08/12/2021 Μάθημα 20

8| Γραμμικός Κλασματικός Προγραμματισμός

Ένα πρόβλημα της μορφής

$$\max \frac{c_1 x_1 + c_2 x_2 + \ldots + c_n x_n + c_0}{d_1 x_1 + d_2 x_2 + \ldots + d_n x_n + d_0}$$

$$\text{υπό} \quad a_{i1} x_1 + a_{i2} x_2 + \ldots + a_{in} x_n \leq b_i \ , \ i = 1, \ldots, m$$

$$x_j \geq 0 \ , \ j = 1, \ldots, n$$

με $c_j, d_j \geq 0$, $j = 0, \ldots, n$, $a_{ij} \in \mathbb{R}$, $b_i \in \mathbb{R}$

ονομάζεται πρόβλημα γραμμικού κλασματικού προγραμματισμού.

Τέτοια προβλήματα προκύπτουν όταν θέλουμε να μεγιστοποιήσουμε την απόδοση (κέρδους/ώρες εργασίας ή απόδοση επένδυσης/χρήματα)

Το πρόβλημα σε πινακική μορφή γράφεται

$$\max \frac{\underline{c}' \cdot \underline{x} + c_0}{\underline{d}' \cdot \underline{x} + d_0}$$

$$\text{υπό} \quad \underline{a}_i' \cdot \underline{x} \leq b_i \ , \ i = 1, \ldots, m$$

$$\underline{x} \geq \underline{0}$$

όπου $\underline{c}, \underline{d}, \underline{x} \in \mathbb{R}^n_+$ και $a_i \in \mathbb{R}^n$, $b_i \in \mathbb{R}$

Θέτω $\underline{y} = \dfrac{\underline{x}}{\underline{d}'\underline{x} + d_0}$ και $t = \dfrac{1}{\underline{d}'\underline{x} + d_0}$.

Έχουμε $\dfrac{\underline{y}}{t} = \underline{x} \Rightarrow \underline{y} = \underline{x}\, t$.

Τότε $\max \underline{c}'\underline{y} + c_0 t$

υπό $\underline{a}_i' \cdot \underline{y} - b_i t \leq 0,\ i = 1, \ldots, m$

$\underline{y} \geq \underline{0}$

$t \geq 0$

Που είναι π.γ.π..

5.3 Προβλήματα χωρίς περιορισμούς

Το πρόβλημα $\max f(\underline{x})$

υπό $\underline{x} \in \mathbb{R}^n$

είναι πρόβλημα χωρίς περιορισμούς.

Ορισμοί (Ολικό - Τοπικό μέγιστο)

Έστω $f: \mathbb{R}^n \to \mathbb{R}$. Το $\underline{x}^* \in \mathbb{R}^n$ είναι

- ολικό μέγιστο αν $f(\underline{x}^*) \geq f(\underline{x})\ \forall \underline{x} \in \mathbb{R}^n$
- τοπικό μέγιστο αν $\exists \varepsilon > 0 : f(\underline{x}^*) \geq f(\underline{x})\ \forall \underline{x}$ με $\|\underline{x} - \underline{x}^*\| < \varepsilon$.

Ορισμοί: Έστω $f: \mathbb{R}^n \to \mathbb{R}$ δύο φορές συνεχώς παραγωγίσιμη.

- Το ανάδελτα ή η βαθμίδα της f στο $\underline{x} \in \mathbb{R}^n$ είναι
  $\nabla f(\underline{x}) = \left(\dfrac{\partial f}{\partial x_1}(\underline{x}), \ldots, \dfrac{\partial f}{\partial x_n}(\underline{x}) \right)'$

- Ο Εσσιανός πίνακας της f στο $\underline{x} \in \mathbb{R}^n$ είναι
  $Hf(\underline{x}) = \left(\dfrac{\partial^2 f}{\partial x_i \partial x_j}(\underline{x}) \right)_{i,j = 1, \ldots, n}$.

Ανάπτυγμα Taylor

Έστω $f: \mathbb{R}^n \to \mathbb{R}$ δυο φορές συνεχώς παραγωγίσιμη.

