Διαλέξεις “Βελτιστοποίησης”
Αναγκαίες και Ικανές Συνθήκες

Ορισμοί:

- Έστω $\Omega \subset \mathbb{R}^n$ ανοικτό. Υποθέτουμε ότι η $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ έχει μερικές παραγωγούς πρώτης τάξης σε σημείο $x \in \Omega$, που συμβολίζεται με $\frac{\partial f}{\partial x_i}(x), i = 1,...,n$. Συμβολίζουμε την κλίση (βαθμίδα) - gradient το διάνυσμα $\nabla f(x) = \left[\frac{\partial f}{\partial x_1}(x), \ldots, \frac{\partial f}{\partial x_n}(x)\right]^T$. Αν η $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ έχει συνεχείς μερικές παραγώγους στο Ω μέχρι και τάξης m τότε συμβολίζουμε με $f \in C^m(\Omega)$.

- Η συνάρτηση $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ λέγεται παραγωγίσιμη στο $x \in \Omega$, αν για $x + h \in \Omega$, ισχύει $f(x+h) - f(x) = v(x)^T h + e(h)\|h\|$, με $v(x) \in \mathbb{R}^n$ όπου $e(h) \to 0$ όταν $\|h\| \to 0$.

- Υπενθυμίζουμε πως αν οι μερικές παράγωγοι $\frac{\partial f}{\partial x_i}, i = 1,...,n,$ υπάρχουν σε μία ανοικτή περιοχή του x, δηλαδή υπάρχουν $\forall y \in B(x, \delta)$ (για κάποιο $\delta > 0$), και είναι συνεχής στο x, τότε $v(x) = \nabla f(x)$.

- Αν η f είναι παραγωγίσιμη σε κάθε σημείο $S \subset \Omega$, τότε η f λέγεται παραγωγίσιμη στο $S \subset \Omega$.

- Έστω $V \subset \Omega$, με V ανοικτό. Αν η $f \in C^1(V)$ τότε η f είναι παραγωγίσιμη στο V.

Ορισμοί ΙΙ

- Η ύπαρξη μερικών παραγώγων της f <u>συνεχών στο σημείο x μόνο</u> δεν εξασφαλίζει την παραγωγισιμότητα της f στο x. Αν όμως $f \in C^1(V)$, όπου V είναι μία ανοικτή περιοχή (γειτονιά) του x, τότε η f είναι παραγωγίσιμη στο V.

- <u>Ορισμός</u> (κατά κατεύθυνση παραγώγων): Έστω $f : S \subset \mathbb{R}^n \to \mathbb{R}$, $x \in S, h \in \mathbb{R}^n$. Αν $x + ah \in S, a \neq 0,$ για κάθε a αρκετά μικρό, και το όριο $\delta f(x, h) = \lim_{a \to 0} \frac{f(x + ah) - f(x)}{a} < \infty$ τότε το $\delta f(x, h)$ καλείται παράγωγος της f στο x κατά κατεύθυνση h.

- <u>Ορισμός</u> (κατά θετική κατεύθυνση παραγώγων): Έστω $f : S \subset \mathbb{R}^n \to \mathbb{R}$, $x \in S, h \in \mathbb{R}^n$. Αν $x + ah \in S, a > 0,$ για κάθε a αρκετά μικρό, και το όριο $\delta_+ f(x, h) = \lim_{a \to 0^+} \frac{f(x + ah) - f(x)}{a} < \infty$ τότε το $\delta_+ f(x, h)$ καλείται παράγωγος της f στο x κατά θετική κατεύθυνση h.

- Αν η $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$, παραγωγίσιμη, με $\Omega \subset \mathbb{R}^n$ ανοικτό τότε ισχύει: $\delta f(x, h) = \delta_+ f(x, h) = \nabla f(x)^T h$, για κάθε $h \in \mathbb{R}^n$.

Ερμηνεία (γεωμετρική) της Κλίσης

- $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ παραγωγίσιμη στο $\Omega \subset \mathbb{R}^n$ ανοικτό. Έστω, $h \in \mathbb{R}^n, \|h\|_2 = 1$. Τότε από ανισότητα Cauchy-Schwarz
$$\begin{aligned} |\delta f(x,h)| = |\nabla f(x)^T h| &\leq \|\nabla f(x)\|_2 \|h\|_2 \\ &= \|\nabla f(x)\|_2 \end{aligned}$$

- Παρατηρούμε, ότι για $u := \dfrac{\nabla f(x)}{\|\nabla f(x)\|_2}$, $\|u\|_2 = 1$ και εύκολα υπολογίζουμε, $\delta f(x,u) = \|\nabla f(x)\|_2$. Επομένως, καταλήγουμε πως $\delta f(x,h) \leq \delta f(x,u), \forall h \in \mathbb{R}^n, \|h\|_2 = 1$

- Δηλαδή, η κατεύθυνση της κλίσης είναι αυτή που εξασφαλίζει το μεγαλύτερο ρυθμό αύξησης της $\delta f(x,h)$ κοντά στο x.

- Άρα, σε έναν επαναληπτικό αλγόριθμο, που προσπαθεί να υπολογίσει σημεία ελαχίστου, η κατεύθυνση $-\nabla f$ αποτελεί βασική επιλογή.

