ΒΕΛΤΙΣΤΟΠΟΙΗΣΗ ΜΕ ΠΕΡΙΟΡΙΣΜΟΥΣ

Το πρόβλημα: $\min F(\vec{x})$, $F: \mathbb{R}^N \to \mathbb{R}$

$$\models \quad c_i(\vec{x}) \le 0 \quad i = 1, \dots, M_2 \text{ ή } i \in I \quad (I = \text{Inequality})$$

$$c_i(\vec{x}) = 0 \quad i = 1, \dots, M_1 \text{ ή } i \in E \quad (E = \text{Equality})$$

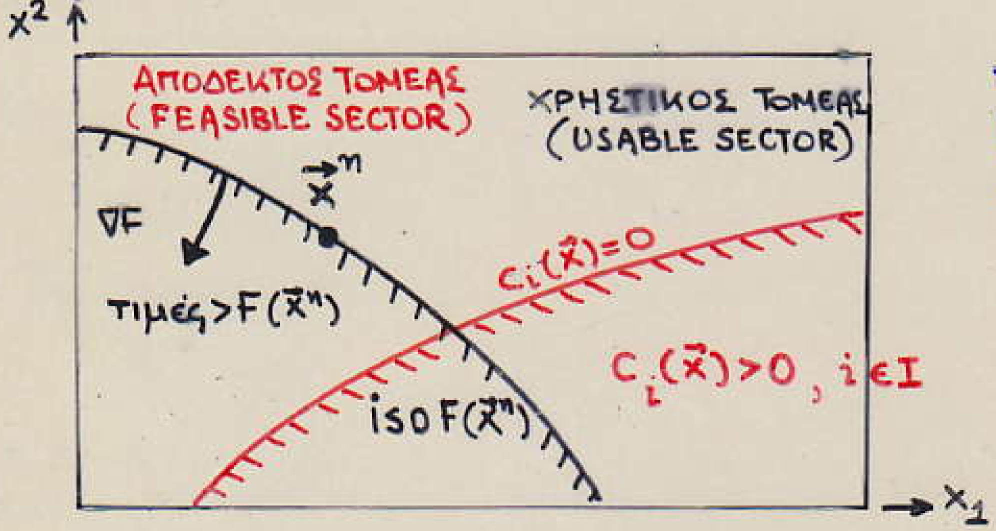
Χώρος Αποδεκτών Λύσεων:

$$\Omega = \{ \vec{x} \mid c_i(\vec{x}) \le 0 , i \in I \; ; \; c_i(\vec{x}) = 0 , i \in E \}$$

Το πρόβλημα: $\min\limits_{\vec{x} \in \Omega} F(\vec{x})$

Ενεργός Περιορισμός Ανισότητας στο $\vec{x}$:

για $\vec{x} \in \Omega$, όταν $c_i(\vec{x}) = 0$



Παράδειγμα με έναν περιορισμό ανισότητας –

Αναγκαία Συνθήκη Ελαχίστου με έναν Περιορισμό Ισότητας:

Αναγκαία συνθήκη ώστε $\vec{x}^*$ λύση του

$$\min_{\vec{x} \in \mathbb{R}^N} F(\vec{x}) \qquad \models \; c_1(\vec{x}) = 0$$

είναι να μην υπάρχει κατεύθυνση $\vec{p}$ όπου ταυτόχρονα να ισχύουν

$$\vec{p}^T \nabla c_1(\vec{x}^*) = 0$$

$$\vec{p}^T \nabla F(\vec{x}^*) < 0$$

ή, αναγκαία η παραλληλία : $\nabla F(\vec{x}^*) = \lambda_1^* \nabla c_1(\vec{x}^*)$

* <u>αλλοιώς</u>, κίνηση κατά :

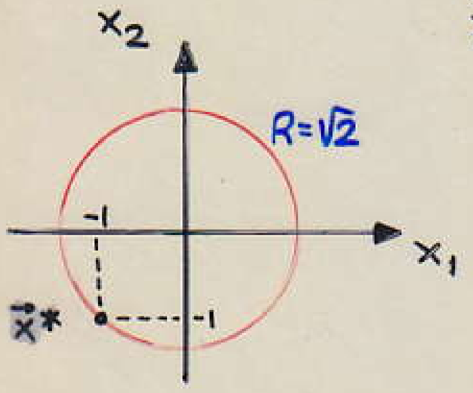
$$\vec{p} = - \left(I - \frac{\nabla c_1^T(\vec{x}^*) \, \nabla c_1(\vec{x}^*)}{\| \nabla c_1(\vec{x}^*) \|^2} \right) \nabla F(\vec{x}^*)$$