- Το ανάπτυγμα Taylor $1^{ης}$ τάξης γύρω από το $\underline{x}^0 \in \mathbb{R}^n$ είναι:

$$f(\underline{x}) = f(\underline{x}^0) + \nabla f(\underline{x}^0)'(\underline{x}-\underline{x}^0) + \|\underline{x}-\underline{x}^0\| a_{\underline{x}^0}(\underline{x}-\underline{x}^0)$$

όπου $\lim_{\underline{x}\to\underline{x}^0} a_{\underline{x}^0}(\underline{x}-\underline{x}^0) = 0$

- Το ανάπτυγμα Taylor $2^{ης}$ τάξης γύρω από το $\underline{x}^0 \in \mathbb{R}^n$ είναι:

$$f(\underline{x}) = f(\underline{x}^0) + \nabla f(\underline{x}^0)'(\underline{x}-\underline{x}^0) + \frac{1}{2}(\underline{x}-\underline{x}^0)' Hf(\underline{x}^0)(\underline{x}-\underline{x}^0) + \|\underline{x}-\underline{x}^0\|^2 a_{\underline{x}^0}(\underline{x}-\underline{x}^0)$$

όπου $\lim_{\underline{x}\to\underline{x}^0} a_{\underline{x}^0}(\underline{x}-\underline{x}^0) = 0$

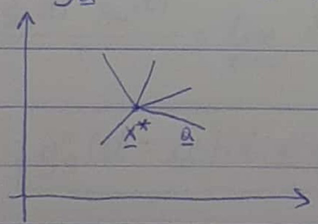
Κίνηση πάνω σε ευθεία

Έστω $f: \mathbb{R}^n \to \mathbb{R}$ δυο φορές συνεχώς παραγωγίσιμη συνάρτηση, $\underline{x}^* \in \mathbb{R}^n$ και $\underline{a} \in \mathbb{R}^n$

↑ σημεία ↑ κατεύθυνση

Έστω $g_{\underline{a}}(t) = f(\underline{x}^* + t\underline{a})$, $t \geq 0$, ο περιορισμός της f στην ημιευθεία που διέρχεται από το $\underline{x}^*$ και έχει κατεύθυνση $\underline{a}$.

Τότε: i) $g_{\underline{a}}'(0) = \nabla f(\underline{x}^*)'\underline{a}$



ii) $g_{\underline{a}}''(0) = \underline{a}' Hf(\underline{x}^*)\underline{a}$.

Αποδείξεις:

i) Το ανάπτυγμα Taylor της f γύρω από το $\underline{x}^*$ είναι

$$f(\underline{x}) = f(\underline{x}^*) + \nabla f(\underline{x}^*)'(\underline{x}-\underline{x}^*) + \|\underline{x}-\underline{x}^*\| a_{\underline{x}^*}(\underline{x}-\underline{x}^*).$$

για $\underline{x} = \underline{x}^* + \underline{a}t$, $t \geq 0$, έχουμε

$$f(\underline{x}^*+\underline{a}t) = f(\underline{x}^*) + \nabla f(\underline{x}^*)'(\underline{x}^*+\underline{a}t-\underline{x}^*) + \|\underline{x}^*+\underline{a}t-\underline{x}^*\| a_{\underline{x}^*}(\underline{x}^*+\underline{a}t-\underline{x}^*)$$

$$\Rightarrow g_{\underline{a}}(t) = g_{\underline{a}}(0) + t\nabla f(\underline{x}^*)'\underline{a} + t\|\underline{a}\| a_{\underline{x}^*}(\underline{a}t)$$

→

$$\Rightarrow \lim_{t\to 0^+} \frac{g_{\underline{a}}(t)-g_{\underline{a}}(0)}{t} = \lim_{t\to 0^+} \frac{\cancel{t}\,\nabla f(\underline{x}^*)'\underline{a}}{\cancel{t}} + \lim_{t\to 0^+} \frac{\cancel{t}\,\|\underline{a}\|\,a_{\underline{x}^*}(\underline{a}t)}{\cancel{t}}$$

$$\Longrightarrow g'_{\underline{a}}(0) = \nabla f(\underline{x}^*)'\underline{a} + 0$$

$$\Longrightarrow g'_{\underline{a}}(0) = \nabla f(\underline{x}^*)'\underline{a}\,.$$