- Ορισμός Εσσιανού πίνακα: Υποθέτουμε ότι η $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ έχει παραγωγούς δεύτερης τάξης $\frac{\partial^2 f}{\partial x_i \partial x_j}(x) \quad i, j = 1,\ldots,n$, ορίζουμε τον Εσσιανό πίνακα της f στο x, ως

$$H(x) := f_{xx}(x) := \begin{pmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1}(x) & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \cdots & \cdots & \cdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(x) & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_n}(x) \end{pmatrix}$$

- Υπενθυμίζουμε πως αν η f είναι C^2 τότε ο $H(x)$ είναι συμμετρικός.

- *<u>Θεώρημα</u>* (*<u>Μέσης Τιμής</u>*): $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$. Αν $x + ah \in \Omega$, για κάθε $a \in [0,1]$, η f είναι συνεχής σε κάθε σημείο $x + ah \in \Omega$ με $a \in [0,1]$ και η f είναι παραγωγίσιμη σε κάθε σημείο $x + ah \in \Omega$, με $a \in (0,1)$, τότε υπάρχει $\mu \in (0,1)$, τέτοιο ώστε $f(x + h) = f(x) + \nabla f(x + \mu h)^T h$.

- *<u>Θεώρημα</u>* *<u>(Taylor)</u>*: $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$, f είναι C^2, Αν $x + ah \in \Omega$, για κάθε $a \in [0,1]$, τότε υπάρχει $\mu \in (0,1)$ ώστε $f(x + h) = f(x) + \nabla f(x)^T h + \frac{1}{2} h^T H(x + \mu h) h$. Αν $\Omega \subset \mathbb{R}^n$ ανοικτό, για $\|h\|$ αρκετά μικρή, $f(x + h) = f(x) + \nabla f(x)^T h + \frac{1}{2} h^T H(x) h + e(h)\|h\|^2$, όπου $\lim_{\|h\| \to 0} e(h) = 0$.

- **<u>Θεώρημα 5 (Αναγκαία Συνθήκη)</u>**: Έστω $f: S \subset \mathbb{R}^n \to \mathbb{R}$, $S \subset \mathbb{R}^n$ κυρτό, $\bar{x} \in S$. Έστω ότι η $\delta_+ f(\bar{x}, x - \bar{x})$ υπάρχει για κάθε $x \in S$. Αν το $\bar{x} \in S$ είναι σημείο τοπικού ή ολικού ελαχίστου, τότε ισχύει η συνθήκη $\delta_+ f(\bar{x}, x - \bar{x}) \geq 0, \quad \forall x \in S$. Αν επιπλέον $S \subset \Omega$, με $\Omega \subset \mathbb{R}^n$ ανοικτό και η $f: \Omega \subset \mathbb{R}^n \to \mathbb{R}$ είναι παραγωγίσιμη στο $\bar{x} \in S$ τότε η αναγκαία συνθήκη παίρνει την μορφή $\nabla f(\bar{x})^T (x - \bar{x}) \geq 0, \quad \forall x \in S$.

- *<u>Απόδειξη Θεωρήματος 5</u>*: Επειδή το $\bar{x} \in S$ είναι σημείο τοπικού ελαχίστου της f στο S, υπάρχει $\delta > 0$, τέτοιο ώστε $\forall y \in S, \|y - \bar{x}\| < \delta$, να ισχύει $f(\bar{x}) \leq f(y)$. Επειδή $x, \bar{x} \in S, a \in (0,1]$, και $S \subset \mathbb{R}^n$ κυρτό έχουμε πως $(1-a)\bar{x} + ax \in S$. Παρατηρούμε ότι για $a \in (0,1]$ αρκετά μικρό το στοιχείο $y := \bar{x} + a(x - \bar{x})$ ικανοποιεί την σχέση $\|\bar{x} + a(x - \bar{x}) - \bar{x}\| = a\|x - \bar{x}\| < \delta$, επομένως ισχύει $f(\bar{x}) \leq f(\bar{x} + a(x - \bar{x}))$, που σημαίνει ότι
$$\delta_+ f(\bar{x}, x - \bar{x}) = \lim_{a \to 0^+} \frac{f(\bar{x} + a(x - \bar{x})) - f(\bar{x})}{a} \geq 0 \quad \text{για κάθε} \quad x \in S.$$
Αν επιπλέον $S \subset \Omega$, με $\Omega \subset \mathbb{R}^n$ ανοικτό και η $f: \Omega \subset \mathbb{R}^n \to \mathbb{R}$ είναι παραγωγίσιμη στο $\bar{x} \in S$ τότε η αναγκαία συνθήκη παίρνει την μορφή $\delta_+ f(\bar{x}, x - \bar{x}) = \nabla f(\bar{x})^T (x - \bar{x}) \geq 0, \quad \forall x \in S$.

- **<u>Θεώρημα 6</u>** (Ικανή και Αναγκαία Συνθήκη): Έστω $f : S \subset \mathbb{R}^n \to \mathbb{R}$, $S \subset \mathbb{R}^n$ κυρτό, $\bar{x} \in S$. Έστω ότι η $\delta_+ f(\bar{x}, x - \bar{x})$ υπάρχει για κάθε $x \in S$. Αν η f είναι κυρτή στο $S \subset \mathbb{R}^n$, τότε η συνθήκη $\delta_+ f(\bar{x}, x - \bar{x}) \geq 0, \quad \forall x \in S$ είναι ικανή και αναγκαία για να είναι το $\bar{x} \in S$, σημείο ελαχίστου της f στο S.