- <u>Διερεύνηση</u> (Πρώτης Τάξης): $\boxed{\vec{x} \in \Omega} \rightarrow \boxed{\vec{x}+\vec{p} \in \Omega}$

$$c_1(\vec{x}+\vec{p}) \approx c_1(\vec{x}) + \vec{p}^T \nabla c_1(\vec{x}) = \vec{p}^T \nabla c_1(\vec{x}) = 0$$

$$F(\vec{x}+\vec{p}) \approx F(\vec{x}) + \vec{p}^T \nabla F(\vec{x}) < F(\vec{x}) \Rightarrow \vec{p}^T \nabla F(\vec{x}) < 0$$

- <u>Παράδειγμα:</u>



$$\min_{\vec{x} \in \mathbb{R}^2} F(\vec{x}) \quad , \quad F(\vec{x}) = x_1 + x_2$$

$$\models \; c_1(\vec{x}) = x_1^2 + x_2^2 - 2 = 0$$

$$\underbrace{\vec{x}^* = (-1,-1)}: \left\{ \begin{array}{l} \nabla F(\vec{x}) = (1,1) \\ \nabla c_1(\vec{x}) = (2x_1, 2x_2) \end{array} \right\}$$

$$\nabla F(\vec{x}^*) = \lambda_1^* \nabla c_1(\vec{x}^*)$$

$$\left. \begin{array}{l} 1 = -2x_1^* \\ 1 = -2x_1^* \end{array} \right\} \quad \lambda_1^* = -\frac{1}{2}$$

Χρήση Πολλαπλασιαστών Lagrange:

$$L(\vec{x}, \lambda_1) = F(\vec{x}) - \lambda_1 c_1(\vec{x})$$

Στάσιμα σημεία: $\nabla L(\vec{x}, \lambda_1) = \underbrace{\nabla F(\vec{x}) - \lambda_1 \nabla c_1(\vec{x})}_{\text{γνωστή ήδη!}} = 0$

<u>Προσοχή</u>: Είναι αναγκαία, όχι ικανή (αφού π.χ. ισχύει και στο $(+1,+1)$)

<u>Προσοχή</u>: Δεν αντλούμε πληροφορία από $\text{sign}(\lambda_1)$!!!

Αναγκαία Συνθήκη Ελαχίστου με έναν περιορισμό Ανισότητας

Αναγκαία συνθήκη ώστε το $\vec{x}^*$ λύση του

$$\min_{\vec{x} \in \mathbb{R}^N} F(\vec{x}) \quad \models \; c_1(\vec{x}) \leq 0$$

είναι η $\nabla F(\vec{x}^*) = \lambda_1^* \nabla c_1(\vec{x}^*)$ με $\lambda_1^* \leq 0$

και η $\lambda_1^* c_1(\vec{x}^*) = 0$

<u>Παράδειγμα:</u>

$$\min_{\vec{x} \in \mathbb{R}^2} F(\vec{x}) \quad , \quad F(\vec{x}) = x_1 + x_2$$

$$\models \; c_1(\vec{x}) = x_1^2 + x_2^2 - 2 \leq 0$$

$\vec{x}^* = (-1,-1)$

ΣΥΝΘΗΚΕΣ KARUSH - KUHN - TUCKER (KKT)

Πρώτης τάξης <u>αναγκαίες</u> συνθήκες ώστε το $\vec{x}^*$ νάναι <u>τοπικά</u> βέλτιστη λύση του $\underline{\min F(\vec{x})}$

$$\models \quad c_i(\vec{x}) \leq 0 \quad , i \in I$$
$$c_i(\vec{x}) = 0 \quad , i \in E$$

για το οποίο ορίζουμε τη συνάρτηση LAGRANGE:

$$L(\vec{x},\vec{\lambda}) = F(\vec{x}) - \sum_{i \in E \cup I} \lambda_i c_i(\vec{x})$$

