{x_i}(\vec{x}) - \lambda g_{x_i}(\vec{x}) \\ L_{\lambda}(\vec{x},\lambda) = -g(\vec{x}) \end{cases}$$

Στον Hessian Matrix της συνάρτησης Lagr.

(A) η τελευταία στήλη είναι το λ.

(B) (A)

i) $\left.\begin{array}{l} \Delta_3 < 0 \\ \Delta_4 < 0 \\ \vdots \\ \Delta_{n+1} < 0 \end{array}\right\}$ τοπ. ελάχιστο

ii) $\left.\begin{array}{l} \Delta_3 < 0 \\ \Delta_4 > 0 \\ \text{εναλλάξ} \end{array}\right\}$ τοπ. μέγιστο

iii) όχι τοπ. βελτιστοποιητές.

$\boxed{\text{SOS}}$

Ασκήσεις όπου βρείτε τα σημεία δεσμ. τοπικών ακροτάτων, δεσμ. τοπ. βελτιστοποιητές

1) f, g C^2 συναρτήσεις στο R^n

2) Υπολογίζω μερ. παραγώγους (f έως 2) (g έως 1)

3) $\nabla f(\vec{x}) - \lambda \nabla g(x) = 0$ και $g(x) = 0$ λύνω το σύστημα για να βρω τα σημεία

4) Γράφω τον Hessian βάσει των σημείων.

5) Βρίσκω με Δ_3, Δ_4 (A)

4/ Αν μπορώ να λύσω ~~για~~ τη $g(x)=0$ ως προς κάποια μεταβλητή και τη βάλω στην συνάρτηση,

δουλεύω με 1η παράγωγο και αδέσμευτη βελτιστοποίηση.

5

Συνθήκες 1ης τάξης

$$\begin{cases} \nabla_{\vec{x}} L(\vec{x}^*, \lambda_1^*) = \vec{0} \\ \lambda_1^* g_1(\vec{x}) = 0 \end{cases}$$

$\lambda = 0$ σημαίνει ανενεργός περιορισμός.

(όπου $L(\vec{x}, \lambda) = f(\vec{x}) - \lambda g_1(\vec{x})$.)

≠ ~~[illegible]~~ KKT conditions

$$\nabla_{\vec{x}} L(\vec{x}^*, \vec{\lambda}^*) = 0$$

$$g_i(\vec{x}) = 0 \quad i \in \mathcal{E}$$

$$g_i(\vec{x}) \geq 0 \quad i \in I$$

$$\lambda_i \geq 0 \quad i \in I$$

$$\lambda_i g_i(\vec{x}^*) \geq 0, \; \forall i \in \mathcal{E} \cup I$$

Θέμα 1 - Θέμα 2.

- Μέθοδο Τέλειας Επαγωγής.
- Γεωμετρική ακολουθία
- Κριτήριο $\left|\frac{a_{n+1}}{a_n}\right| \to a$

$0 < a < 1 \to 0$

$a > 1 \to +\infty$

$a = 1$ ίδιο

- Αναδρομική ακολουθία.
- Θεώρημα κριτήριο λόγου $q = \frac{|a_{n+1}|}{|a_n|}$ (ίδιο με κριτήριο ρίζας $\sqrt[n]{|a_n|}$)

$q > 1$ αποκλίνει

$q < 1$ απόλυτα συγκλ.

$q =$ ίδιο.

- Σύνολο σύγκλισης δυναμοσειράς

δουλεύω με $R = \frac{1}{\lim \sqrt[n]{|a_n|}}$ και $\begin{cases} R = 0 & \text{συγκλίνει για } x = 0 \\ R = +\infty & \text{απόλυτα συγκλίνει} \\ R \in (0, +\infty) & |x| < R \text{ απολ. συγκλ.} \\ & |x| > R \text{ αποκλίνει} \\ & |x| = R \text{ ίδιο.} \end{cases}$

ίδιο

- Όρια συναρτήσεων sin, cos.

Sandwich criteria ή δύο υπακολουθίες

$2n\pi \to +\infty$

$2n\pi + \pi/2 \to +\infty$ ... $\left.\begin{matrix} f(x_n) \to b_1 \\ f(y_n) \to b_2 \end{matrix}\right\} b_1 \neq b_2.$

- Leibniz φθίνουσα και μηδενική συγκλίνει.

$$\left|\sum_{k=n+1}^{\infty} (-1)^{k+1} a_k\right| \leq a_{n+1}$$

- Φραγμένη επί μηδενική = μηδενική
- $\frac{a_{n+1}}{a_n} \to b \Rightarrow \sqrt[n]{a_n} \to b.$
- Μονότονη + φραγμένη = συγκλίνει

-//- + μη φραγμ. = συγκλίνει $+\infty$ ή $-\infty$.

- Εναλλάσσουσες σειρές

Leibnitz → φθίνουσα και μηδενική ακολουθία.

$$\sum_{k=n+1}^{\infty}(-1)^{k+1}a_k \,\Big| \le a_{n+1}.$$

	ημ	συν
0°	0	1
90°	1	0

Θέμα 3. - Θέμα 4

- Θεώρημα Rolle + Απόδειξη
- Διατύπωση Taylor
- Σειρές Mclaurin για sinx, cosx.

$\Rrightarrow$ Να βρω $\cos(2x) = \sum_{n=0}^{+\infty} \frac{(-1)^n 2^{2n} x^{2n}}{(2n)!}$

- Αόριστα ολοκληρώματα.
- Θεώρημα Μέσης Τιμής + Απόδειξη
- Από τον τύπο $e^x = \sum_{n=0}^{+\infty} \frac{x^n}{n!}$
- για το $\cos(x) = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n}}{(2n)!}$

παίρνω την Taylor σειρά για το cos x
Mclaurin κλπ κλπ.

Κριτήριο Δεύτερης Παραγώγου για ακρότατα.

Αν $f'(x_0) = 0$ και $f''(x_0) > 0$ τότε η f έχει τοπικό ελάχιστο στο x_0.

Απόδειξη

$$f''(x_0) = \lim_{h \to 0} \frac{f'(x_0+h) - f'(x_0)}{h} = \lim_{h \to 0} \frac{f'(x_0+h)}{h} > 0.$$

Επομένως για μικρό h και θετικό έχω $f'(x_0+h) > 0$ ενώ για μικρό h και αρνητικό, έχω $f'(x_0+h) < 0$. Ισοδύναμα μπορώ να γράψω

- για μικρό h και θετικό έχω $f'(x_0+h) > 0$
- για μικρό h και θετικό έχω $f'(x_0-h) < 0$.

Μπορώ να βρω ένα $\varepsilon > 0$ με $\varepsilon < h$ τέτοιο ώστε

- η f να είναι γν. φθίνουσα στο $(x_0 - \varepsilon, x_0)$ και
- η f να είναι γν. αύξουσα στο $(x_0, x_0 + \varepsilon)$

Άρα το x_0 είναι τοπικό ελάχιστο.