- *<u>Απόδειξη</u>*: Επειδή το S είναι κυρτό, έχουμε ότι για $x \in S$, και $a \in (0,1]$, το στοιχείο, $(1-a)\bar{x} + ax = \bar{x} + a(x - \bar{x}) \in S$. Τότε, αφού f κυρτή,
$$\begin{aligned} f(\bar{x} + a(x - \bar{x})) &= f((1-a)\bar{x} + ax) \\ &\leq (1-a)f(\bar{x}) + af(x) \\ &\leq f(\bar{x}) + a\left(f(x) - f(\bar{x})\right) \end{aligned}$$
Επομένως, έχουμε λόγω του προηγούμενου θεωρήματος 5 έχουμε,
$\delta_+ f(\bar{x}, x - \bar{x}) = \lim_{a \to 0^+} \dfrac{f(\bar{x} + a(x - \bar{x})) - f(\bar{x})}{a} \geq 0$. Συνδυάζοντας τις δύο ανισότητες έχουμε πως $a(f(x) - f(\bar{x})) \geq 0$, που σημαίνει πως $f(\bar{x}) \leq f(x), \forall x \in S$.

- Σημείωση: Υπενθυμίζουμε πως Θεώρημα 3, έχουμε πως τα τοπικά σημεία ελαχίστου είναι (ολικά) σημεία ελαχίστου.

- Σημείωση: Αν το $S = R^n$, τότε η αναγκαία συνθήκη του Θεωρήματος 5, γίνεται $\nabla f(\bar{x}) = 0$, είναι η κλασική συνθήκη βελτιστοποίησης / βελτιστότητας / optimality condition, για προβλήματα χώρις περιορισμούς.

Ικανή Συνθήκη

- Ορισμός: Ένας πίνακας $A \in \mathbb{R}^{n\times n}$ λέγεται θετικός αν $x^T A x \geq 0, \quad \forall x \in \mathbb{R}^n$ και θετικά ορισμένος αν $x^T A x > 0, \quad \forall x \in \mathbb{R}^n, \quad x \neq 0$.

- **<u>Θεώρημα 7</u>**: Έστω $\Omega \subset \mathbb{R}^n$ ανοικτό, $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$, $\bar{x} \in \Omega$, και η f είναι C^2 σε μία περιοχή κοντά στο $\bar{x}$. Αν $\nabla f(\bar{x}) = 0$ και ο $H(\bar{x})$ είναι θετικά ορισμένος, τότε το $\bar{x}$ είναι σημείο τοπικού ελαχίστου της f στο Ω.

- <u>*Απόδειξη*</u>: Εφαρμόζοντας το θεώρημα Taylor, λαμβάνουμε

$$f(\bar{x}+h) = f(\bar{x}) + \nabla f(\bar{x})^T h + \frac{1}{2} h^T H(\bar{x}) h + e(h)\|h\|^2$$

$$= f(\bar{x}) + \frac{1}{2} h^T H(\bar{x}) h + e(h)\|h\|^2$$

όπου $0 \neq h \in \mathbb{R}^n, \|h\| \to 0$ και $\lim_{\|h\|\to 0} e(h) = 0$. Υπενθυμίζουμε πως επειδή $\Omega \subset \mathbb{R}^n$ ανοικτό και $0 \neq h \in \mathbb{R}^n, \|h\| \to 0$, έχουμε $x + h \in \Omega$. Θέτουμε $u = \frac{h}{\|h\|}$, επομένως, $\|u\| = 1$ και $f(\bar{x}+h) = f(\bar{x}) + \frac{1}{2}\left(u^T H(\bar{x}) u + e(h)\right)\|h\|^2$.

- *(Συνέχεια Απόδειξης)*: Παρατηρούμε ότι η συνάρτηση $u \in S := \{u \in \mathbb{R}^n \,|\, \|u\| = 1\} \to u^T H(\bar{x}) u \in \mathbb{R}_+$ είναι συνεχής και ορισμένη στο συμπαγές (ως κλειστό και φραγμένο) υποσύνολο του $\mathbb{R}^n$. Επομένως υπάρχει τουλάχιστον ένα σημείο ελαχίστου, $\bar{u} \in S$, και επειδή ο Εσσιανός πίνακας είναι θετικά ορισμένος έχουμε $u^T H(\bar{x}) u \geq \bar{u}^T H(\bar{x}) \bar{u} = \mu > 0$. Συνεπώς έχουμε,

$$f(\bar{x} + h) = f(\bar{x}) + \frac{1}{2}\left(u^T H(\bar{x}) u + e(h)\right) \|h\|^2$$

$$\geq f(\bar{x}) + \frac{1}{2}\left(\mu + e(h)\right) \|h\|^2 \geq f(\bar{x})$$

καθώς $0 \neq h \in \mathbb{R}^n, \|h\| \to 0$ και επειδή $\lim_{\|h\| \to 0} e(h) = 0$.