είναι να υπάρχει διάνυσμα συντελεστών $\vec{\lambda}^*$ ώστε

$$\nabla L(\vec{x}^*,\vec{\lambda}^*) = 0 \;\# \qquad (KKT-1)$$
$$c_i(\vec{x}^*) = 0 \quad \forall i \in E \qquad (KKT-2)$$
$$c_i(\vec{x}^*) \leq 0 \quad \forall i \in I \qquad (KKT-3)$$
$$\lambda_i^* \leq 0 \quad \forall i \in I \qquad (KKT-4)$$
$$\lambda_i c_i(\vec{x}^*) = 0 \quad \forall i \in E \cup I \qquad (KKT-5)$$

$$\# \quad \nabla F(\vec{x}^*) - \sum_{i \in E \cup I} \lambda_i \nabla c_i(\vec{x}) = 0$$

<u>Άσκηση:</u>

$$\left.\begin{aligned} &\min F(x_1,x_2) = \left(x_1 - \tfrac{3}{2}\right)^2 + \left(x_2 - \tfrac{1}{8}\right)^2 \\ \models\; &c_1(x_1,x_2) = x_1 + x_2 - 1 \leq 0 \\ &c_2(x_1,x_2) = x_1 - x_2 - 1 \leq 0 \\ &c_3(x_1,x_2) = -x_1 + x_2 - 1 \leq 0 \\ &c_4(x_1,x_2) = -x_1 - x_2 - 1 \leq 0 \end{aligned}\right\} \quad \text{λύση } \vec{x}^* = (1,0)$$

$$\left.\begin{aligned} \nabla F(\vec{x}^*) &= (-1, -\tfrac{1}{2})^T \\ \nabla c_1(\vec{x}^*) &= (1,1)^T \\ \nabla c_2(\vec{x}^*) &= (1,-1)^T \end{aligned}\right\} \Rightarrow \left\{\begin{aligned} -1 - \lambda_1 - \lambda_2 &= 0 \\ -\tfrac{1}{2} - \lambda_1 + \lambda_2 &= 0 \end{aligned}\right\} \Rightarrow \left\{\begin{aligned} \lambda_1 &= -3/4 \\ \lambda_2 &= -1/4 \end{aligned}\right.$$

$$(\lambda_1,\lambda_2,\lambda_3,\lambda_4)^* = \left(\underbrace{-\tfrac{3}{4}, -\tfrac{1}{4}}, \emptyset, \emptyset\right)$$

Ενεργοί
$c_1(\vec{x}^*) = 0$, $c_2(\vec{x}^*) = 0$

$c_3(\vec{x}^*) = -2 < 0$ $\quad$ $c_4(\vec{x}^*) = -2 < 0$

ΕΛΑΧΙΣΤΟΠΟΙΗΣΗ ΜΕ ΧΡΗΣΗ ΣΥΝ|ΣΤΩΝ ΠΟΙΝΗΣ

Ψευδοαντικειμενική : $\Phi(\vec{x},\omega_p) = F(\vec{x}) + \omega_p P(\vec{x})$

SUMT = Sequential Unconstrained Minimization Technique

• ΜΕΘΟΔΟΣ ΕΞΩΤΕΡΙΚΗΣ ΠΟΙΝΗΣ (exterior penalty)

$$P(\vec{x}) = \sum_{i\in I}\left[\max\left(0, c_i(\vec{x})\right)\right]^2 + \sum_{i\in E}\left[c_i(\vec{x})\right]^2$$

δηλ. $P(\vec{x}) = 0 \quad \forall \vec{x} \in \Omega$ (αποδεκτή)