Θεώρημα Μέσης Τιμής.

Η f συνεχής στο $[\alpha, b]$ και παραγωγίσιμη στο (α, b).
Τότε υπάρχει $\xi \in (\alpha, b)$ τέτοιο ώστε $f'(\xi) = \dfrac{f(b) - f(\alpha)}{b - \alpha}$

Απόδειξη

Εφαρμόζω Rolle στη συνάρτηση

$$g(x) = f(x) - \frac{f(b) - f(\alpha)}{b - \alpha}(x - \alpha)$$

Ισχύει $g(\alpha) = f(\alpha)$

$g(b) = f(b) - (f(b) - f(\alpha)) = f(\alpha)$

άρα $g(\alpha) = g(b)$. Οπότε $\exists\ \xi \in (\alpha, b)$ τέτοιο ώστε

$$g'(\xi) = 0 \Rightarrow f'(\xi) = \frac{f(b) - f(\alpha)}{b - \alpha}.$$

Πόρισμα 1 του Θ.Μ.Τ.

Αν $f'(x) = 0$ για κάθε $x \in (\alpha, \beta)$ τότε $f = constant$.

Απόδειξη

Έστω $x_0 \in (\alpha, \beta)$. Με ΘΜΤ στο (x_0, x) (ή στο (x, x_0)) υπάρχει $\xi \in (x_0, x)$ με $f(x) - f(x_0) = f'(\xi)(x - x_0)$

Άρα $f(x) = f(x_0)$ ισχύει για κάθε x άρα η f έχει σταθερή τιμή $f(x_0)$.

Rolle

Η f συνεχής στο $[\alpha, b]$ και παραγωγίσιμη στο (α, b) με $f(\alpha) = f(b)$ τότε υπάρχει $\xi \in (\alpha, b)$ τέτοιο ώστε $f'(x) = 0$

Απόδειξη

Εφόσον η f είναι συνεχής σε κλειστό διάστημα έχει ακρότατα. Έστω $\xi \in (\alpha, b)$ το σημείο μεγίστου. Τότε $f'(\xi) = 0$. Αν $\xi \in (\alpha, b)$ σημείο ελαχίστου τότε επίσης από το θεώρημα του Fermat $f'(\xi) = 0$

Αν υποθέσουμε τέλος ότι η μέγιστη και η ελάχιστη τιμή λαμβάνονται στα άκρα, τότε $f(\alpha) = f(b)$ η μέγιστη και η ελάχιστη τιμή θα συμπίπτουν, άρα η f είναι σταθερή, επομένως οποιοδήποτε $\xi \in (\alpha, b)$ θα ικανοποιεί την $f'(\xi) = 0$

Fermat.

Η f ορισμένη στο $[\alpha, b]$ και παραγωγίσιμη στο $x \in (\alpha, b)$. Αν x σημείο τοπικού ακροτάτου, τότε $f'(x) = 0$.

Απόδειξη

Έστω x σημείο τοπικού μεγίστου.

$\exists h \in \mathbb{R}$ μικρό τέτοιο ώστε $f(x) \geq f(x+h)$. Επομένως για

· $h > 0$ και μικρό, προκύπτει $f(x+h) - f(x) \leq 0$

δηλαδή $$\lim_{h \to 0} \frac{f(x+h) - f(x)}{h} \leq 0 \quad (1)$$

· παρόμοια για $h < 0$ προκύπτει $\dfrac{f(x+h) - f(x)}{h} \geq 0$

δηλαδή $$\lim_{h \to 0} \frac{f(x+h) - f(x)}{h} \geq 0 \quad (2)$$

(1), (2) συνεπάγονται $f'(x) = 0$.

Κανόνας της αλυσίδας.

$$[f(g(x))]' = f'(g(x)) \cdot g'(x).$$

για το θεώρημα σελ 8.

Θ Θεώρημα Μέσης Τιμής.

f συνεχής $[\alpha, b]$

f παραγωγίσιμη στο (α, b)

$$\exists x_0 \in (\alpha, b) : f'(x_0) = \frac{f(b) - f(\alpha)}{b - \alpha}.$$

Θεώρημα.

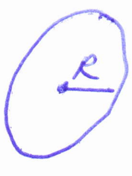
f συνεχής στο $[\alpha, b]$

f παραγωγίσιμη στο (α, b)

$f(\alpha) = f(b)$

$\exists x_0 \in (\alpha, b).\ f'(x_0) = 0$

Θ Γραμμικοποίηση (γραμμική προσέγγιση μιας συνάρτησης)



$R = 3cm$ προσέγγιση: $E = 9\pi$

$E = \pi R^2$ πραγματικό $R_{\pi\rho} = R + \Delta R \rightarrow$ σφάλμα.

$E_{\pi\rho} = \pi(R + \Delta R)^2$

Παράδειγμα 1

Βρείτε διαφορικό ΔV (V όγκος σφαίρας) με ακτίνα $r = 6cm$

$\Delta r = 0.003$ cm $\qquad \Delta V = ;$

$V = \frac{4}{3}\pi R^3$, $V' = 4\pi R^2$

$$\Delta V = V'(6)\cdot\Delta r = 36\cdot 4\cdot\pi\cdot 0.0003 \simeq \ldots$$

Παράδειγμα 2

$\sqrt{100,05}$ $\qquad x_0 = 100 \quad \Delta x = 0.05$

$f(x) = \sqrt{x}$ $\qquad f'(x) = \frac{1}{2\sqrt{x}}$

$$\Delta f(x) = f'(x_0)\,\Delta x = \frac{1}{2\sqrt{100}}\cdot 0.05 = \frac{1}{20}\cdot 0.05$$

$$\Delta f(x) = f(x_0 + \Delta x) - f(x_0) = \frac{0.05}{20}$$

Θ Γραμμικοποίηση Ορισμός: f παραγωγίσιμη στο α, ονομάζω γραμμικοποίηση της f στο α

$$L_\alpha(x) = f(\alpha) + f'(\alpha)(x-\alpha)$$

$$\Delta f \simeq f'(\alpha)(x-\alpha)$$

Παράδειγμα.

Βρείτε τη γραμμικοποίηση της $f = \sqrt{1+x}$, $\alpha = 0$

$f'(x) = \frac{1}{2\sqrt{1+x}}$ $\qquad L_0(x) = 1 + \frac{1}{2}x.$

Θ Θεώρημα Bolzano

αν f συνεχής στο $[\alpha, b]$ και $f(\alpha)\cdot f(b) < 0$
τότε $\exists\, x_0 \in (\alpha, b)$ τέτοιο ώστε $f(x_0) = 0$.

Θ Σύμβολα του Landau

f, g δύο συναρτήσεις ορισμένες στο $J \subseteq \mathbb{R}$.