Δηλαδή, αποδείξαμε πως το $\bar{x}$ είναι σημείο *τοπικού ελαχίστου.*
Σημείωση: Το πρόσημο της παράστασης $(\mu + e(h)) > 0$ καθώς $0 \neq h \in \mathbb{R}^n, \|h\| \to 0$ και $\lim_{\|h\| \to 0} e(h) = 0$.

- *Άσκηση*: Να δείξετε ότι η συνάρτηση $f(x, y) = 5x^2 + y^2 + 4xy - 2x - 2y$ έχει μοναδικό σημείο ελαχίστου.

- *Λύση*: (υπόδειξη) Παρατηρούμε πως έχουμε ένα πρόβλημα βελτιστοποίησης χωρίς περιορισμούς. Επίσης η συνάρτηση είναι ομαλή C^∞ και $\Omega = \mathbb{R}^2$. Τα υποψήφια σημεία (τοπικού ή ολικού) ελαχίστου, υπολογίζονται από τη αναγκαία συνθήκη: $\nabla f(\bar{x}) = 0$. Εύκολα καταλήγουμε, πως το μοναδικό υποψήφιο σημείο είναι το $\bar{x}$=$[-1,3]^T$. Στη συνέχεια υπολογίζουμε τον H($\bar{x}$)= $\begin{bmatrix} 10 & 4 \\ 4 & 2 \end{bmatrix}$ (παρατηρούμε πως είναι ο Εσσιανός είναι σταθερός για κάθε $[x, y]^T$). Εύκολα αποδεικνύεται ο είναι θετικά ορισμένος, επομένως από Θεώρημα 7, έχουμε την ύπαρξη μοναδικού ελαχίστου.

Το θεώρημα Kuhn-Tucker-Lagrange

- Έστω, συναρτήσεις, f_i και σύνολο U (περιορισμών / δεσμεύσεων), τέτοια ώστε $f_i : \mathbb{R}^n \to \mathbb{R}, i = , \ldots, m, S \subset \mathbb{R}^n$ κυρτό

  $$U = \{x \in S \quad | \quad f_i(x) \leq 0, \quad i = 1,\ldots,l \quad | \quad f_i(x) = 0, \quad i = l+1,\ldots,m\}.$$

- *<u>Ορισμός Προβλήματος Βελτιστοποίησης (Π)</u>*: Να βρεθεί $\bar{x} \in U$ τέτοιο ώστε, $f_0(\bar{x}) = \min_{x \in U} f_0(x)$.

- Οι σχέσεις ανισοτήτων $f_i, \quad i = 1,\ldots,l$ καλούνται περιορισμοί - ανισότητας ενώ $f_i, \quad i = l+1,\ldots,m$ καλούνται περιορισμοί - ισότητας (δεσμεύσεις).

- Στόχος: Εύρεση αναγκαίο συνθηκών για το Πρόβλημα (Π).

- Θα ασχοληθούμε κυρίως με την περίπτωση $S = \mathbb{R}^n$.

- Όταν $l = 0$, δεν έχουμε κάποιο περιορισμό ισότητας.

- **Θεώρημα 8** (Πολλαπλασιαστών KTL): Έστω $\bar{x} \in U, S \subset \mathbb{R}^n$ κυρτό και συναρτήσεις $f_i, i = 0,\ldots,m$ παραγωγίσιμες στο $\bar{x}$, και συνεχείς σε μία ανοικτή περιοχή κοντά στο $\bar{x}$ (π.χ $f_i \in C^1, i = 0,\ldots,m$). Αν το $\bar{x} \in U$ είναι σημείο ελαχίστου (ή σημείο τοπικού ελαχίστου) της f_0 στο U, τότε υπάρχουν πολλαπλασιαστές $\lambda_i \in \mathbb{R}, i = 0,\ldots,m$, που δεν είναι όλοι μηδέν και $\lambda_i \geq 0, i = 0,\ldots,l$, ώστε:
$$\left[\sum_{i=0}^{m} \lambda_i \nabla f_i(\bar{x})\right]^T (x - \bar{x}) \geq 0, \quad \forall x \in S$$
$$\lambda_i f_i(\bar{x}) = 0, \quad i = 1,\ldots l$$

- Αν $S = \mathbb{R}^n$, τότε οι αναγκαίες συνθήκες μετατρέπονται σε:
$$\sum_{i=0}^{m} \lambda_i \nabla f_i(\bar{x}) = 0,$$
$$\lambda_i f_i(\bar{x}) = 0, \quad i = 1,\ldots l$$

- Οι πολλαπλασιαστές $\lambda_i \in \mathbb{R}, i = 0,\ldots,m$ δεν είναι μοναδικοί. Εκτός από εξαιρετικές περιπτώσεις, έχουμε $\lambda_0 \neq 0$ και επομένως θέτουμε $\lambda_0 = 1$.

- Περίπτωση $S = \mathbb{R}^n$: Αναγκαίες Συνθήκες:

$$\nabla f_0(\bar{x}) + \sum_{i=1}^{m} \lambda_i \nabla f_i(\bar{x}) = 0, \qquad (n \text{ εξισώσεις })$$

$$\lambda_i f_i(\bar{x}) = 0, \quad i = 1,\ldots,l \qquad (l \text{ εξισώσεις })$$

$$f_i(\bar{x}) = 0, \quad i = l+1,\ldots,m \qquad (m - l \text{ εξισώσεις})$$

- Παρατηρούμε πως το παραπάνω σύστημα είναι μη-γραμμικό, με n+m εξισώσεις και n+m αγνώστους (τις n συντεταγμένες του διανύσματος $\bar{x}$, και τους πολλαπλασιαστές $\lambda_\iota \in \mathbb{R}, i = 1,\ldots,m$ (προσοχή το $\lambda_0 = 1$.)

