

Τεχνικές Βελτιστοποίησης

Συλλογή Παλαιών Θεμάτων & Λύσεων Γαριδομακαροναδα

Επιμέλεια: VM+DK

Credits στα παιδιά που βγάλανε τις φωτογραφίες, στα παιδιά που βοήθησαν με τις λύσεις και στον
Αστακομακαροναδα που ξεκίνησε την ιδέα.

Disclaimer: Οι παρούσες λύσεις και εκφωνήσεις ενδέχεται να περιέχουν λάθη. Κάντε comment στο github μου για τα λάθη.

Περιεχόμενα

- 1. Ιούνιος 2014**
- 2. Ιούλιος 2018**
- 3. Φεβρουάριος 2020**
- 4. Φεβρουάριος 2021**
- 5. Σεπτέμβριος 2021**
- 6. Φεβρουάριος 2022**
- 7. Ιούνιος 2022**
- 8. Σεπτέμβριος 2022**
- 9. Φεβρουάριος 2023**
- 10. Ιούνιος 2023**
- 11. Σεπτέμβριος 2023**
- 12. Ιούνιος 2024**
- 13. Σεπτέμβριος 2024**
- 14. Φεβρουάριος 2025**
- 15. Ιούλιος 2025**
- 16. Σεπτέμβριος 2025**

Λυμένες Ασκήσεις Βιβλίου & Θεωρία

Ασκήσεις 3.6.3 -- Σελίδα 71

Ασκηση 3.6.3 (α)

Δείξτε ότι:

$$\frac{x}{4} + \frac{3y}{4} \leq \sqrt{\ln\left(\frac{e^{x^2}}{4} + \frac{3}{4}e^{y^2}\right)}, \quad \forall x, y > 0$$

Λύση

Λύση:

Θέλουμε να δείξουμε ότι:

$$\frac{x}{4} + \frac{3y}{4} \leq \sqrt{\ln\left(\frac{e^{x^2}}{4} + \frac{3}{4}e^{y^2}\right)} \Rightarrow \left(\frac{x+3y}{4}\right)^2 \leq \ln\left(\frac{e^{x^2}}{4} + \frac{3}{4}e^{y^2}\right) \quad (1)$$

Θεωρούμε τη συνάρτηση $\phi(t) = t^2$, η οποία είναι **κυρτή**. Από την ανισότητα Jensen:

$$f\left(\sum_{i=1}^k \lambda_i x_i\right) \leq \sum_{i=1}^k \lambda_i f(x_i)$$

Για $\lambda_1 = \frac{1}{4}$, $\lambda_2 = \frac{3}{4}$ ($\lambda_1 + \lambda_2 = 1$):

$$\left(\frac{1}{4}x + \frac{3}{4}y\right)^2 \leq \frac{1}{4}x^2 + \frac{3}{4}y^2 \quad (2)$$

Επίσης, θεωρούμε τη συνάρτηση $\psi(t) = -\ln t$, η οποία είναι **κυρτή** (ή $\phi(t) = \ln t$ κοιλη). Εφαρμόζοντας πάλι την ανισότητα Jensen για την $\ln t$:

$$\ln\left(\sum \lambda_i z_i\right) \geq \sum \lambda_i \ln z_i$$

Για $z_1 = e^{x^2}$ και $z_2 = e^{y^2}$:

$$\ln\left(\frac{1}{4}e^{x^2} + \frac{3}{4}e^{y^2}\right) \geq \frac{1}{4}\ln(e^{x^2}) + \frac{3}{4}\ln(e^{y^2}) = \frac{1}{4}x^2 + \frac{3}{4}y^2 \quad (3)$$

Από τις ανισότητες (2) και (3) προκύπτει:

$$\ln\left(\frac{1}{4}e^{x^2} + \frac{3}{4}e^{y^2}\right) \geq \frac{1}{4}x^2 + \frac{3}{4}y^2 \geq \left(\frac{x+3y}{4}\right)^2$$

Άρα:

$$\frac{x}{4} + \frac{3y}{4} \leq \sqrt{\ln\left(\frac{e^{x^2}}{4} + \frac{3}{4}e^{y^2}\right)}$$

Ασκηση 3.6.3 (β)

Δείξτε ότι:

$$\left(\frac{x}{2} + \frac{y}{3} + \frac{z}{12} + \frac{w}{12}\right)^4 \leq \frac{1}{2}x^4 + \frac{1}{3}y^4 + \frac{1}{12}z^4 + \frac{1}{12}w^4$$

με την ισότητα να ισχύει αν και μόνο αν $x = y = z = w$.