$f(x) = O(g(x))$ (ή $f(x) \in O g(x)$)

$\Longleftrightarrow$ $\exists\, M > 0$ και $x_0 \in \mathbb{R}$ τέτοιο ώστε $|f(x)| \leq M(g(x))$
$\forall\, x \geq x_0$.

$O(n) \rightarrow$ εκθετικός χρόνος συνάρτησης.

Παράδειγμα. (Σύμβολα Landau)

$f(x) = 6x^4 - 2x^3 + 5x + 2$ να δείξετε ότι $O(f(x)) = x^4$
$\Longleftrightarrow |f(x)| \leq M x^4$
$M \in \mathbb{R},\ x \geq x_0$.

$(M, x_0) = (15, 2)$

$|6x^4 - 2x^3 + 5x + 2| \leq |6x^4| + |-2x^3| + |5x| + 2$

για $x \geq 2 \Longrightarrow \leq 6x^4 + 2x^4 + 5x^4 + 2x^4 = 15x^4$

Παράδειγμα (Θεώρημα Bolzano)

$f: [0,1] \rightarrow [0,1]$ νδο $\exists\, x_0 \in [0,1] : f(x_0) = x_0$

Λύση

Υπόθεση $f(0) \neq 0$ και $f(1) \neq 1$.

$f(0) > 0$ και $f(1) < 1$.

Θέτω $g(x) = f(x) - x$.
$[0, 1]$

$$\left.\begin{aligned} g(0) &= f(0) > 0 \\ g(1) &= f(1) - 1 < 0 \end{aligned}\right\} \overset{\text{Θ. Bolzano}}{\Rightarrow} \exists x_0 \in (0,1) \text{ τέτοιο ώστε } g(x_0) = 0$$

ή $f(x_0) - x_0 = 0 \Rightarrow f(x_0) = x_0$.

Θ. Παράγωγος

Ορισμός: Αν υπάρχει το $\lim\limits_{h \to 0} \dfrac{f(x+h) - f(x)}{h}$ λέμε ότι η f είναι παραγωγίσιμη στο σημείο $x \in D_f$.

Κριτήριο: f παραγωγίσιμη στο $x_0 \in D_f \Leftrightarrow$

$$f'_\delta(x_0) = \lim_{x \to x_0^+} \frac{f(x) - f(x_0)}{x - x_0}$$

$$f'_\alpha(x_0) = \lim_{x \to x_0^-} \frac{f(x) - f(x_0)}{x - x_0}$$

$$f'_\alpha(x_0) = f'_\delta(x_0)$$

Ιδιότητες

i) $(f+g)'(x) = f'(x) + g'(x)$

ii) $(f \cdot g)'(x) = f'(x) \cdot g(x) + f(x) \cdot g'(x)$

iii) $\left(\dfrac{f}{g}\right)'(x) = \dfrac{f'(x)g(x) - f(x)g'(x)}{(g(x))^2}$

Αποδείξεις από τον ορισμό.

iv) f παραγωγίσιμη στο $x_0 \Rightarrow f$ συνεχής στο x_0.

Θεώρημα Taylor.

$$f(x) = f(\alpha) + f'(\alpha)(x-\alpha) + \ldots + \frac{f^{k}(\alpha)}{k!}(x-\alpha)^{k} + \text{υπόλοιπο}$$

ή ως σειρά

$$\sum_{n=0}^{\infty} \frac{f^{n}(\alpha)}{n!}(x-\alpha)^{n}.$$ με κέντρο το α

για $\alpha = 0$ λέγεται σειρά McLaurin.

$$e^{x} = \sum_{n=0}^{\infty} \frac{x^{n}}{n!}$$

$e^{0} = 1$ $\quad \frac{de^{x}}{dx} = e^{x}$

McLaurin για $e^{x} = \sum_{n=0}^{\infty} \frac{f^{n}(0)}{n!}(x-0)^{n} = \sum_{n=0}^{\infty} \frac{1 \cdot x^{n}}{n!}.$

1) Αποδείξτε τον τύπο της μεθόδου Newton με Taylor $f(x_k+p)$

2) Αποδείξτε τον τετρ. ρυθμό σύγκλισης της μεθόδου Newton. $e_k = x_k - x^*$ βάσει Taylor

3) Γράψτε τον αλγόριθμο της μεθόδου Newton..

4) Βρείτε τα r και C για τις

1) $x_k = 1+10^{-k}$ 2) $x_{k+1} = \frac{x_k}{2} + \frac{2}{x_k}$

$e_k = x_k - x^*$ (η ακολουθία το e_{k+1} βάσει του e_k). $x_0 = 4$.

5) Δείξτε ότι η ακολουθία Newton στη $f(x) = x^{1/3}$ αποκλίνει για $x_0 \neq 0$ και $a_k = 1$? δουλεύω με $\frac{|x_{k+1}|}{|x_k|}$

1) Γράψτε τον αλγόριθμο Newton για optimization.

2) Επιβεβαιώστε ότι για την τετρ. συνάρτηση, η Newton συγκλίνει σε ένα βήμα $(x_1 = x^*)$

3) Για ποιο a_k θα συγκλίνει με Newton η $f = x^{4/3}$ $\frac{|x_{k+1}|}{|x_k|} < 1$

4) +, -- μεθόδου Newton.

5) $f(x) = x^4 - 1$, $x_0 = 4$, $a_k = 1$

i) 3 πρώτες επαναλήψεις

ii) x^*

iii) ρυθμό σύγκλισης.

$\frac{x_{k+1}}{x_k} = \frac{2}{3}$

άρα $x_k = \left(\frac{2}{3}\right)^k$

6) Εφαρμογή Newton για $a_k = 1$ βελτιστοποίησης $\nabla f(x^*) = 0$.

$$x_0 = [1 \quad 1]^T$$

όταν $f(x) = 5x_1^4 + 6x_2^4 - 6x_1^2 + 2x_1x_2 + 5x_2^2 + 15x_1 - 7x_2 + 13$

7) Βρείτε την κατεύθυνση p, της Newton, $x_0 [0 \quad 1]^T$ με Cholesky.

$$f(x_1, x_2) = x_1^4 + 2x_1^3 + 2x_1^2 + x_2^2 - 2x_1x_2.$$

$\to \nabla^2 f(x_0)$

$$LL^T p = -\nabla f(x_0)$$

Προβλήματα Βελτιστοποίησης με δύο δεσμεύσεις ισότητας

$$\left.\begin{array}{l} \min f(\vec{x}) \\ \vec{x} \in X \subseteq \mathbb{R}^n \\ g_1(\vec{x}) = 0 \\ g_2(\vec{x}) = 0 \end{array}\right\}$$

Θεώρημα Αναγκαία Συνθήκη Πρώτης Τάξης

Έστω $f, g_1, g_2 : X \subseteq \mathbb{R}^n \to \mathbb{R}$ (X ανοικτό) C^1 συναρτήσεις.