Scanned with CamScanner

<u>Σημείωση</u> Άρα το $\nabla f(\underline{x}^*)'\underline{a}$ δίνει τον ρυθμό μεταβολής της f όταν ξεκινάμε από το $\underline{x}^*$ και κινούμαστε προς τη κατεύθυνση $\underline{a}$.

ii) Το ανάπτυγμα Taylor $2^{ης}$ τάξης γύρω από το $\underline{x}^*$ είναι

$$f(x) = f(\underline{x}^*) + \nabla f(\underline{x}^*)'(\underline{x}-\underline{x}^*) + \frac{1}{2}(\underline{x}-\underline{x}^*)' Hf(\underline{x}^*)(\underline{x}-\underline{x}^*)$$

Για $\underline{x} = \underline{x}^* + \underline{a}t$, $t \geq 0$

$$f(\underline{x}^*+\underline{a}t) = f(\underline{x}^*) + \nabla f(\underline{x}^*)'(\cancel{\underline{x}^*}+\underline{a}t-\cancel{\underline{x}^*}) + \frac{1}{2}(\cancel{\underline{x}^*}+\underline{a}t-\cancel{\underline{x}^*})' Hf(\underline{x}^*)(\cancel{\underline{x}^*}+\underline{a}t-\cancel{\underline{x}^*})$$

Εφόσον $g_{\underline{a}}(t) = f(\underline{x}^*+\underline{a}t)$,

$$g_{\underline{a}}(t) = g_{\underline{a}}(0) + t\nabla f(\underline{x}^*)'\underline{a} + \frac{t^2}{2}\underline{a}' Hf(\underline{x}^*)\underline{a}$$

και $$g_{\underline{a}}(t+h) = g_{\underline{a}}(0) + (t+h)\nabla f(\underline{x}^*)'\underline{a} + \frac{(t+h)^2}{2}\underline{a}' Hf(\underline{x}^*)\underline{a}$$

Αφαιρώ κατά μέλη:

$$g_{\underline{a}}(t+h) - g_{\underline{a}}(t) = h\nabla f(\underline{x}^*)'\underline{a} + \frac{1}{2}(2th+h^2)\underline{a}' Hf(\underline{x}^*)'\underline{a}$$

$$\Rightarrow \lim_{h\to 0^+} \frac{g_{\underline{a}}(t+h) - g_{\underline{a}}(t)}{h} = \lim_{h\to 0^+} \frac{\cancel{h}\nabla f(\underline{x}^*)'\underline{a}}{\cancel{h}} + \lim_{h\to 0^+} \frac{\frac{1}{2}(2t\cancel{h}+h^2)\underline{a}' Hf(\underline{x}^*)'\underline{a}}{\cancel{h}}$$

$$\Rightarrow g_{\underline{a}}'(t) = \underbrace{\nabla f(\underline{x}^*)\underline{a}}_{g_{\underline{a}}'(0)} + \frac{1}{\cancel{2}}\cancel{2}t\underline{a}' Hf(\underline{x}^*)'\underline{a}$$

$$\Rightarrow \lim_{t\to 0^+} \frac{g_{\underline{a}}'(t) - g_{\underline{a}}'(0)}{t} = \lim_{t\to 0^+} \frac{\cancel{t}\underline{a}' Hf(\underline{x}^*)'\underline{a}}{\cancel{t}}$$

$$\boxed{g_{\underline{a}}''(0) = \underline{a}' Hf(\underline{x}^*)\underline{a}}$$

Άρα, το $\underline{a}' Hf(\underline{x}^*)\underline{a}$ μας δείχνει αν είναι κυρτή ή κοίλη η f όταν ξεκινάμε από το $\underline{x}^*$ και κινούμαστε προς το $\underline{a}$.

<u>Ορισμοί</u> (Θετικά - Αρνητικά (ημι)ορισμένος πίνακας)

Έστω $A = (a_{ij})$ ένας $n\times n$ πίνακας.