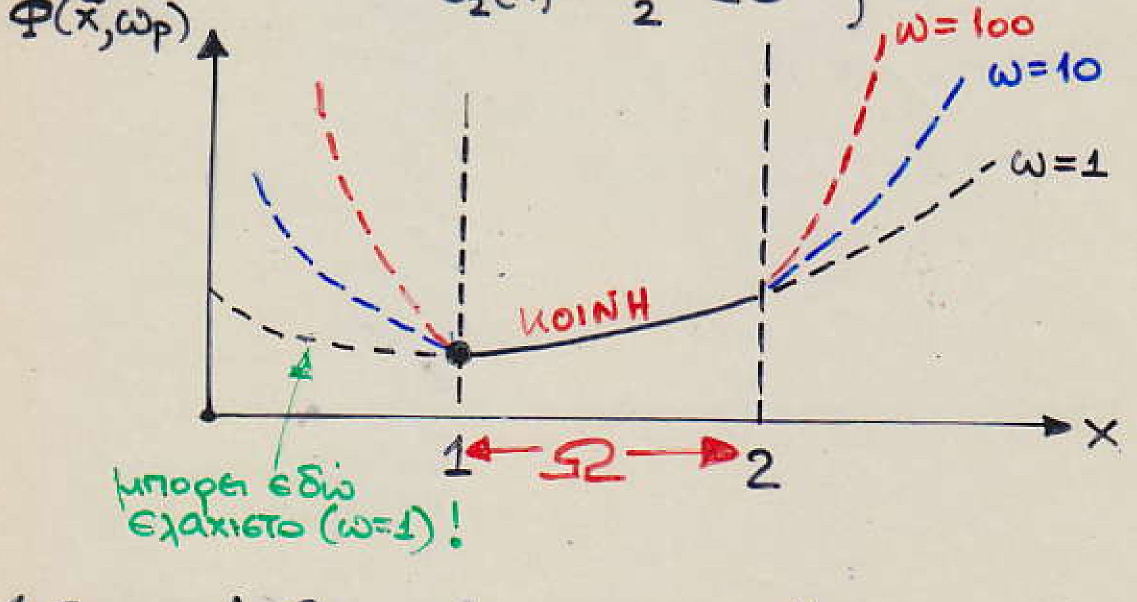
ω_p:

- $\omega_p \ll$: εύκολη επίλυση – ενδεχόμενη μη-ικανοποίηση των περιορισμών
- $\omega_p \gg$: ενδεχόμενες δυσκολίες σύγκλισης – σιγουριά μάλλον για τους περιορισμούς

Πρακτικά : αρχικά $\omega_p \ll$, $\forall$ κύκλο $\omega_p \leftarrow \omega_p * 3$!

Παράδειγμα : $\min\limits_{\vec{x}\in R} F(\vec{x}) = \dfrac{(x+2)^2}{20}$

$$\left.\begin{aligned} \models\ c_1(x) &= \frac{1-x}{2} \le 0 \\ c_2(x) &= \frac{x-2}{2} \le 0 \end{aligned}\right\} \quad \Omega = [1,2] \quad \vec{x}^* = 1$$



Κίνδυνος: Αν διακοπεί νωρίτερα να δώσει $1-\epsilon$, $\epsilon \lll$, δηλ. μή-αποδεκτή αλλά κοντά στην πραγμ. βέλτιστη

Θεραπεία: Μέθοδος εσωτερικής ποινής !

♦ ΜΕΘΟΔΟΣ ΕΣΩΤΕΡΙΚΗΣ ΠΟΙΝΗΣ (interior penalty)