η f έχει στο $\vec{x}^* \in X$ δεσμ. τοπ. βελτιστοποίηση υπό τις δεσμεύσεις $g_1(\vec{x}) = 0$, $g_2(\vec{x}) = 0$, και τα διανύσματα $\nabla g_1(\vec{x})$ και $\nabla g_2(\vec{x})$ είναι γραμμ. ανεξάρτητα.

Τότε υπάρχουν $\lambda_1^*, \lambda_2^* \in \mathbb{R}$ τ.ω. $\nabla f(\vec{x}^*) = \lambda_1^* \nabla g_1(\vec{x}^*) + \lambda_2^* \nabla g_2(\vec{x}^*)$

και $g_1(\vec{x}^*) = 0$ και $g_2(\vec{x}^*) = 0$.

Τα λ_1^*, λ_2^* ονομάζονται πολλαπλασιαστές Lagrange.

Ορισμός Δεσμευμένο Κρίσιμο Σημείο

Έστω $f, g_1, g_2 : X \subseteq \mathbb{R}^n \to \mathbb{R}$ (X ανοικτό) C^1 συναρτήσεις. Ένα σημείο $\vec{x}^*$ ονομάζεται δ.κ.σ. της f υπό τις δεσμεύσεις $g_1(\vec{x}) = 0$ και $g_2(\vec{x}) = 0$ όταν υπάρχουν λ_1^* και $\lambda_2^* \in \mathbb{R}$ τ.ω. το σημείο είναι η λύση του συστήματος

$(\vec{x}^*, \lambda_1^*, \lambda_2^*)$

$$\begin{cases} \nabla f(\vec{x}) = \lambda_1 \nabla g_1(\vec{x}) + \lambda_2 \nabla g_2(\vec{x}) \\ g_1(\vec{x}) = 0, \quad g_2(\vec{x}) = 0. \end{cases}$$

Τα σημεία τοπ. δεσμ. βελτιστοποίησης βρίσκονται μεταξύ των κρίσιμων σημείων.

άρα διαφέρονται:

Θεώρημα: Αναγκαία Συνθήκη Πρώτης Τάξης.

Αν η f έχει στο $\vec{x}^* \in X$ δεσμ. τοπ. βελτιστοποιητή υπό τις δεσμεύσεις $g_1(\vec{x})=0$ και $g_2(\vec{x})=0$ και τα διανύσματα $\nabla g_1(\vec{x}^*)$ και $\nabla g_2(\vec{x}^*)$ είναι γραμμ. ανεξάρτητα, τότε το $\vec{x}^*$ είναι δ.κ.σ. της f.

ορισμός: Επιφάνεια του $\mathbb{R}^n$

Έστω $g_1, g_2 : X \subseteq \mathbb{R}^n \to \mathbb{R}$. Το υποσύνολο του $\mathbb{R}^n$

$S = \{ \vec{x} \in X : g_1(\vec{x}) = 0$ και $g_2(\vec{x}) = 0 \}$ ονομάζεται επιφάνεια του $\mathbb{R}^n$ με εξισώσεις $g_1(\vec{x})=0$ και $g_2(\vec{x})=0$.

ορισμός: Εφαπτόμενος Κώνος / Χώρος.

Έστω $g_1, g_2 : X \subseteq \mathbb{R}^n \to \mathbb{R}$ C^1 συναρτήσεις και

$\vec{a} \in S = \{ \vec{x} \in X : g_1(\vec{x}) = 0$ και $g_2(\vec{x}) = 0 \}$ με $\nabla g_1(\vec{a})$ και $\nabla g_2(\vec{a})$ γραμμ. ανεξάρτητα διανύσματα. Το σύνολο

$C(\vec{a}) = \{ \vec{y} \in \mathbb{R}^n : \nabla g_1(\vec{a}) \cdot \vec{y} = 0$ και $\nabla g_2(\vec{a}) \cdot \vec{y} = 0 \}$

ονομάζεται Εφαπτόμενος κώνος της S στο σημείο $\vec{a}$.

- Συνάρτηση Lagrange -

Θεωρούμε τις πραγμ. συναρτήσεις $f, g_1, g_2 : X \subset \mathbb{R}^n \to \mathbb{R}$

Η πραγμ. συνάρτηση

$$L = L(\vec{x}, \lambda_1, \lambda_2) : X \times \mathbb{R}^2 \subseteq \mathbb{R}^{n+2} \to \mathbb{R}$$

η οποία ορίζεται από τον τύπο

$$L(\vec{x}, \lambda_1, \lambda_2) = f(\vec{x}) - \lambda_1 g_1(\vec{x}) - \lambda_2 g_2(\vec{x}), \quad x \in X, \ \lambda_1, \lambda_2 \in \mathbb{R}$$

ονομάζεται συνάρτηση Lagrange των f, g_1, g_2.

Αν οι συναρτήσεις είναι παραγωγίσιμες στο ανοιχτό X, τότε η L είναι παραγωγίσιμη στο $X \times \mathbb{R}^2$ και ισχύουν

$$L_{x_i}(\vec{x}, \lambda_1, \lambda_2) = f_{x_i}(\vec{x}) - \lambda_1 g_{1x_i}(\vec{x}) - \lambda_2 g_{2x_i}(\vec{x})$$

$$L_{\lambda_1}(\vec{x}, \lambda_1, \lambda_2) = -g_1(\vec{x})$$

$$L_{\lambda_2}(\vec{x}, \lambda_1, \lambda_2) = -g_2(\vec{x})$$

Οπότε $\nabla_{\vec{x}} L(\vec{x}, \lambda_1, \lambda_2) = \nabla f(\vec{x}) - \lambda_1 \nabla g_1(\vec{x}) - \lambda_2 \nabla g_2(\vec{x})$

Έτσι ένα σημείο $\vec{x}^* \in X$ είναι δ.σ.σ. ακρ f υπό τις δεσμεύσεις $g_1(\vec{x}) = 0$ και $g_2(\vec{x}) = 0$ τότε και μόνο τότε όταν υπάρχουν $\lambda_1^*, \lambda_2^* \in \mathbb{R}$ τ.ω.

$$\nabla L(\vec{x}^*, \lambda_1^*, \lambda_2^*) = \vec{0}$$

/ δηλαδή το $(\vec{x}^*, \lambda_1^*, \lambda_2^*)$ είναι κρίσιμο σημείο της L

Θεώρημα: Αναγκαίες Συνθήκες Δεύτερης τάξης

Έστω $f, g_1, g_2 : X \subseteq \mathbb{R}^n \to \mathbb{R}$ (X ανοιχτό) C^2 συναρτήσεις

Υποθέτουμε $x^* \in X \to$ σημ. τοπ. ελαχιστοποιητής υπό τις δεσμεύσεις

$$\begin{cases} g_1(\vec{x}) = 0 \\ g_2(\vec{x}) = 0 \end{cases}$$ και $\nabla g_1(\vec{x}^*), g_2(\vec{x})$ γραμμ. ανεξάρτητα.