Τότε η συνάρτηση $Q_A(\underline{x}) = \underline{x}' A \underline{x}$ είναι η τετραγωνική μορφή του πίνακα A. Ο A λέγεται

- <u>Θετικά ορισμένος</u> ανν $Q_A(\underline{x}) = \underline{x}' A \underline{x} > 0 \quad \forall \underline{0} \neq \underline{x} \in \mathbb{R}^n$
- <u>Θετικά ημιορισμένος</u> ανν $Q_A(\underline{x}) = \underline{x}' A \underline{x} \geq 0 \quad \forall \underline{x} \in \mathbb{R}^n$

- $\square$ <u>αρνητικά ορισμένος</u> ανν $Q_A(\underline{x}) = \underline{x}'A\underline{x} < 0 \quad \forall \underline{0} \neq \underline{x} \in \mathbb{R}^n$
- $\square$ <u>αρνητικά ημιορισμένος</u> ανν $Q_A(\underline{x}) = \underline{x}'A\underline{x} \leq 0 \quad \forall \underline{x} \in \mathbb{R}^n$
- $\square$ <u>μη-ορισμένος</u> ανν $\exists \underline{x}, \underline{y} \in \mathbb{R}^n : \underline{x}'A\underline{x} > 0$ και $\underline{y}'A\underline{y} < 0$.

<u>Θεώρημα</u>

- Ο A είναι θετικά (ημι)ορισμένος ανν ο $-A$ είναι αρνητικά (ημι)ορισμένος.
- Ο A είναι θετικά ορισμένος (ή ημιορισμένος) αν οι άνω τετραγωνικοί υποπίνακές του έχουν θετικές (ή μη αρνητικές) ορίζουσες.
- Ο A είναι θετικά (ημι)ορισμένος αν οι ιδιοτιμές του είναι θετικές (≥ 0).
- $d_1 \leq \frac{\underline{x}'A\underline{x}}{\underline{x}'\underline{x}} \leq d_n$ όπου d_1 η μικρότερη ιδιοτιμή του A
  d_n η μεγαλύτερη ιδιοτιμή του A

<u>Θεώρημα</u> (Αναγκαία συνθήκη τοπικού μέγιστου)

Έστω $f: \mathbb{R}^n \to \mathbb{R}$ 2 φορές συνεχώς παραγωγίσιμη με τοπικό μέγιστο $\underline{x}^*$. Τότε

(i) $\nabla f(\underline{x}^*) = \underline{0}$

(ii) $Hf(\underline{x}^*)$ αρνητικά ημιορισμένος

<u>Απόδειξη:</u>

(i) Εφόσον $\underline{x}^*$ τοπικό μέγιστο $\Rightarrow g_{\underline{a}}'(0) \leq 0 \Rightarrow \nabla f(\underline{x}^*)'\underline{a} \leq 0 \quad \forall \underline{a} \in \mathbb{R}^n$.

Θεωρώντας $\underline{a} = e_i$ και $\underline{a} = -e_i$ παίρνω

$$\left.\begin{array}{l} \nabla f(\underline{x}^*)'e_i \leq 0 \Rightarrow \frac{\partial f}{\partial x_i}(\underline{x}^*) \leq 0 \\ \nabla f(\underline{x}^*)'(-e_i) \leq 0 \Rightarrow -\frac{\partial f}{\partial x_i}(\underline{x}^*) \leq 0 \end{array}\right\} \Longrightarrow \frac{\partial f}{\partial x_i}(\underline{x}^*) = 0,\ \forall i = 1, 2, \dots, n.$$ Άρα, $\nabla f(\underline{x}^*) = 0$.

(ii) Θέλω ν.δ.ο. $\underline{a}'Hf(\underline{x}^*)\underline{a} \leq 0 \quad \forall \underline{a} \in \mathbb{R}^n$.