Τελική λύση ← Αλληλουχία αποδεκτών λύσεων

$$P(\vec{x}) = \sum_{i\in I} \frac{-1}{c_i(\vec{x})} \quad (\text{ανισοτητα})$$

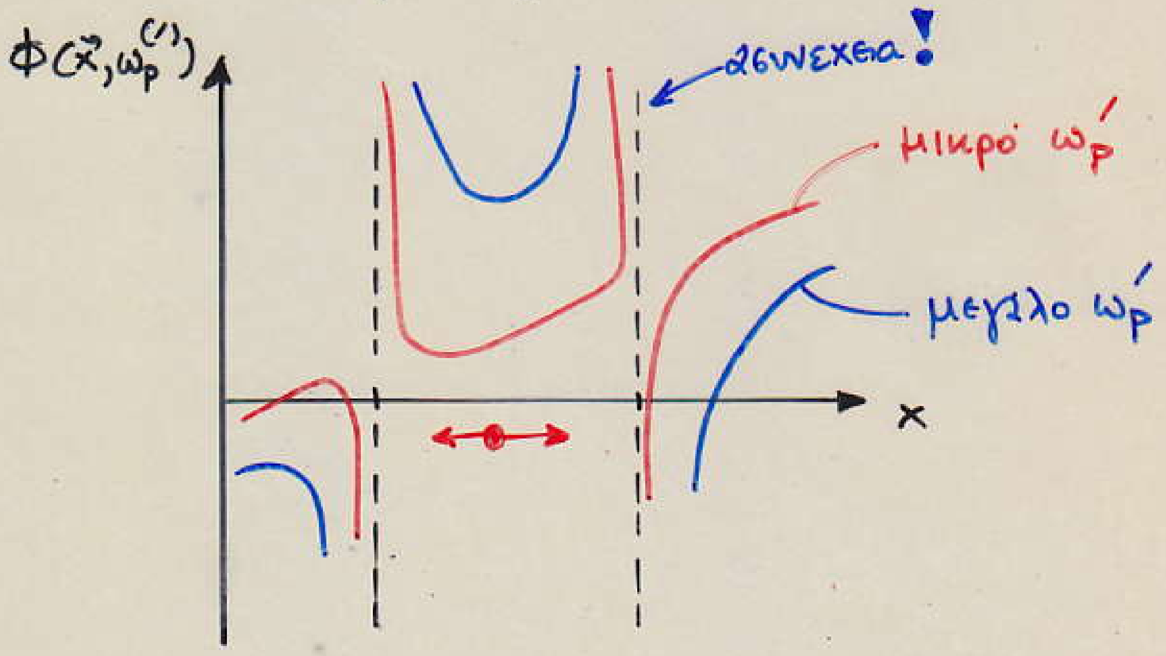
$$\Phi(\vec{x},\omega_p^{(\prime)}) = F(\vec{x}) + \omega_p' P(\vec{x}) + \omega_p \sum_{i\in E} [c_i(\vec{x})]^2$$

$$= F(\vec{x}) + \omega_p' \underbrace{\sum_{i\in I}\frac{-1}{c_i(\vec{x})}} + \omega_p \sum_{i\in E}[c_i(\vec{x})]^2$$

$$P(\vec{x}) = \sum_{i\in I}\left[-\ln(-c_i(\vec{x}))\right]$$, εναλλακτικά

Concept: $\left(\underset{i\in I}{c_i(\vec{x})\to 0}\right) \to \left[P(\vec{x})\to\infty\right]$!!

ω_p' → Αρχικά: μεγάλη
ω_p' → Διαρκώς: μειώνεται



ΜΕΘΟΔΟΣ ΔΙΕΥΡΥΜΕΝΗΣ ΕΣΩΤΕΡΙΚΗΣ ΠΟΙΝΗΣ

- Γραμμική διευρυμένη εσωτερική ποινή:

$$P(\vec{x}) = \sum_{i \in I} \tilde{c}(\vec{x}) \quad \text{όπου}$$

$$\tilde{c}(\vec{x}) = -\frac{1}{c_i(\vec{x})} \quad , \text{ αν } c_i(\vec{x}) \le \varepsilon$$

$$\tilde{c}(\vec{x}) = -\frac{2\varepsilon - c_i(\vec{x})}{\varepsilon^2} \quad , \text{ αν } c_i(\vec{x}) > \varepsilon \qquad \text{όπου } |\varepsilon| \ll \ , \ \varepsilon < 0 \ !$$

με συνέχεια τιμής & πρώτης παραγώγου στο $c_1(\vec{x}) = \varepsilon$!

Ασυνέχεια δεύτερης παραγώγου

Συνιστάται: $\varepsilon = -C\,(\omega_p')^{\alpha}$, $\frac{1}{2} \le \alpha \le \frac{1}{3}$, $C = \text{const}$

- Μ. Τετραγωνικής διευρυμένης εσωτερικής ποινής:

$$\tilde{c}(\vec{x}) = -\frac{1}{c_i(\vec{x})} \quad , \text{ αν } c_i(\vec{x}) \le \varepsilon$$

$$\tilde{c}(\vec{x}) = -\frac{1}{\varepsilon}\left[\left(\frac{c_i(\vec{x})}{\varepsilon}\right)^2 - 3\left(\frac{c_i(\vec{x})}{\varepsilon}\right) + 3\right] \quad , \text{ αν } c_i(\vec{x}) > \varepsilon$$

με συνέχεια δεύτερης παραγώγου !