Τότε υπάρχουν $\lambda_1^*, \lambda_2^* \in \mathbb{R}$ ε.ω.

α) $\nabla_{\vec{x}} L(\vec{x}^*, \lambda_1^*, \lambda_2^*) = \vec{0}$ και

$g_1(\vec{x}^*) = 0$, $g_2(\vec{x}^*) = 0$.

β) ο πίνακας $\nabla^2_{\vec{x}\vec{x}} L(\vec{x}^*, \lambda_1^*, \lambda_2^*)$ είναι θετικά ημιορισμένος στον εφαπτόμενο χώρο $C(\vec{x}^*)$ δηλαδή ισχύει

$$\vec{u}^T \cdot \nabla^2_{\vec{x}\vec{x}} L(\vec{x}^*, \lambda_1^*, \lambda_2^*)\, \vec{u} \geq 0, \text{ για κάθε } \vec{u} \in C(\vec{x}^*)$$

$$= \{ \vec{y} \in \mathbb{R}^n : \nabla g_1(\vec{x}^*) \cdot \vec{y} = 0 \text{ και } \nabla g_2(\vec{x}^*) \cdot \vec{y} = 0 \}$$

Θεώρημα: Ικανές Συνθήκες Δεύτερης Τάξης ③

Έστω $f, g_1, g_2 : X \subseteq \mathbb{R}^n \to \mathbb{R}$ (X ανοιχτό) C^2 συναρτήσεις

και $\vec{x}^* \in X$ με τις ιδιότητες

1) $g_1(\vec{x}^*) = 0$ και $g_2(\vec{x}^*) = 0$ και τα $\nabla g_1(\vec{x}^*)$ και

$g_2(\vec{x}^*)$ είναι γραμμ. ανεξάρτητα.

2) υπάρχουν $\lambda_1^*, \lambda_2^* \in \mathbb{R}$ ώστε

$$\nabla_{\vec{x}} L(\vec{x}^*, \lambda_1^*, \lambda_2^*) = \vec{0}$$

3) ο πίνακας $\nabla^2_{\vec{x}\vec{x}} L(\vec{x}^*, \lambda_1^*, \lambda_2^*)$ είναι θετικά ορισμένος

στον εφαπτόμενο χώρο $C(\vec{x}^*)$

→ Τότε το $\vec{x}^*$ είναι (γνησίως) δεσμ. τοπ. ελαχιστοποιητής της f

υπό τις δεσμεύσεις $g_1(\vec{x}) = 0$ και $g_2(\vec{x}) = 0$.

Θεώρημα: (Εναλλακτική) Ικανή Συνθήκη Δεύτερης Τάξης (Εφαρμογή)

(για $n = 3$)

Έστω $f = f(x_1, x_2, x_3)$, $g_1 = g_1(x_1, x_2, x_3)$ $g_2 = g_2(x_1, x_2, x_3)$

C^2 συναρτήσεις.

και $\vec{x}^*$ ένα δεσμευμένο κρίσιμο σημείο της f υπό τις

δεσμεύσεις $g_1(\vec{x}) = 0$ και $g_2(\vec{x}) = 0$ με $\nabla g_1(\vec{x}^*)$, $\nabla g_2(\vec{x}^*)$

γραμμ. ανεξάρτητα, και L η συνάρτηση Lagrange της f, g_1, g_2

με Hessian Matrix: της L στο σημείο $(\vec{x}^*, \lambda_1^*, \lambda_2^*)$

$$\nabla^2 L(\vec{x}^*, \lambda_1^*, \lambda_2^*) =$$

$$\begin{bmatrix} L_{x_1x_1}(\vec{x}^*, \lambda_1^*, \lambda_2^*) & L_{x_1x_2}(\ldots) & L_{x_1x_3}(\ldots) & -g_{1x_1}(\vec{x}^*) & -g_{2x_1}(\vec{x}^*) \\ L_{x_2x_1}(\ldots) & L_{x_2x_2}(\ldots) & L_{x_2x_3}(\ldots) & -g_{1x_2}(\vec{x}^*) & -g_{2x_2}(\vec{x}^*) \\ L_{x_3x_1}(\ldots) & L_{x_3x_2}(\ldots) & L_{x_3x_3}(\ldots) & -g_{1x_3}(\vec{x}^*) & -g_{2x_3}(\vec{x}^*) \\ -g_{1x_1}(\vec{x}^*) & -g_{1x_2}(\vec{x}^*) & -g_{1x_3}(\vec{x}^*) & 0 & 0 \\ -g_{2x_1}(x^*) & -g_{2x_2}(\vec{x}^*) & -g_{2x_3}(\vec{x}^*) & 0 & 0 \end{bmatrix}$$

Τότε ισχύουν οι ισχυρισμοί:

1) Αν $|\nabla^2 L(\vec{x}^*, \lambda_1^*, \lambda_2^*)| > 0$ η f έχει στο x^* δεσμ. τοπικό ελάχιστο υπό τις δεσμεύσεις $g_1(\vec{x}) = 0$, $g_2(\vec{x}) = 0$

2) Αν $|\nabla^2 L(\vec{x}^*, \lambda_1^*, \lambda_2^*)| < 0$ f στο x^* δεσμ. τοπ. μέγιστο υπό g_1 g_2

3) Αν $|\nabla^2 L(\vec{x}^*, \lambda_1^*, \lambda_2^*)| = 0$ τότε η f μπορεί να έχει ή μπορεί να μην έχει το σημείο $\vec{x}^*$ ως δεσμ. τοπ. βελτιστοποιητή υπό τις δεσμεύσεις $g_1(\vec{x}) = 0$ και $g_2(\vec{x}) = 0$.