Το ανάπτυγμα Taylor $2^{ης}$ τάξης για την $g_{\underline{a}}(t)$ $= f(\underline{x}^* + \underline{a}t)$ γύρω από το 0

$$g_{\underline{a}}(t) = g_{\underline{a}}(0) + g_{\underline{a}}'(0) \cdot t + \frac{1}{2} g_{\underline{a}}''(0) \cdot t^2 + t^2 a_0(t)$$

Όμως, $g_{\underline{a}}'(0) = \nabla f(\underline{x}^*)'\underline{a} = 0$

$$g_{\underline{a}}(t) - g_{\underline{a}}(0) = \frac{1}{2} g_{\underline{a}}''(0) t^2 + t^2 a_0(t)$$

Αν $g_{\underline{a}}''(0)>0$ τότε για μικρά t θα είχαμε $\overset{f(\underline{x}^*+\underline{a}t)}{g_{\underline{a}}(t)} - \overset{f(\underline{x}^*)}{g_{\underline{a}}(0)} > 0$

Άτοπο γιατί $\underline{x}^*$ βέλτιστο.

Άρα $g_{\underline{a}}''(0)\leq 0 \quad \forall \underline{a}\in\mathbb{R}^n \Rightarrow \underline{a}'Hf(\underline{x}^*)\underline{a}\leq 0 \quad \forall \underline{a}\in\mathbb{R}^n \Rightarrow$

$\Rightarrow Hf(\underline{x}^*)$ αρνητικά ημιορισμένος.

13/12/2021 Μάθημα 21

<u>Θεώρημα</u> (Ικανή συνθήκη για τοπικό μέγιστο)

Έστω $f:\mathbb{R}^n\longrightarrow\mathbb{R}$ δύο φορές συνεχώς παραγωγίσιμη και $\underline{x}^*\in\mathbb{R}^n$.

$$\left.\begin{array}{l}\nabla f(\underline{x}^*)=0\\ Hf(\underline{x}^*) \text{ αρνητικά ορισμένος}\end{array}\right\}\Rightarrow \underline{x}^* \text{ είναι τοπικό μέγιστο}$$

<u>Απόδειξη:</u>

Το ανάπτυγμα Taylor 2ης τάξης γύρω από το $\underline{x}^*$:

$$f(\underline{x})=f(\underline{x}^*)+\underbrace{\nabla f(\underline{x}^*)}_{0}(\underline{x}-\underline{x}^*)+\frac{1}{2}(\underline{x}-\underline{x}^*)Hf(\underline{x}^*)(\underline{x}-\underline{x}^*)+$$
$$+\|\underline{x}-\underline{x}^*\|^2 a_{\underline{x}^*}(\underline{x}-\underline{x}^*).$$

για $\underline{x}=\underline{x}^*+\underline{a}$:

$$f(\underline{x}^*+\underline{a})=f(\underline{x}^*)+\frac{1}{2}\underline{a}'Hf(\underline{x}^*)\underline{a}+\|\underline{a}\|^2 a_{\underline{x}^*}(\underline{a}) \quad (1)$$

$Hf(\underline{x}^*)$ αρνητικά ορισμένος $\Rightarrow$

$-Hf(\underline{x}^*)$ θετικά ορισμένος $\Rightarrow$

$-Hf(\underline{x}^*)$ έχει θετικές ιδιοτιμές.

Όμως, $d_1\leq \dfrac{\underline{a}'(-Hf(\underline{x}^*))\underline{a}}{\underline{a}'\underline{a}}$ όπου $d_1>0$ η μικρότερη ιδιοτιμή

$$\Rightarrow d_1\|\underline{a}\|^2\leq -\underline{a}'Hf(\underline{x}^*)\underline{a}\Rightarrow$$
$$\underline{a}'Hf(\underline{x}^*)\underline{a}\leq -d_1\|\underline{a}\|^2 \quad (2)$$

$$(1)\overset{(2)}{\Longrightarrow} f(\underline{x}^*+\underline{a})\leq f(\underline{x}^*)-\frac{1}{2}d_1\|\underline{a}\|^2+\|\underline{a}\|^2 a_{\underline{x}^*}(\underline{a})\Rightarrow$$

$$\Rightarrow f(\underline{x}^*+\underline{a})-f(\underline{x}^*)\leq -\frac{1}{2}d_1\|\underline{a}\|^2+\|\underline{a}\|^2 a_{\underline{x}^*}(\underline{a})$$

Για μικρά $\underline{a}$ $(\|\underline{a}\|<\varepsilon)$,
θα έχουμε $f(\underline{x}^*+\underline{a})-f(\underline{x}^*)\le 0$

$\Rightarrow f(\underline{x}^*+\underline{a})\le f(\underline{x}^*)$ για $\underline{a}$ με $\|\underline{a}\|<\varepsilon$

$\Rightarrow \underline{x}^*$ τοπικό μέγιστο.