- Στις περισσότερες περιπτώσεις είναι εξαιρετικά δύσκολο (ίσως και αδύνατο) να λυθεί αναλυτικά το μη-γραμμικό σύστημα των αναγκαίων συνθηκών πρώτης τάξης. Γι' αυτό το λόγο θα κατασκευάσουμε αριθμητικές μεθόδους για την προσέγγιση των υποψήφιων σημείων ελαχίστου, δηλαδή προσέγγιση λύσεων του συστήματος (KTL).

- *Άσκηση*: Να υπολογιστεί (αν υπάρχει) με βάση το Θεώρημα (KTL) υποψήφιο σημείο (τοπικού / ολικού) ελαχίστου για το πρόβλημα, $min_{x \in U} f(x, y, z)$ όπου $f(x, y, z) = 2x + 3y + 4z$ και

  $$U = \{[x, y, z]^T \in \mathbb{R}^3 / x^2 + y^2 + z^2 \leq 1\}$$

  Είναι το σημείο μοναδικό;

- *Λύση*: Παρατηρούμε πως το υποσύνολο $U \subset \mathbb{R}^3$ είναι κλειστό και φραγμένο (άρα συμπαγές), ενώ η f είναι συνεχής. Επομένως από το Θεώρημα 1, υπάρχει τουλάχιστον ένα σημείο ελαχίστου. Επίσης το $U \subset \mathbb{R}^3$ είναι κυρτό, ενώ η συνάρτηση f είναι κυρτή. Επομένως αναζητούμε μόνο σημείο(α) ελαχίστου. Χρησιμοποιώντας το θεώρημα (KTL)

  $$f_0(x, y, z) = f(x, y, z), f_1(x, y, z) = x^2 + y^2 + z^2 - 1, l = m = 1,$$

  έχουμε,

  $$\nabla f_0(\bar{x}) + \lambda_1 \nabla f_1(\bar{x}) = 0,$$

  $$\lambda_1 f_1(\bar{x}) = 0$$

  $\lambda_1 \geq 0$ (αντιστοιχεί σε ανισότητα)

  ή ισοδύναμα

- (*συνέχεια λύσης*):

$$\begin{bmatrix} 2+2\lambda_1\bar{x} \\ 3+2\lambda_1\bar{y} \\ 4+2\lambda_1\bar{z} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\lambda_1(\bar{x}^2+\bar{y}^2+\bar{z}^2-1)=0, \quad \lambda_1 \geq 0$$

Περίπτωση 1: $\lambda_1 = 0$. Τότε έχουμε άτοπο.

Περίπτωση 2: $\lambda_1 > 0$. Τότε,

$$\begin{bmatrix} \bar{x} \\ \bar{y} \\ \bar{z} \end{bmatrix} = \begin{bmatrix} -\frac{2}{2\lambda_1} \\ -\frac{3}{2\lambda_1} \\ -\frac{4}{2\lambda_1} \end{bmatrix}$$

$$(\bar{x}^2+\bar{y}^2+\bar{z}^2-1)=0.$$

Αντικαθιστώντας τα x,y,z στην τελευταία σχέση, λαμβάνουμε

$$\frac{29}{4\lambda_1^2}=1, \quad \lambda_1 = \frac{\sqrt{29}}{2}, \bar{x} = -\frac{2}{\sqrt{29}}, \bar{y} = -\frac{3}{\sqrt{29}}, \bar{z} = -\frac{4}{\sqrt{29}}.$$

- Παρατηρούμε πως καταλήξαμε σε ένα μόνο υποψήφιο σημείο ελαχίστου (Μοναδικό !) το οποίο βρίσκεται πάνω στο σύνορο.

- *<u>Άσκηση</u>*:Να υπολογιστεί (αν υπάρχει) με βάση το Θεώρημα (KTL) υποψήφιο σημείο (τοπικού / ολικού) ελαχίστου για το πρόβλημα, $min_{x \in U} f(x, y, z)$ όπου $f(x, y, z) = 2(x-1)^2 + 2(y-1)^2 + z^2$ και
$U = \{[x, y, z]^T \in \mathbb{R}^3 / x^2 + y^2 \leq 1, z = 1\}$
Είναι το σημείο μοναδικό;