Παράδειγμα

Βρείτε τα τοπικά δεσμευμένα τοπικά βελτιστοποιητικά της $f(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2$ υπό τις δεσμεύσεις

$$g_1(x_1, x_2, x_3) = x_1^2 + x_2^2 - x_3^2 = 0 \text{ και}$$

$$g_2(x_1, x_2, x_3) = x_1 + x_2 - x_3 + 2 = 0.$$

Λύση

f, g_1, g_2 C^2 στο $\mathbb{R}^3$. με

$f_{x_1} = 2x_1 \quad f_{x_2} = 2x_2 \quad f_{x_3} = 2x_3$

$f_{x_1x_1} = f_{x_2x_2} = f_{x_3x_3} = 2 \quad , f_{x_ix_j} = 0 \quad , i \neq j$

$g_{1x_1} = 2x_1 \quad , g_{1x_2} = 2x_2 \quad , g_{1x_3} = 2x_3$

$g_{2x_1} = 1 \quad , g_{2x_2} = 1 \quad , g_{2x_3} = -1$

Τα δεσμευμένα κρίσιμα σημεία προκύπτουν ως λύσεις του συστήματος

$$\left.\begin{aligned} 2x_1 - 2\lambda_1 x_1 - 2\lambda_2 &= 0 \\ 2x_2 - 2\lambda_1 x_2 - 2\lambda_2 &= 0 \\ 2x_3 - 2\lambda_1 x_3 - 2\lambda_2 &= 0 \\ x_1^2 + x_2^2 - x_3^2 &= 0 \\ x_1 + x_2 - x_3 + 2 &= 0 \end{aligned}\right\} \ldots$$

$(\vec{x}_1^*, \lambda_1^*, \lambda_2^*) = (-2+\sqrt{2}, -2+\sqrt{2}, -2+2\sqrt{2}, -3+2\sqrt{2}, -24+16\sqrt{2})$

$(\vec{x}_2^*, \lambda_1^*, \lambda_2^*) = (-2-\sqrt{2}, -2-\sqrt{2}, -2-2\sqrt{2}, -3+2\sqrt{2}, -24-16\sqrt{2})$

Άρα ο Hessian Matrix είναι:

$$\begin{bmatrix} 2-2\lambda_1 & 0 & 0 & -2x_1 & -1 \\ 0 & 2-2\lambda_1 & 0 & -2x_2 & -1 \\ 0 & 0 & 2+2\lambda_1 & 2x_3 & 1 \\ -2x_1 & -2x_2 & 2x_3 & 0 & 0 \\ -1 & -1 & 1 & 0 & 0 \end{bmatrix}$$

Στο $\vec{x}_1^*$: $|\nabla^2 L(\vec{x}_1^*, \lambda_1^*, \lambda_2^*)| = 37.49 > 0$

Άρα $\vec{x}_1^*$ → σημ. τοπ. ελαχιστοποίησης της f υπό τις δεσμεύσεις $g_1(\vec{x}) = 0$ και $g_2(\vec{x}) = 0$

Στο $\vec{x}_2^*$

$|\nabla^2 L(\vec{x}_2^*, \lambda_1^*, \lambda_2^*)| = 218.51 > 0$

Άρα $\vec{x}_2^*$ → σημ. τοπ. ελαχιστοποίησης της f υπό τις δεσμεύσεις $g_1(\vec{x}) = 0$ και $g_2(\vec{x}) = 0$

Παράδειγμα.

Βρείτε τους δεσμευμένους τοπ. βελτιστοποιητές της

$f(x_1, x_2, x_3) = 2x_1 + x_2^2 - x_3^2$ υπό τους δεσμούς

$g_1(x_1, x_2, x_3) = x_1 - 2x_2 = 0$ και

$g_2(x_1, x_2, x_3) = x_1 + x_3 = 0$.

Λύση

οι δεσμοί που περιγράφουν οι g_1 και g_2 μπορούν να λυθούν ως προς κάποια απ' τις μεταβλητές.

$$g_1(x_1, x_2, x_3) = x_1 - 2x_2 = 0 \Rightarrow x_2 = \frac{x_1}{2}$$

ενώ $g_2(x_1, x_2, x_3) = x_1 + x_3 = 0 \Rightarrow x_3 = -x_1$

οπότε με αντικατάσταση στη f λαμβάνουμε

$$h(x_1) = f\left(x_1, \frac{x_1}{2}, -x_1\right) = 2x_1 + \frac{x_1^2}{4} - x_1^2 = 2x_1 - \frac{3}{4}x_1^2.$$

και αναζητούμε τους δεσμευμένους τοπ. βελτιστοποιητές της συνάρτησης μιας μεταβλητής $h(x_1)$.

$$h'(x_1) = 2 - \frac{3}{2}x_1$$

η παράγωγος είναι $=0$ εκεί όπου αλλάζει η φορά!

με $h'\left(\frac{4}{3}\right) = 0$. και $h'(x_1) < 0$ για $x_1 > 4/3$

ενώ $h'(x_1) > 0$ για $x_1 < 4/3$

Άρα στο σημείο

$\vec{X}^{*} = \left[\frac{4}{3} \quad \frac{2}{3} \quad -\frac{4}{3} \right]^T$ η συνάρτηση f έχει δέσμ. τοπ. μέγιστο υπό τις δεσμεύσεις $g_1(x_1, x_2, x_3) = 0$ και $g_2(x_1, x_2, x_3) = 0$.

Βελτιστοποίηση

⊚ Προβλήματα Βελτιστοποίησης με ανισοτικούς περιορισμούς

minimize $f(\vec{x})$

$\vec{x} \in X \subseteq \mathbb{R}^n$

subject to $g_i(\vec{x}) = 0, \ i \in \mathcal{E}$

$g_i(\vec{x}) \geq 0, \ i \in I$

ορίζω $\Omega = \{\vec{x} \in X : g_i(\vec{x}) = 0, i \in \mathcal{E}$ και $g_i(\vec{x}) \geq 0, i \in I\}$

άρα

minimize $f(\vec{x})$

$\vec{x} \in \Omega.$

Ορισμός

Το ενεργό σύνολο $A(\vec{x})$ σε ένα εφικτό σημείο $\vec{x} \in \Omega$ αποτελείται από τους δείκτες των ισοτικών περιορισμών του $\mathcal{E}$, και από εκείνους τους δείκτες των ανισοτικών περιορισμών του I για τους οποίους ισχύει $g_i(\vec{x}) = 0$ δηλαδή $A(\vec{x}) = \mathcal{E} \cup \{i \in I : g_i(\vec{x}) = 0\}$

Σε ένα εφικτό σημείο $\vec{x} \in \Omega$, ο ανισοτικός περιορισμός $i \in I$ λέγεται ότι είναι

active αν $g_i(\vec{x}) = 0$

inactive αν $g_i(\vec{x}) > 0$

Παράδειγμα ενός ανισοτικός περιορισμού.

minimize $f(x_1, x_2) = x_1 + x_2$

subject to $2 - x_1^2 - x_2^2 \geq 0$

Λύση $\Omega = \{(x_1, x_2) \in \mathbb{R}^2 : x_1^2 + x_2^2 \leq 2\}$

που είναι ο κύκλος $x_1^2 + x_2^2 = 2$, και το εσωτερικό του.

όπως $\nabla f(x_1, x_2) = [1 \;\; 1]^T$, $\nabla g_1(x_1, x_2) = [-2x_1 \;\; -2x_2]^T$

δηλαδή το $\nabla g_1(x_1, x_2)$ δείχνει προς το εσωτερικό του Ω.