Λύση

Θεωρία: Θεώρημα 3.1.3 (Jensen). Έστω $f(x)$ κυρτή συνάρτηση ορισμένη σε κυρτό υποσύνολο του \mathbb{R}^n . Αν $\lambda_i \geq 0$ για $i = 1, \dots, k$ και $\sum_{i=1}^k \lambda_i = 1$, τότε:

$$f\left(\sum_{i=1}^k \lambda_i x_i\right) \leq \sum_{i=1}^k \lambda_i f(x_i) \quad (1)$$

Αν η f είναι **γνήσια κυρτή** και $\lambda_i > 0, i = 1, \dots, k$, τότε η ισότητα στην (1) ισχύει αν και μόνο αν τα $x_i, i = 1, \dots, k$ είναι **ίσα**.

Άσκηση 3.5.8 -- Σελίδα 55**Ανισότητα Αριθμητικού-Γεωμετρικού Μέσου**

Έστω $x_i, i = 1, 2, \dots, n$ θετικοί πραγματικοί αριθμοί και $\delta_i > 0, i = 1, 2, \dots, n$ με $\sum_{i=1}^n \delta_i = 1$. Τότε:

$$\prod_{i=1}^n x_i^{\delta_i} \leq \sum_{i=1}^n \delta_i x_i$$

με την ισότητα να ισχύει αν και μόνο αν $x_1 = x_2 = \dots = x_n$.

Λύση**Απόδειξη της Ανισότητας με τη χρήση του Θεωρήματος 3.1.3:**

Θεωρώ την $f(z) = -\ln z, z > 0$. Η $f(z)$ είναι γνήσια κυρτή. Έτσι, μπορώ να εφαρμόσω το Θεώρημα 3.1.3 για κάποια x_i και δ_i τα οποία $\sum_{i=1}^n \delta_i = 1$.

Ισχύει ότι:

$$-\ln \left(\sum_{i=1}^n \delta_i x_i \right) \leq \sum_{i=1}^n \delta_i (-\ln x_i) = -\sum_{i=1}^n \delta_i \ln(x_i) \implies$$

$$\ln \left(\sum_{i=1}^n \delta_i x_i \right) \geq \sum_{i=1}^n \delta_i \ln(x_i) \implies$$

χρησιμοποιώ την ιδιότητα λογαρίθμου $\ln(a) + \ln(b) = \ln(ab)$:

$$\ln \left(\sum_{i=1}^n \delta_i x_i \right) \geq \sum_{i=1}^n \ln(x_i^{\delta_i}) = \ln \left(\prod_{i=1}^n x_i^{\delta_i} \right) \implies$$

Επειδή η \ln είναι γνησίως αύξουσα ($\ln \uparrow$):

$$\sum_{i=1}^n \delta_i x_i \geq \prod_{i=1}^n x_i^{\delta_i} \implies \boxed{\prod_{i=1}^n x_i^{\delta_i} \leq \sum_{i=1}^n \delta_i x_i}$$

με την ισότητα να ισχύει αν και μόνο αν $x_1 = x_2 = \dots = x_n$.

Άσκηση 3.6.9 -- Σελίδα 72 (Άλυτες Ασκήσεις)

Χρήση της Ανισότητας Hölder

Κάνοντας χρήση της ανισότητας του Hölder, δείξτε ότι αν $a_i, i = 1, 2, \dots, m$ είναι σταθερά διανύσματα στο \mathbb{R}^n και αν $c_i > 0, i = 1, 2, \dots, m$, τότε η:

$$f(x) = \ln \left(\sum_{i=1}^m c_i e^{a_i^T x} \right)$$

είναι κυρτή στο \mathbb{R}^n .

Λύση

Ουσιαστικά η εκφώνηση μας δίνει μία συνάρτηση $f(x)$ και μας λέει να αποδείξουμε ότι είναι κυρτή στο πεδίο ορισμού της. (Η απόδειξη της κυρτότητας της $f(x)$ σύμφωνα με την άσκηση πρέπει να γίνει με την χρήση της ανισότητας Hölder).

Ανισότητα Hölder (σελ. 61)

Αν $p, q > 1$ και $\frac{1}{p} + \frac{1}{q} = 1$, τότε $\forall x, y \in \mathbb{R}^n$:

$$\sum_{i=1}^n |x_i y_i| \leq \left(\sum_{i=1}^n |x_i|^p \right)^{\frac{1}{p}} \left(\sum_{i=1}^n |y_i|^q \right)^{\frac{1}{q}}$$

Λύση:

Για να είναι κυρτή η $f(x)$ πρέπει να ισχύει το Θ. 3.1.3 (Jensen). Θα γράψουμε την ανισότητα αυτή του Θ. 3.1.3 και αν αυτή ικανοποιείται, τότε η $f(x)$ θα είναι κυρτή. (Αυτό θα το κάνουμε για 2 σημεία με 2 συντελεστές).