- *<u>Λύση</u>* (υπόδειξη): Παρατηρούμε πως το υποσύνολο $U \subset \mathbb{R}^3$ είναι κλειστό και φραγμένο (άρα συμπαγές), ενώ η f είναι συνεχής. Επομένως από το Θεώρημα 1, υπάρχει τουλάχιστον ένα σημείο ελαχίστου. Επίσης το $U \subset \mathbb{R}^3$ είναι κυρτό, ενώ η συνάρτηση f είναι κυρτή. Επομένως αναζητούμε μόνο σημείο(α) ελαχίστου Χρησιμοποιώντας το θεώρημα (KTL)
$f_0(x, y, z) = f(x, y, z), f_1(x, y, z) = x^2 + y^2 - 1, f_2(x, y, z) = z - 1, l = 1, m = 2.$
έχουμε,
$$\nabla f_0(\bar{x}) + \lambda_1 \nabla f_1(\bar{x}) + \lambda_2 \nabla f_2(\bar{x}) = 0,$$
$$\lambda_1 f_1(\bar{x}) = 0$$
$$f_2(\bar{x}) = 0$$
$\lambda_1 \geq 0$ αντιστοιχεί σε ανισότητα , $\lambda_2 \in \mathbb{R}$ αντιστοιχεί σε ισότητα

- (συνέχεια λύσης):

$$\begin{bmatrix} 4(\bar{x}-1)+2\lambda_1\bar{x} \\ 4(\bar{y}-1)+2\lambda_1\bar{y} \\ 2\bar{z}+\lambda_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\lambda_1(\bar{x}^2+\bar{y}^2-1)=0, \quad \lambda_1 \geq 0$$

$$\bar{z}=1.$$

Περίπτωση 1: Αν $\lambda_1 = 0$, τότε έχουμε $\bar{x} = 1, \bar{y} = 1, \bar{z} = 1, \lambda_2 = -2$. Παρατηρούμε όμως ότι αυτό το σημείο δεν ανήκει στο U. Άτοπο.

Περίπτωση 2: $\lambda_1 > 0$, τότε

$$\begin{bmatrix} \bar{x} \\ \bar{y} \\ \bar{z} \end{bmatrix} = \begin{bmatrix} \frac{2}{2+\lambda_1} \\ \frac{2}{2+\lambda_1} \\ -\frac{\lambda_2}{2} \end{bmatrix}$$

$$(\bar{x}^2+\bar{y}^2-1)=0, \bar{z}=1$$

Επομένως λαμβάνουμε, αντικαθιστώντας τις τιμές των $\bar{x}, \bar{y}$, στην τέταρτη εξίσωση,

$\dfrac{8}{(2+\lambda_1)^2} = 1, \lambda_1 = 2(\sqrt{2}-1)$ (η αρνητική τιμή του λ_1 απορρίπτεται)….

- *Άσκηση*:Να υπολογιστεί (αν υπάρχει) με βάση το Θεώρημα (KTL) υποψήφιο σημείο (τοπικού / ολικού) ελαχίστου για το πρόβλημα, $min_{x \in U} f(x, y, z)$ όπου
$f(x, y, z) = x + y + z$ και

$$U = \{[x, y, z]^T \in \mathbb{R}^3 / \frac{x^2}{4} + y^2 \leq 1, z = 2\}$$

Είναι το σημείο μοναδικό;

- *Λύση* (υπόδειξη): Παρατηρούμε πως το υποσύνολο $U \subset \mathbb{R}^3$ είναι κλειστό και φραγμένο (άρα συμπαγές), ενώ η f είναι συνεχής. Επομένως από το Θεώρημα 1, υπάρχει τουλάχιστον ένα σημείο ελαχίστου. Επίσης το $U \subset \mathbb{R}^3$ είναι κυρτό, ενώ η συνάρτηση f είναι κυρτή. Επομένως αναζητούμε μόνο σημείο(α) ελαχίστου. Χρησιμοποιώντας το θεώρημα (KTL)

$$f_0(x, y, z) = f(x, y, z), f_1(x, y, z) = \frac{x^2}{4} + y^2 - 1, f_2(x, y, z) = z - 2, l = 1, m = 2.$$

έχουμε,

$$\nabla f_0(\bar{x}) + \lambda_1 \nabla f_1(\bar{x}) + \lambda_2 \nabla f_2(\bar{x}) = 0,$$

$$\lambda_1 f_1(\bar{x}) = 0$$

$$f_2(\bar{x}) = 0$$

$\lambda_1 \geq 0$ αντιστοιχεί σε ανισότητα , $\lambda_2 \in \mathbb{R}$ αντιστοιχεί σε ισότητα

- συνέχεια λύσης):

$$\begin{bmatrix} 1+\lambda_1\frac{\bar{x}}{2} \\ 1+2\lambda_1\bar{y} \\ 1+\lambda_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\lambda_1(\frac{\bar{x}^2}{4}+\bar{y}^2-1)=0, \quad \lambda_1 \geq 0$$

$$\bar{z}=2.$$

Περίπτωση 1: Αν $\lambda_1 = 0$, τότε έχουμε άτοπο (βλ. δεύτερη εξίσωση) .

Περίπτωση 2: $\lambda_1 > 0$, τότε

$$\begin{bmatrix} \bar{x} \\ \bar{y} \\ \lambda_2 \end{bmatrix} = \begin{bmatrix} -\frac{2}{\lambda_1} \\ -\frac{1}{2\lambda_1} \\ -1 \end{bmatrix}$$

$$(\frac{\bar{x}^2}{4}+\bar{y}^2-1)=0, \bar{z}=2$$

Επομένως λαμβάνουμε, αντικαθιστώντας τις τιμές των $\bar{x}, \bar{y}$, στην τέταρτη εξίσωση,

$\lambda_1 = \frac{\sqrt{17}}{2}, \bar{x} = -\frac{4}{\sqrt{17}}, \bar{y} = -\frac{1}{\sqrt{17}}, \bar{z}=2, \lambda_2 = -1$ (η αρνητική τιμή του λ_1 απορρίπτεται)….