→ Στα σημεία του κύκλου, γνωρίζουμε ότι: local ελαχιστοποιητής $\vec{x}^* = [-1 \;\; -1]^T$ για το εσωτερικό του κύκλου, δεν υπάρχει τοπ. βελτιστοποιητής για τη συνάρτηση $f(x_1, x_2)$.

Στο σημείο x^* τα ∇f και ∇g_1 είναι παράλληλα και ομόρροπα.

→ Οι συνθήκες πρώτης τάξης τις οποίες ικανοποιεί ο βελτιστοποιητής x^* είναι

$$\boxed{\begin{array}{ll} (1) & \nabla_{\vec{x}} L(\vec{x}^*, \lambda_1^*) = \vec{0} \qquad \lambda_1^* \geq 0 \\ (2) & \lambda_1^* g_1(\vec{x}^*) = 0 \\ & \\ \text{όπου} & L(\vec{x}, \lambda) = f(\vec{x}) - \lambda g_1(\vec{x}) \end{array}}$$

η δεύτερη συνθήκη εγγυάται ότι ο λ_1 είναι άγνωστα θετικός, μόνο όταν ο περιορισμός g_1 είναι ενεργός

δηλαδή $g_1(\vec{x}) = 0$, $\vec{x} \in \Omega$

→ Αν $\vec{x}^*$ είναι στο εσωτερικό του Ω τότε $g_1(\vec{x}^*) > 0$ και άρα από την δεύτερη συνθήκη (2) $\lambda_1^* = 0$. και η (1) αναγεται σε $\nabla f(\vec{x}^*) = 0$

Βελτιστοποίηση β.

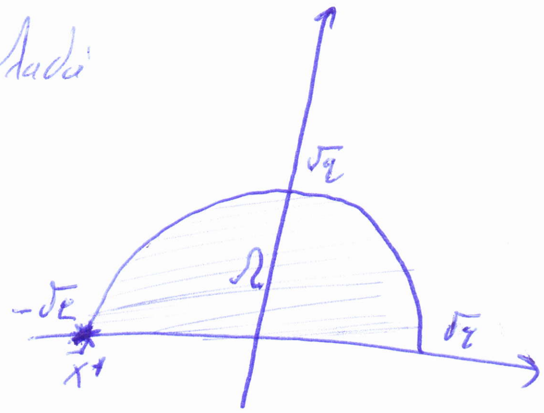
Παράδειγμα: δύο ανισοτικοί περιορισμοί.

$$\left.\begin{array}{ll} \text{minimize} & f(x_1, x_2) = x_1 + x_2 \\ \text{subject to} & 2 - x_1^2 - x_2^2 \geq 0 \\ & x_2 \geq 0 \end{array}\right\}$$

Λύση

$$\Omega = \{(x_1, x_2) \in \mathbb{R}^2 : x_1^2 + x_2^2 \leq 2 \;,\; x_2 \geq 0\}.$$

δηλαδή



$$\nabla f(x_1, x_2) = [1 \quad 1]^T$$

$$\nabla g_1(x_1, x_2) = [-2x_1 \quad -2x_2]^T$$

$$\nabla g_2(x_1, x_2) = [0 \quad 1]^T$$

Από το σχήμα προκύπτει ότι η λύση του προβλήματος είναι $\vec{x}^* = [-\sqrt{2} \quad 0]$. Στο σημείο αυτό, και οι δύο περιορισμοί είναι ενεργοί, διότι $g_1(\vec{x}^*) = 0$, $g_2(\vec{x}^*) = 0$.

Για δύο δεσμεύσεις, η συνάρτηση Lagrange έχει ως εξής.

$$L(\vec{x}, \lambda_1, \lambda_2) = f(\vec{x}) - \lambda_1 g_1(\vec{x}) - \lambda_2 g_2(\vec{x})$$

Άρα έχουμε

$$\left.\begin{array}{l} L(\vec{x}^*, \lambda_1^*, \lambda_2^*) = \vec{0} \qquad \lambda_1^*, \lambda_2^* \geq 0 \\ \lambda_1^* g_1(\vec{x}^*) = 0, \\ \lambda_2^* g_2(\vec{x}^*) = 0. \end{array}\right\}$$

Για $\vec{x}^* = [-\sqrt{2} \quad 0]^T$

$\nabla f(\vec{x}^*) = [1 \quad 1]^T$

$\nabla g_1(\vec{x}^*) = [2\sqrt{2} \quad 0]^T$

$\nabla g_2(\vec{x}^*) = [0 \quad 1]^T$

και $\nabla_{\vec{x}} L(\vec{x}^*, \lambda_1^*, \lambda_2^*) = \vec{0} \Rightarrow$

$$\begin{bmatrix} 1 - 2\sqrt{2}\lambda_1 \\ 1 - \lambda_2^* \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow$$

$\Rightarrow \lambda_1^* = \dfrac{1}{2\sqrt{2}}$ και $\lambda_2^* = 1$ $\Big\}$ $\lambda_1^*, \lambda_2^* > 0$ δηλαδή ικανοποιείται η 1η συνθήκη.

Κάτι που δεν ισχύει για τα γειτ. σημεία (αφού δεν είναι βελτιστοποιητές).

Βελτιστοποίηση J.

Αναγκαίες συνθήκες 1ης τάξης του γεν. προβλήματος βελτιστοποίησης.

Ορισμός (κανονικό σημείο).

$g_i : X \subseteq \mathbb{R}^n \to \mathbb{R}$ C^1 συναρτήσεις με $i \in \mathcal{E} \cup \mathcal{I}$

Ένα σημείο $x^* \in X$ για το οποίο ισχύει ότι το σύνολο $\{\nabla g_i(\vec{x}^*), i \in A(\vec{x}^*)\}$ είναι γραμμ. ανεξάρτητο, ονομάζεται κανονικό σημείο.

$$(A(\vec{x}^*) = \mathcal{E} \cup \{i \in \mathcal{I} : g_i(\vec{x}^*) = 0\})$$

Θεώρημα. (Αναγκαίες συνθήκες 1ης τάξης).

Έστω $f, g_i : X \subseteq \mathbb{R}^n \to \mathbb{R}$ C^1 συναρτήσεις με $i \in \mathcal{E} \cup \mathcal{I}$ και $\vec{x}^* \in X$ τοπ. βελτιστοποιητής και $\vec{x}^*$ κανονικό σημείο.

Τότε υπάρχει ένα διάνυσμα $\vec{\lambda}^*$ με συνιστώσες $\lambda_i^* \in \mathbb{R}$, $i \in \mathcal{E} \cup \mathcal{I}$ ώστε.

!!!

$$\left.\begin{aligned}
&\nabla_{\vec{x}} L(\vec{x}^*, \vec{\lambda}^*) = 0 \\
&g_i(\vec{x}^*) = 0 \quad i \in \mathcal{E} \\
&g_i(\vec{x}^*) \geq 0 \quad i \in \mathcal{I} \\
&\lambda_i^* \geq 0 \quad i \in \mathcal{I} \\
&\lambda_i^* g_i(\vec{x}^*) = 0, \quad \forall i \in \mathcal{E} \cup \mathcal{I}
\end{aligned}\right\}$$

KKT conditions.