1) Επιλέγουμε τους συντελεστές λ_i με $i = 1, 2$ ως εξής:

$$\lambda_1 = \frac{1}{p} \quad \text{και} \quad \lambda_2 = \frac{1}{q}$$

Ισχύει ότι: $\lambda_1 + \lambda_2 = \frac{1}{p} + \frac{1}{q} = 1$.

Από το Θεώρημα 3.1.3 θα πρέπει να ελέγχω κατά πόσο το παρακάτω είναι αληθές:

$$\ln \left(\sum_{i=1}^m c_i e^{a_i^T (\frac{1}{p}x + \frac{1}{q}y)} \right) \leq \frac{1}{p} \ln \left(\sum_{i=1}^m c_i e^{a_i^T x} \right) + \frac{1}{q} \ln \left(\sum_{i=1}^m c_i e^{a_i^T y} \right) \quad (1)$$

Β' ΜΕΛΟΣ =

$$\begin{aligned} & \frac{1}{p} \ln \left(\sum_{i=1}^m c_i e^{a_i^T x} \right) + \frac{1}{q} \ln \left(\sum_{i=1}^m c_i e^{a_i^T y} \right) = \\ & = \ln \left(\sum_{i=1}^m c_i e^{a_i^T x} \right)^{\frac{1}{p}} + \ln \left(\sum_{i=1}^m c_i e^{a_i^T y} \right)^{\frac{1}{q}} = \end{aligned}$$

χρησ. ιδιότητα λογαρίθμων: $\ln a + \ln b = \ln(ab)$

$$= \ln \left[\left(\sum_{i=1}^m c_i e^{a_i^T x} \right)^{\frac{1}{p}} \left(\sum_{i=1}^m c_i e^{a_i^T y} \right)^{\frac{1}{q}} \right] \quad (2)$$

Αυτό έχει αρχίσει με την ανισότητα του Hölder που λέει η εκφώνηση να χρησιμοποιήσω. Δουλεύω λίγο ακόμη το Β' μέλος, για να το κάνω ακριβώς να μοιάζει με το Hölder... Παίρνω λοιπόν λίγο την ποσότητα του ln του Β' μέλους:

$$\left(\sum_{i=1}^m c_i e^{a_i^T x} \right)^{\frac{1}{p}} \left(\sum_{i=1}^m c_i e^{a_i^T y} \right)^{\frac{1}{q}} \quad (3)$$

Άρα, έχουμε ότι:

$$\begin{aligned} & \left(\sum_{i=1}^m c_i e^{a_i^T x} \right)^{\frac{1}{p}} \left(\sum_{i=1}^m c_i e^{a_i^T y} \right)^{\frac{1}{q}} = \\ & = \left(\sum_{i=1}^m \left| c_i^{\frac{1}{p}} e^{\frac{1}{p} a_i^T x} \right|^p \right)^{\frac{1}{p}} \left(\sum_{i=1}^m \left| c_i^{\frac{1}{q}} e^{\frac{1}{q} a_i^T y} \right|^q \right)^{\frac{1}{q}} \end{aligned}$$

Αυτή η ποσότητα είναι ακριβώς στην ίδια μορφή με την ανισότητα Hölder. Οπότε, εφαρμόζοντας την ανισότητα Hölder:

$$\begin{aligned} & \left(\sum_{i=1}^m \left| c_i^{\frac{1}{p}} e^{\frac{1}{p} a_i^T x} \right|^p \right)^{\frac{1}{p}} \left(\sum_{i=1}^m \left| c_i^{\frac{1}{q}} e^{\frac{1}{q} a_i^T y} \right|^q \right)^{\frac{1}{q}} \geq \\ & \geq \sum_{i=1}^m \left| c_i^{\frac{1}{p}} e^{\frac{1}{p} a_i^T x} \cdot c_i^{\frac{1}{q}} e^{\frac{1}{q} a_i^T y} \right| \end{aligned}$$