- *Άσκηση*:Να υπολογιστεί (αν υπάρχει) με βάση το Θεώρημα (KTL) υποψήφιο σημείο (τοπικού / ολικού) ελαχίστου για το πρόβλημα,

$min_{x \in U} f(x, y, z)$ όπου $f(x, y, z) = x^3 - y^3 + z$ και $U = \{[x, y, z]^T \in \mathbb{R}^3 / x^2 + y^2 + z^2 \leq 1\}$

Είναι το σημείο μοναδικό;

- *Λύση* (υπόδειξη): Παρατηρούμε πως το υποσύνολο $U \subset \mathbb{R}^3$ είναι κλειστό και φραγμένο (άρα συμπαγές), ενώ η f είναι συνεχής. Επομένως από το Θεώρημα 1, υπάρχει τουλάχιστον ένα σημείο ελαχίστου. Επομένως αναζητούμε σημείο(α) τοπικού / ολικού ελαχίστου χρησιμοποιώντας το θεώρημα (KTL)

$f_0(x, y, z) = f(x, y, z), f_1(x, y, z) = x^2 + y^2 + z^2 - 1$

έχουμε,

$$\nabla f_0(\bar{x}) + \lambda_1 \nabla f_1(\bar{x}) = 0,$$

$$\lambda_1 f_1(\bar{x}) = 0$$

$\lambda_1 \geq 0$ αντιστοιχεί σε ανισότητα .

- (*συνέχεια λύσης*):

$$\begin{bmatrix} 3\bar{x}^2 + 2\lambda_1\bar{x} \\ -3\bar{y}^2 + 2\lambda_1\bar{y} \\ 1 + 2\lambda_1\bar{z} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\lambda_1(\bar{x}^2 + \bar{y}^2 + \bar{z}^2 - 1) = 0, \quad \lambda_1 \geq 0$$

Περίπτωση 1: $\lambda_1 = 0$. Τότε έχουμε άτοπο (από την τρίτη εξίσωση).

Περίπτωση 2: $\lambda_1 > 0$. Τότε,

$$\begin{bmatrix} \bar{x}(3\bar{x} + 2\lambda_1) \\ \bar{y}(-3\bar{y} + 2\lambda_1) \\ \bar{z} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ -\frac{1}{2\lambda_1} \end{bmatrix}$$

$$(\bar{x}^2 + \bar{y}^2 + \bar{z}^2 - 1) = 0.$$

Περίπτωση 2α: $\bar{x} = 0, \bar{y} = 0$. Τότε έχουμε $\bar{z} = 1$ ή $\bar{z} = -1$. Όταν $\bar{z} = 1$, παρατηρούμε πως $\lambda_1 = -(1/2)$, άτοπο.

Επομένως καταλήγουμε στο $\bar{x} = 0, \bar{y} = 0, \bar{z} = -1, \lambda_1 = 1/2$. (Είναι τοπικό ή ολικό ελάχιστο ?)

Περίπτωση 2β: $\bar{x} = 0, \bar{y} \neq 0$. Τότε

$$\begin{bmatrix} \bar{x} \\ \bar{y} \\ \bar{z} \end{bmatrix} = \begin{bmatrix} 0 \\ \frac{2\lambda_1}{3} \\ -\frac{1}{2\lambda_1} \end{bmatrix}$$

$$(\bar{x}^2 + \bar{y}^2 + \bar{z}^2 - 1) = 0.$$

Αντικαθιστώντας στην τελευταία εξίσωση λαμβάνουμε:

- Συνέχεια (Λύσης): Τότε έχουμε $\dfrac{4\lambda_1^2}{9}+\dfrac{1}{4\lambda_1^2}=1$. Θέτοντας $4\lambda_1^2=\mu$ έχουμε $\mu^2-9\mu+9=0$. Λύνοντας ως προς μ, και αντικαθιστώντας έχουμε, $4\lambda_1^2=\dfrac{9+3\sqrt{5}}{2}$ ή $4\lambda_1^2=\dfrac{9-3\sqrt{5}}{2}$ που σημαίνει (επειδή $\lambda_1>0$) , $\lambda_1=\sqrt{\dfrac{9+3\sqrt{5}}{8}}$ ή $\lambda_1=\sqrt{\dfrac{9-3\sqrt{5}}{8}}$. Επομένως καταλήγουμε, στα σημεία: $\begin{bmatrix}\bar{x}\\ \bar{y}\\ \bar{z}\end{bmatrix}=\begin{bmatrix}0\\ \dfrac{2\sqrt{\frac{9+3\sqrt{5}}{8}}}{3}\\ -\dfrac{1}{2\sqrt{\frac{9+3\sqrt{5}}{8}}}\end{bmatrix}$ $\begin{bmatrix}\bar{x}\\ \bar{y}\\ \bar{z}\end{bmatrix}=\begin{bmatrix}0\\ \dfrac{2\sqrt{\frac{9-3\sqrt{5}}{8}}}{3}\\ -\dfrac{1}{2\sqrt{\frac{9-3\sqrt{5}}{8}}}\end{bmatrix}$ αντιστοίχως.