Karush-Kuhn-Tucker conditions

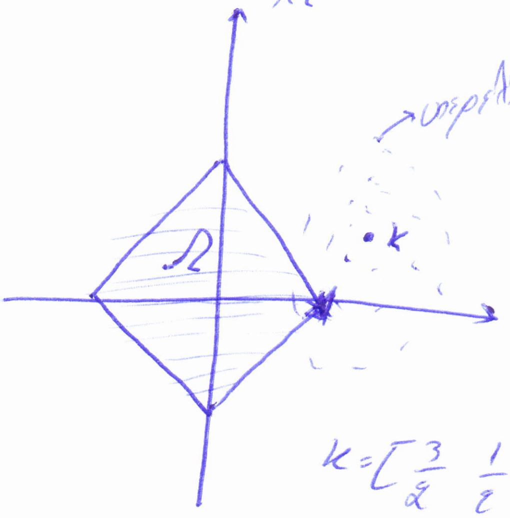
↳ σημαίνει ότι είτε ο περιορισμός είναι ενεργός, είτε $\lambda_i^* = 0$ είτε και τα δύο.

με $\nabla_{\vec{x}} L(\vec{x}^*, \vec{\lambda}^*) = \nabla f(\vec{x}^*) - \lambda_i \nabla g_i(\vec{x}^*) = 0$.

Παράδειγμα: !!!

Ποιοι περιορισμοί είναι ενεργοί.

minimize $f(x_1, x_2) = (x_1 - 3/2)^2 + (x_2 - 1/2)^4$



subject to

$g_1(x_1, x_2) = 1 - x_1 - x_2 \geq 0$

$g_2(x_1, x_2) = 1 - x_1 + x_2 \geq 0$

$g_3(x_1, x_2) = 1 - x_1 - x_2 \geq 0$

$\kappa = [\frac{3}{2} \ \frac{1}{2}]^T$ $g_4(x_1, x_2) = 1 + x_1 + x_2 \geq 0$.

$\nabla f(\vec{x}) = [2(x_1 - 3/2) \quad 4(x_2 - 1/2)^3]^T \neq [0 \ 0]^T$.

Το $\vec{x}^* = [1 \ 0]$ είναι ο βελτιστοποιητής.

Στο $\vec{x}^*$ είναι ενεργοί μόνο οι δύο πρώτοι περιορισμοί.

$\nabla f(\vec{x}^*) = [-1 \ \ -1/2]^T$ Οι KKT συνθήκες ικανοποιούνται

διότι οι λοιποί είναι ανενεργοί

$\nabla g_1(\vec{x}^*) = [-1 \ \ -1]^T$ για $\lambda^* = [3/4 \ \ 1/4 \ \ 0 \ \ 0]$

$\nabla g_2(\vec{x}^*) = [-1 \ \ 1]^T$

λ_1^* λ_2^* → active constraints

λ_3^* λ_4^* → inactive constraints.

Βελτιστοποίηση δ. Παράδειγμα !!! Ποια είναι τα εφικτά σημεία.

minimize $f(x_1, x_2) = (x_1 - 2)^2 + 2(x_2 - 1)^2$

Subject to

$$g_1(x_1, x_2) = 3 - x_1 - 4x_2 \geq 0$$

$$g_2(x_1, x_2) = x_1 - x_2 \geq 0 .$$

Λύση

$$L(x_1, x_2, \lambda_1, \lambda_2) = f(x_1, x_2) - \lambda_1 g_1(x_1, x_2) - \lambda_2 g_2(x_1, x_2)$$

$$\nabla_{\vec{x}} L(x_1, x_2, \lambda_1, \lambda_2) = \cancel{\nabla f(x_1, x_2)} [2(x_1 - 2) + \lambda_1 - \lambda_2 \quad 4(x_2 - 1) + 4\lambda_1 + \lambda_2]^T$$

άρα KKT conditions:

$$\begin{cases} 2(x_1 - 2) + \lambda_1 - \lambda_2 = 0 \\ 4(x_2 - 1) + 4\lambda_1 + \lambda_2 = 0 \\ 3 - x_1 - 4x_2 \geq 0 \\ x_1 - x_2 \geq 0 \\ \lambda_1 \geq 0, \ \lambda_2 \geq 0 \\ \lambda_1 (3 - x_1 - x_2 4) = 0 \\ \lambda_2 (x_1 - x_2) = 0 \end{cases}$$

άρα

α) $\lambda_1 = 0, \ \lambda_2 = 0$

$\Rightarrow x_1 = 2 \quad x_2 = 1$ όμως $g_1(2, 1) = -3 < 0$

άρα το $[2 \ 1]^T$ δεν είναι λύση.

γ) $\lambda_1 = 0$, $x_1 - x_2 = 0$

$\Rightarrow x_1 = 4/3$, $x_2 = 4/3$

όμως $g_1(4/3, 4/3) = -4/3 < 0$

άρα το $[4/3 \;\; 4/3]$ δεν είναι εφικτό σημείο

δ) $\lambda_2 = 0$, $3 - x_1 - x_2 = 0$

$\Rightarrow x_1 = 5/3$, $x_2 = 1/3$

άρα $\lambda_1 = 2/3 \geq 0$ ✓

$x_1 - x_2 \geq 0$ ✓.

άρα το $\vec{x}^* = [5/3 \;\; 1/3]$ είναι εφικτό σημείο.

το μόνο σημείο που ικανοποιεί τις ΚΚΤ. Είναι βέλτιστο? βλέπε επόμενη σελίδα...

δ) $3 - x_1 - 4x_2 = 0$, $x_1 - x_2 = 0$

$\Rightarrow x_1 = 3/5$, $x_2 = 3/5$

$\lambda_1 = \frac{22}{25} > 0$ ✓

$\lambda_2 = -\frac{48}{25} < 0$ ✗

(πρέπει $\lambda_2 \geq 0$)

άρα το $[3/5 \;\; 3/5]$ δεν είναι εφικτό.

Βελτιστοποίηση Ε

αρκεί

1) $\vec{x}^*$: ομαλό σημείο

2) υπάρχει διάνυσμα $\vec{\lambda}^*$ τ.ω. να ικανοποιούνται οι ΚΚΤ στο $(\vec{x}^*, \vec{\lambda}^*)$

3) ο $\nabla^2_{\vec{x}\vec{x}}(\vec{x}^*, \vec{\lambda}^*)$ να είναι θετικά ορισμένος.

αρκεί ότι το x^* είναι τοπ. ελαχιστοποιητής του γεν. προβλ. βελτιστοποίησης.