ΕΠΑΥΞΗΜΕΝΕΣ ΜΕΘΟΔΟΙ ΜΕ ΧΡΗΣΗ ΠΟΛ/ΣΤΩΝ LAGRANGE

AUGMENTED LAGRANGE MULTIPLIER (ALM) METHODS

(1) Με περιορισμούς ισότητας μόνο:

$$\min_{\vec{x}\in R^N} F(\vec{x}) \quad \models \; c_i(\vec{x})=0 \;, \; i\in E$$

▶ Τι θα γίνονταν αν υπήρχαν N περιορισμοί ισότητας?

βέλτιστη λύση: στάσιμο σημείο της

$$L(\vec{x},\vec{\lambda}) = F(\vec{x}) - \sum_{i\in E} \lambda_i c_i(\vec{x}) \qquad \text{(KKT)}$$

ή (εξωτ. ποινή) επαυξημένη συνάρτηση Lagrange:

$$\Phi^*(\vec{x},\vec{\lambda},\omega_p) = F(\vec{x}) + \sum_{i\in E}\left(-\lambda_i c_i(\vec{x}) + \omega_p\left(c_i(\vec{x})\right)^2\right)$$

ΒΗΜΑ 1: $\vec{x}^o = \ldots$, $\vec{\lambda}^o = \ldots$, $\omega_p = \ldots \ll$
$\gamma = \ldots$ (πολλαπλασιαστής του ω_p) , $\omega_p^{max} = \ldots$

ΒΗΜΑ 2: Ελαχιστοποίηση της $\Phi^*(\vec{x},\vec{\lambda},\omega_p)$.
(Πρόβλημα χωρίς περιορισμούς)

ΒΗΜΑ 3: Σύγκλιση? (Κριτήρια: μικρή δF, ικανοποίηση περιορισμών εντός ανοχής) STOP or CONTINUE

ΒΗΜΑ 4:

$$\lambda_i \leftarrow \lambda_i - 2\omega_p c_i(\vec{x}) \quad , \; i\in E$$

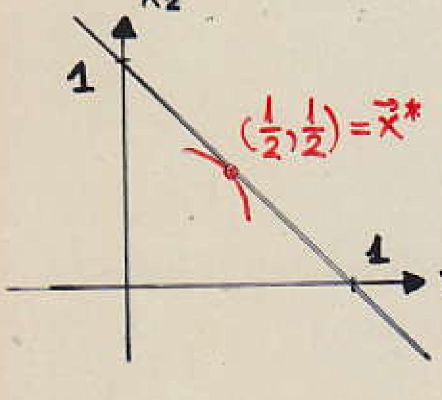
$$\omega_p \leftarrow \min(\gamma\omega_p, \omega_p^{max})$$

Go το ΒΗΜΑ 2

ΑΣΚΗΣΗ:

$$\min_{\vec{x}\in R^2} F(\vec{x}) \quad , \quad F(\vec{x}) = x_1^2 + x_2^2$$

$$\models c_1(\vec{x}) = x_1 + x_2 - 1 = 0$$



$$\Phi^*(\vec{x},\lambda_1,\omega_p) = x_1^2 + x_2^2 - \lambda_1(x_1+x_2-1) + \omega_p(x_1+x_2-1)^2$$

ΛΥΣΗ:

$$\frac{\partial \Phi^*}{\partial x_1}=0 \Rightarrow 2x_1-\lambda_1+2\omega_p(x_1+x_2-1)=0$$

$$\frac{\partial \Phi^*}{\partial x_2}=0 \Rightarrow 2x_2-\lambda_1+2\omega_p(x_1+x_2-1)=0$$

$$\Rightarrow \boxed{x_1=x_2=\frac{2\omega_p+\lambda_1}{2+4\omega_p}}$$

άρα για $\vec{x}^*=(\frac{1}{2},\frac{1}{2}) \Rightarrow \lambda_1^*=1$ (ανεξάρτητο του ω_p !)