Ελέγχουμε αν τα δύο αυτά σημεία βρίσκονται στο σύνολο U.

- Περίπτωση 2γ: $\bar{x} \neq 0, \bar{y} = 0$. Τότε

$$\begin{bmatrix} \bar{x} \\ \bar{y} \\ \bar{z} \end{bmatrix} = \begin{bmatrix} -\frac{2\lambda_1}{3} \\ 0 \\ -\frac{1}{2\lambda_1} \end{bmatrix}$$

$(\bar{x}^2 + \bar{y}^2 + \bar{z}^2 - 1) = 0.$

Αντικαθιστώντας στην τελευταία εξίσωση λαμβάνουμε: Τότε έχουμε $\dfrac{4\lambda_1^2}{9} + \dfrac{1}{4\lambda_1^2} = 1$. Θέτοντας $4\lambda_1^2 = \mu$ έχουμε $\mu^2 - 9\mu + 9 = 0$. Λύνοντας ως προς μ, και αντικαθιστώντας έχουμε,

$4\lambda_1^2 = \dfrac{9 + 3\sqrt{5}}{2}$ ή $4\lambda_1^2 = \dfrac{9 - 3\sqrt{5}}{2}$ που σημαίνει (επειδή $\lambda_1 > 0$) ,

$\lambda_1 = \sqrt{\dfrac{9 + 3\sqrt{5}}{8}}$ ή $\lambda_1 = \sqrt{\dfrac{9 - 3\sqrt{5}}{8}}$. Επομένως καταλήγουμε, στα σημεία:

$$\begin{bmatrix} \bar{x} \\ \bar{y} \\ \bar{z} \end{bmatrix} = \begin{bmatrix} -\frac{2\sqrt{\frac{9+3\sqrt{5}}{8}}}{3} \\ 0 \\ -\frac{1}{2\sqrt{\frac{9+3\sqrt{5}}{8}}} \end{bmatrix} \qquad \begin{bmatrix} \bar{x} \\ \bar{y} \\ \bar{z} \end{bmatrix} = \begin{bmatrix} -\frac{2\sqrt{\frac{9-3\sqrt{5}}{8}}}{3} \\ 0 \\ -\frac{1}{2\sqrt{\frac{9-3\sqrt{5}}{8}}} \end{bmatrix}$$

αντιστοίχως. Ελέγχουμε αν τα δύο αυτά σημεία βρίσκονται στο σύνολο U.

- Περίπτωση 2δ: $\bar{x} \neq 0, \bar{y} \neq 0$. Τότε

$$\begin{bmatrix} \bar{x} \\ \bar{y} \\ \bar{z} \end{bmatrix} = \begin{bmatrix} -\frac{2\lambda_1}{3} \\ \frac{2\lambda_1}{3} \\ -\frac{1}{2\lambda_1} \end{bmatrix}$$

$$(\bar{x}^2 + \bar{y}^2 + \bar{z}^2 - 1) = 0.$$

Αντικαθιστώντας στην τελευταία εξίσωση λαμβάνουμε: Τότε έχουμε $\frac{8\lambda_1^2}{9} + \frac{1}{4\lambda_1^2} = 1$. Θέτοντας $4\lambda_1^2 = \mu$ έχουμε $2\mu^2 - 9\mu + 9 = 0$. Λύνοντας ως προς μ, και αντικαθιστώντας έχουμε,

$4\lambda_1^2 = \frac{9+3}{4} = 3$ ή $4\lambda_1^2 = \frac{9-3}{4} = \frac{3}{2}$ που σημαίνει (επειδή $\lambda_1 > 0$) , $\lambda_1 = \frac{\sqrt{3}}{2}$ ή $\lambda_1 = \frac{\sqrt{3}}{2\sqrt{2}}$. .

Επομένως καταλήγουμε, στα σημεία: $\begin{bmatrix} \bar{x} \\ \bar{y} \\ \bar{z} \end{bmatrix} = \begin{bmatrix} -\frac{\sqrt{3}}{3} \\ \frac{\sqrt{3}}{3} \\ -\frac{1}{\sqrt{3}} \end{bmatrix}$ $\begin{bmatrix} \bar{x} \\ \bar{y} \\ \bar{z} \end{bmatrix} = \begin{bmatrix} -\frac{\sqrt{3}}{3\sqrt{2}} \\ \frac{\sqrt{3}}{3\sqrt{2}} \\ -\frac{\sqrt{2}}{\sqrt{3}} \end{bmatrix}$ αντιστοίχως. Ελέγχουμε αν τα δύο αυτά σημεία βρίσκονται στο σύνολο U. Παρατηρούμε πως το δεύτερο σημείο ανήκει!.

- Βιβλιογραφία:

- Αλ. Μπακόπουλος και Ιων. Χρυσοβέργης, “Εισαγωγή στην Αριθμητική Ανάλυση”, Εκδόσεις Συμμετρία.
- Γ. Ακριβής και Β. Δουγαλής, “Εισαγωγή στην Αριθμητική Ανάλυση”, Πανεπιστημιακές Εκδόσεις Κρήτης.
- J. Nocedal and S. Wright, “Numerical Optimization”, Springer-Verlag 2006.
- E. Polak, “Optimization”, Springer-Verlag 1997