<u>Παράδειγμα</u>

$f(x_1,x_2,x_3)=x_1+2x_3+x_2x_3-x_1^2-x_2^2-x_3^2$.

Να βρεθεί κάποιο τοπικό μέγιστο.

<u>Λύση</u>

$\nabla f(x_1,x_2,x_3)=(1-2x_1,\ x_3-2x_2,\ 2+x_2-2x_3)$

$$Hf(x_1,x_2,x_3)=\begin{bmatrix}-2&0&0\\0&-2&1\\0&1&-2\end{bmatrix}$$

$$-Hf(x_1,x_2,x_3)=\begin{bmatrix}2&0&0\\0&2&-1\\0&-1&2\end{bmatrix}$$

$$\left.\begin{aligned}&|2|=2>0\\&\begin{vmatrix}2&0\\0&2\end{vmatrix}=4-0=4>0\\&\begin{vmatrix}2&0&0\\0&2&-1\\0&-1&2\end{vmatrix}=2\cdot\begin{vmatrix}2&-1\\-1&2\end{vmatrix}=2(4-1)=6>0\end{aligned}\right\}$$

$\Rightarrow -Hf(\underline{x})$ θετικά ορισμένος

$\Rightarrow Hf(\underline{x})$ αρνητικά ορισμένος $\forall \underline{x}\in\mathbb{R}^3$

$$\nabla f(x_1,x_2,x_3)=\underline{0}\Rightarrow\begin{cases}1-2x_1=0\\x_3-2x_2=0\\2+x_2-2x_3=0\end{cases}\Rightarrow\begin{cases}x_1=\frac{1}{2}\\x_3=2x_2\\2-3x_2=0\end{cases}\Rightarrow\begin{cases}x_1=\frac{1}{2}\\x_3=\frac{4}{3}\\x_2=\frac{2}{3}\end{cases}$$

Άρα, στο $\underline{x}^*=\begin{pmatrix}\frac{1}{2}\\\frac{2}{3}\\\frac{4}{3}\end{pmatrix}$ $\nabla f(\underline{x}^*)=\underline{0}$ και $Hf(\underline{x}^*)$ αρν. ορισμένος $\Rightarrow$

$\Rightarrow \underline{x}^*$ τοπικό μέγιστο.

Θεώρημα

$\left.\begin{array}{l} f:\mathbb{R}^n \longrightarrow \mathbb{R} \\ f \text{ κοίλη} \\ F \text{ κυρτό σύνολο} \\ \underline{x}^* \text{ τοπικό μέγιστο} \end{array}\right\} \Longrightarrow \underline{x}^* \text{ ολικό μέγιστο}$

Απόδειξη:

Εφόσον $\underline{x}^*$ τοπικό μέγιστο, $\exists \varepsilon>0 : f(\underline{x}^*) \geq f(\underline{z})$ $\forall \underline{z}\in\mathbb{R}^n$ με $\|\underline{x}^*-\underline{z}\|<\varepsilon$.

Έστω ότι $\underline{x}^*$ δεν είναι ολικό μέγιστο. Τότε,

$\exists y\in F : f(y) > f(\underline{x}^*)$.

Εφόσον F κυρτό, τα $d\underline{x}^*+(1-d)y\in F$ $\forall d\in(0,1)$.

Θεωρούμε $\underline{z}_0 = d_0\underline{x}^*+(1-d_0)y\in F$ με $\|\underline{x}^*-\underline{z}_0\|<\varepsilon$. Τότε, $f(\underline{x}^*)\geq f(\underline{z}_0)$.

Όμως $f(\underline{z}_0) = f(d_0\underline{x}^*+(1-d_0)y) \overset{f \text{ κοίλη}}{\geq} d_0 f(\underline{x}^*)+(1-d_0)f(y) \geq$

$\overset{f(y)>f(x^*)}{>} d_0 f(\underline{x}^*)+(1-d_0)f(\underline{x}^*) = f(\underline{x}^*)$.