▲ ΕΚΤΕΛΕΣΗ ΕΠΑΝΑΛΗΨΕΩΝ με $\omega_p=1=$σταθερό (δηλ. $\gamma=1$):

(Κ1) Έστω $\lambda_1=0 \Rightarrow x_1=x_2=\frac{1}{3}$, $F(\vec{x})=\frac{2}{9}$

$c_1(\vec{x})=-\frac{1}{3}\neq 0 \Rightarrow \lambda_1=0-2(-\frac{1}{3})=\frac{2}{3}$

(Κ2) $\lambda_1=\frac{2}{3} \Rightarrow x_1=x_2=\frac{4}{9}$, $F(\vec{x})=\frac{32}{81}$

$c_1(\vec{x})=-\frac{1}{9}\neq 0 \Rightarrow \lambda_1=\frac{2}{3}-2(-\frac{1}{9})=\frac{8}{9}$

(Κ3) $\lambda_1=\frac{8}{9} \Rightarrow x_1=x_2=\frac{13}{27}$, $F(\vec{x})=\frac{338}{729}$

$c_1(\vec{x})=-\frac{1}{27}\neq 0 \Rightarrow \lambda_1=\frac{8}{9}-2(-\frac{1}{27})=\frac{26}{27}$

(Κ4) $\lambda_1=\frac{26}{27} \Rightarrow x_1=x_2=\frac{40}{81}$, $F(\vec{x})=\frac{3200}{6561}$

$c_1(\vec{x})=-\frac{1}{81}$ κ.ο.κ.

▲ Έστω ότι γνωρίζαμε το βέλτιστο $\lambda_1^*=1$:

$$x_1=x_2=\frac{2\omega_p+1}{2+4\omega_p}=1/2$$ (χωρίς ΚΥΚΛΟΥΣ !!!)

(2) Με περιορισμούς ανισότητας :

$$\min_{\vec{x}\in R^N} F(\vec{x}) \quad \models \quad c_i(\vec{x})\leq 0\,,\ i\in I$$

Μετατροπή σε ισοδύναμο πρόβλημα με περιορισμούς ισότητας :

$$c_i(\vec{x})+\underbrace{z_i^2}_{\oplus}=0 \quad ,\ i\in I$$

$$\Phi^*(\vec{x},\vec{\lambda},\vec{z},\omega_p)=F(\vec{x})+\sum_{i\in I}\Big(-\lambda_i\big[c_i(\vec{x})+z_i^2\big]+\omega_p\big[c_i(\vec{x})+z_i^2\big]^2\Big)$$

ασυνεχή 2η παρ. στο $c_i=\frac{\lambda_i}{2\omega_p}$

ή εναλλακτικά :

$$\Phi^*(\vec{x},\vec{\lambda},\omega_p)=F(\vec{x})+\sum_{i\in I}\left(-\lambda_i\psi_i+\omega_p\psi_i^2\right)$$

όπου $\psi_i=\max\left[c_i(\vec{x}),\frac{\lambda_i}{2\omega_p}\right]$

Τρόπος ανανέωσης λ_i : $\quad \lambda_i\leftarrow\lambda_i-2\omega_p\psi_i(\vec{x})\,,\ i\in I$

(3) Γενίκευση:

$$\min_{\vec{x}\in R^N} F(\vec{x}) \quad \models c_i(\vec{x})=0\,,\ i\in E \qquad c_i(\vec{x})\leq 0\,,\ i\in I$$

$$\Phi^*(\vec{x},\vec{\lambda},\omega_p)=F(\vec{x})+\sum_{i\in E}\left(-\lambda_i c_i(\vec{x})+\omega_p c_i^2(\vec{x})\right)+\sum_{i\in I}\left(-\lambda_i\psi_i+\omega_p\psi_i^2\right)$$

όπου $\psi_i=\max\left[c_i(\vec{x}),\frac{\lambda_i}{2\omega_p}\right]$

Ανανεώσεις : $\quad \lambda_i\leftarrow\lambda_i-2\omega_p(c_i(\vec{x})) \quad i\in E$

$\qquad \lambda_i\leftarrow\lambda_i-2\omega_p\max\left[c_i(\vec{x}),\frac{\lambda_i}{2\omega_p}\right] \quad i\in I$

ΣΧΟΛΙΑ/ΣΥΝΤΑΓΕΣ:

- ή ALM δεν εξαρτάται ιδιαίτερα από ω_p !
- Το σημείο ευκίνησης μπορεί να είναι αποδεκτό ή μή !
- Στη βέλτιστη λύση, ή τιμή του λ_i δείχνει ποιός περιορισμός ανισότητας είναι ενεργός !