Βγάζω την απόλυτη τιμή γιατί όλες οι ποσότητες είναι θετικές:

$$\begin{aligned} & = \sum_{i=1}^m c_i^{\left(\frac{1}{p} + \frac{1}{q}\right)} e^{\frac{1}{p} a_i^T x + \frac{1}{q} a_i^T y} \xrightarrow{\ln \uparrow} \dots \\ \implies & \ln \left(\sum_{i=1}^m c_i e^{a_i^T (\frac{1}{p} x + \frac{1}{q} y)} \right) \leq \ln \left[\left(\sum_{i=1}^m c_i e^{a_i^T x} \right)^{\frac{1}{p}} \left(\sum_{i=1}^m c_i e^{a_i^T y} \right)^{\frac{1}{q}} \right] \\ \implies & \ln \left(\sum_{i=1}^m c_i e^{a_i^T (\frac{1}{p} x + \frac{1}{q} y)} \right) \leq \frac{1}{p} \ln \left(\sum_{i=1}^m c_i e^{a_i^T x} \right) + \frac{1}{q} \ln \left(\sum_{i=1}^m c_i e^{a_i^T y} \right) \end{aligned}$$

Η αρχική ανισότητα (1) είναι αληθής, άρα η f είναι κυρτή.

Βασικά Θεωρήματα & Συμπεράσματα

Λύση

Ξεκινάμε με τα θεωρήματα που θα χρησιμοποιήσουμε σήμερα, για να λύσουμε τις ασκήσεις...

Θεώρημα 3.1.4

Κάθε τοπικό ελάχιστο μιας κυρτής συνάρτησης f ορισμένης σε κάποιο κυρτό υποσύνολο του $S \subseteq \mathbb{R}^n$ είναι και ολικό ελάχιστο. Αν η f είναι γνήσια κυρτή, τότε το ολικό ελάχιστο είναι και μοναδικό.

Συμπέρασμα 3.1.1

Αν η f είναι κυρτή με συνεχείς μερικές παραγώγους $1^{\text{ης}}$ τάξης ορισμένη στο $S \subseteq \mathbb{R}^n$, τότε κάθε κρίσιμο σημείο της f , αποτελεί και ολικό ελάχιστο.

Θεώρημα 3.1.6

Αν η f έχει συνεχείς παραγώγους $2^{\text{ης}}$ τάξης σε ανοιχτό κυρτό υποσύνολο $S \subseteq \mathbb{R}^n$ και αν ο $\nabla^2 f(x)$ είναι θετικά ημιορισμένος, τότε η f είναι κυρτή, ενώ αν ο $\nabla^2 f(x)$ είναι θετικά ορισμένος η f είναι γνήσια κυρτή.

Θεώρημα 3.1.7

- i) Αν έχω κάποιες συναρτήσεις $f_i(x), i = 1, \dots, k$ που είναι κυρτές σε κυρτό υποσύνολο $S \subseteq \mathbb{R}^n$, τότε η $f(x) = \sum_{i=1}^k f_i(x)$ είναι κυρτή συνάρτηση. Αν τουλάχιστον μία από τις $f_i(x)$ είναι γνήσια κυρτή, τότε και η $f(x)$ είναι γνήσια κυρτή.
- ii) Για κάποιο $a > 0$ η $af(x)$ είναι (γνήσια) κυρτή, αν η $f(x)$ είναι (γνήσια) κυρτή.
- iii) Αν η f είναι (γνήσια) κυρτή και η g είναι κυρτή και (γνησίως) αύξουσα, τότε η $g(f(x))$ είναι (γνήσια) κυρτή.

Άσκηση 3.6.1 -- Σελίδα 70

Κυρτότητα Συναρτήσεων

Δείξτε αν οι παρακάτω συναρτήσεις είναι κυρτές ή αυστηρά (γνήσια) κυρτές:

$$\delta) f(x_1, x_2) = 4e^{3x_1 - x_2} + 5e^{x_1^2 + x_2^2}, \quad x = [x_1 \ x_2]^T \in \mathbb{R}^2$$

Λύση

Λύση:

Ορίζω την $f_1(x_1, x_2) = 4e^{3x_1 - x_2}$ και την $g_1(y) = 4e^y$. Επίσης ορίζω την $h_1(x_1, x_2) = 3x_1 - x_2$.

Γνωρίζουμε ότι κάθε γραμμικός συνδυασμός της μορφής $f(x) = a^T x + b$, $x \in \mathbb{R}^n$, $b \in \mathbb{R}$, $a \in \mathbb{R}^n$ είναι κυρτή συνάρτηση (δες παραδ. 3.3.1 σελ. 25). Επομένως, η $h_1(x_1, x_2) = 3x_1 - x_2$ είναι **κυρτή** συνάρτηση.