Άρα, $f(\underline{z}_0) > f(\underline{x}^*)$.

Άτοπο. ∎

Θεώρημα (Κριτήριο Κυρτότητας)

Έστω $f:\mathbb{R}^n\longrightarrow\mathbb{R}$. Η f είναι κυρτή (κοίλη) αν και μόνο αν οι $Hf(\underline{x})$ είναι θετικά (αρνητικά) ημιορισμένοι $\forall \underline{x}\in\mathbb{R}^n$.

Θεώρημα

Έστω $g_i:\mathbb{R}^n\longrightarrow\mathbb{R}$, $i=1,\dots,p$ είναι κυρτές συναρτήσεις.

Τότε το $F=\{\underline{x}\in\mathbb{R}^n : g_i(\underline{x})\leq 0,\ i=1,2,\dots,p\}$ είναι κυρτό σύνολο

Απόδειξη:

Έστω $\underline{x}_1,\underline{x}_2\in F$.

$\underline{x}_1\in F \Rightarrow g_i(\underline{x}_1)\leq 0$ $\forall i=1,\dots,p$

$\underline{x}_2\in F \Rightarrow g_i(\underline{x}_2)\leq 0$ $\forall i=1,2,\dots,p$

Έστω $\underline{y}=d\underline{x}_1+(1-d)\underline{x}_2$, $d\in(0,1)$.

Για ν.δ.ο. το F είναι κυρτό, αρκεί ν.δ.ο. $\underline{y}\in F$.

Έχουμε $g_i(y) = g_i(\lambda \underline{x}_1 + (1-\lambda)\underline{x}_2) \overset{g_i \text{ κυρτή}}{\leq} \lambda g_i(\underline{x}_1) + (1-\lambda) g_i(\underline{x}_2) \leq 0, \forall i=1,\dots,p.$
Άρα, $\underline{y} \in F$. ∎

<u>Παράδειγμα</u> (συνέχεια)

$\left.\begin{array}{l} \underline{x}^* \text{ τοπικό μέγιστο} \\ Hf(\underline{x}) \text{ αρνητικά ορισμένος} \Rightarrow f \text{ κοίλη} \end{array}\right\} \underline{x}^* \text{ ολικό μέγιστο}$

<u>Παρατήρηση</u>

Δεν είναι πάντα εύκολο να λύνουμε το σύστημα $\nabla f(\underline{x}) = \underline{0}$.

5.3.1 <u>Αλγόριθμος βαθμίδας</u>

Έστω ότι έχουμε μια κοίλη συνάρτηση.
Ξεκινάμε από σημείο $\underline{x}_0$ και επιλέγουμε κάποια ακρίβεια ε.

<u>Προς ποια κατεύθυνση θα κινηθούμε;</u>
Θέλουμε να κινηθούμε προς την κατεύθυνση a με τη μέγιστη κλίση. Η κλίση όταν κινούμαστε προς το $\underline{a}$ είναι

$\nabla f(\underline{x}_0)' \underline{a} = \|\nabla f(\underline{x}_0)\| \cdot \overset{=1}{\|\underline{a}\|} \cdot \cos\theta$.

Η κατεύθυνση που μεγιστοποιεί την κλίση είναι η κατεύθυνση της βαθμίδας $\nabla f(\underline{x}_0)$.
Άρα, θα κινηθούμε προς τη κατεύθυνση $\nabla f(\underline{x}_0)$, δηλαδή στα σημεία $\underline{x}_0 + t\nabla f(\underline{x}_0)$, $t \geq 0$.
Θα σταματήσουμε στο μέγιστο σημείο αυτής της ημιευθείας

$$g_{\nabla f(\underline{x}_0)}(t) = g_0(t) = f(\underline{x}_0 + t\nabla f(\underline{x}_0))$$

Έστω t^*: $g_0(t^*) = \max_{t \geq 0} g_0(t)$.

Θα σταματήσω στο $\underline{x}_1 = \underline{x}_0 + t^* \nabla f(\underline{x}_0)$.

$\longrightarrow$

Να σταματήσουμε στο x_1 ή να συνεχίσουμε?

Αν $\|\nabla f(\underline{x}_1)\| < \varepsilon$, η πιο μεγάλη κλίση δεν είναι αρκετά απότομη ώστε να συνεχίσουμε.

Αν $\|\nabla f(\underline{x}_1)\| \geq \varepsilon$, συνεχίζουμε.

Αλγόριθμος Βαθμίδας

Εκκίνηση: Επιλέγουμε σημείο εκκίνησης $\underline{x}_0$ και μια επιθυμητή ακρίβεια ε.

Επανάληψη: Έστω ότι έχω υπολογίσει το $\underline{x}_n$

1. Ορίζουμε $g_n(t) = f(\underline{x}_n + t\nabla f(\underline{x}_n))$, $t \geq 0$.
2. Βρίσκουμε $t^*: g_n(t^*) = \max\limits_{t \geq 0} g_n(t)$
3. Θέτουμε $\underline{x}_{n+1} = \underline{x}_n + t^* \nabla f(\underline{x}_n)$
4. $\|\nabla f(\underline{x}_{n+1})\| < \varepsilon$, σταματάμε.
   $\|\nabla f(\underline{x}_{n+1})\| \geq \varepsilon$, συνεχίζουμε.

Παράδειγμα

$f(x_1, x_2) = 4x_1 + 6x_2 - 2x_1^2 - 2x_1x_2 - 2x_2^2$, $x_1, x_2 \in \mathbb{R}$

ασκήσεις
1| Να υπολογιστούν τα $\nabla f(\underline{x})$, $Hf(\underline{x})$ και ν.δ.ο. f κοίλη.
2| Να λυθεί το $\max\limits_{\underline{x} \in \mathbb{R}^2} f(\underline{x})$. $\left(\text{λύση } \underline{x}^* = \begin{pmatrix} \frac{1}{3} \\ \frac{4}{3} \end{pmatrix}\right)$ (αναλυτικά)

3| Να λυθεί το $\max\limits_{\underline{x} \in \mathbb{R}^2} f(\underline{x})$ προσεγγιστικά με τον αλγόριθμο βαθμίδας, ξεκινώντας από το $\underline{x}_0 = (1,1)$ και $\varepsilon = 0.3$.

Λύση $\nabla f(x_1, x_2) = (4 - 4x_1 - 2x_2,\ 6 - 2x_1 - 4x_2)$.

$\underline{x}_0 = (1,1)$

$\nabla f(\underline{x}_0) = (-2, 0)$, $\|\nabla f(\underline{x}_0)\| = \sqrt{(-2)^2 + 0^2} = 2 > \varepsilon$, συνεχίζουμε.

Επανάληψη 1

1. Υπολογίζαμε $g_0(t)$

$$g_0(t) = f(\underline{x}_0 + t\nabla f(\underline{x}_0)) = f((1,1) + t(-2,0)) = f(1-2t, 1) =$$
$$= 4(1-2t) + 6\cdot 1 - 2(1-2t)^2 - 2(1-2t)\cdot 1 - 2\cdot 1^2 =$$

$= -8t^2 + 4t + 4 \;, \; t \geq 0$

2. Βρίσκουμε το t^* : $g_o(t^*) = \max\limits_{t\geq 0} g_o(t)$

$g_o'(t) = -16t + 4$

$g_o'(t^*) = 0 \Rightarrow -16t^* + 4 = 0 \Rightarrow t^* = \frac{1}{4}$

3. Βρίσκουμε $\underline{x}_1$

$\underline{x}_1 = \underline{x}_o + t^* \nabla f(\underline{x}_o) = (1,1) + \frac{1}{4}(-2,0) = \left(\frac{1}{2}, 1\right)$

4. Συνεχίζουμε;

$\nabla f(\underline{x}_1) = \nabla f\left(\frac{1}{2}, 1\right) = (0,1)$

$\|\nabla f(\underline{x}_1)\| = \sqrt{1^2 + 0^2} = 1 > \varepsilon \Rightarrow$ Συνεχίζουμε

Επανάληψη 2 ... (άσκηση)