Επίσης, η $g_1(y)$ είναι **γνησίως αύξουσα** και **γνήσια κυρτή**, διότι:

$$g'_1(y) = 4e^y > 0 \quad \text{και} \quad g''_1(y) = 4e^y > 0$$

Παρατηρώ ότι $f_1(x_1, x_2) = g_1(h_1)$. Επειδή η h_1 είναι κυρτή συνάρτηση και η g_1 είναι γνήσια κυρτή και γνησίως αύξουσα, η $f_1(x_1, x_2) = g_1(h_1)$ είναι **κυρτή** (Εδώ χρησιμοποιούμε ουσιαστικά το Θ. 3.1.7 (iii)).

Ορίζω τώρα τις: $f_2(x_1, x_2) = 5e^{x_1^2 + x_2^2}$ $g_2(y) = 5e^y$ $h_2(x_1, x_2) = x_1^2 + x_2^2$

Έχουμε ότι:

- Η $h_2(x_1, x_2)$ είναι **γνήσια κυρτή** συνάρτηση ως άθροισμα δύο γνήσιων κυρτών συναρτήσεων ($q_1(x_1) = x_1^2$, $q_2(x_2) = x_2^2$). (Σημείωση: Οι $q_1(x_1)$, $q_2(x_2)$ είναι γνήσια κυρτές συναρτήσεις γιατί είναι της μορφής $(a^T x)^2$).
- Η $g_2(y)$ είναι **γνήσια κυρτή** και **γνησίως αύξουσα**.

Έτσι, λόγω του Θ. 3.1.7 (iii) η $f_2(x_1, x_2) = g_2(h_2)$ είναι **γνήσια κυρτή** συνάρτηση.

Παρατηρούμε ότι η $f(x) = f_1(x) + f_2(x)$. Η $f_1(x)$ (όπως δείξαμε) είναι κυρτή, ενώ η $f_2(x)$ είναι γνήσια κυρτή. Έτσι, από το Θ. 3.1.7 (i) η $f(x)$ είναι **γνήσια κυρτή**.

Άσκηση 3.6.5

Χρήση Α-Γ Ανισότητας

Χρησιμοποιείστε την (Α-Γ) ανισότητα για να λύσετε τα παρακάτω προβλήματα:

a) $\min(x^2 + y + z)$ υπό τους περιορισμούς: $xyz = 1$, $x, y, z > 0$

Λύση

Λύση:

Αρχικά, πάμε να θυμήσουμε την (Α-Γ) → Ανισότητα Αριθμητικού Γεωμετρικού Μέσου:

Ανισότητα Α-Γ

Έστω $x_i, i = 1, \dots, n, x_i > 0$ και $\delta_i, i = 1, \dots, n, \delta_i > 0$ και $\sum_{i=1}^n \delta_i = 1$, τότε:

$$\prod_{i=1}^n x_i^{\delta_i} \leq \sum_{i=1}^n \delta_i x_i$$

- Έστω ότι υπάρχουν κάποια $\lambda_1, \lambda_2, \lambda_3 \geq 0$ τέτοια ώστε $\sum_{i=1}^3 \lambda_i = 1$.

- Από (Α-Γ):

$$\begin{aligned} \lambda_1(x^2) + \lambda_2(y) + \lambda_3(z) &\geq (x^2)^{\lambda_1} \cdot y^{\lambda_2} \cdot z^{\lambda_3} \\ \implies (x^2)^{\lambda_1} \cdot y^{\lambda_2} \cdot z^{\lambda_3} &\leq \lambda_1(x^2) + \lambda_2(y) + \lambda_3(z) \end{aligned} \quad (1)$$

- Θα πρέπει: $2\lambda_1 = \lambda_2 = \lambda_3$ (2) και $\lambda_1 + \lambda_2 + \lambda_3 = 1$ (3).

- Από τις (2) και (3) έχουμε ότι:

$$5\lambda_1 = 1 \implies \lambda_1 = \frac{1}{5} \quad \text{και} \quad \lambda_2 = \lambda_3 = \frac{2}{5}$$

Αντικαθιστώντας στην (1) έχουμε ότι:

$$\begin{aligned} x^2 + y + z &= 5 \left(\frac{1}{5}x^2 + \frac{2}{5} \left(\frac{y}{2} \right) + \frac{2}{5} \left(\frac{z}{2} \right) \right) \geq \\ &\geq 5 \left(x^2 \right)^{\frac{1}{5}} \cdot \left(\frac{y}{2} \right)^{\frac{2}{5}} \cdot \left(\frac{z}{2} \right)^{\frac{2}{5}} \implies \\ \implies x^2 + y + z &\geq 5 \left(\frac{x \cdot y \cdot z}{2 \cdot 2} \right)^{\frac{2}{5}} \implies x^2 + y + z \geq 5 \frac{(xyz)^{\frac{2}{5}}}{4^{\frac{2}{5}}} \end{aligned}$$

Από τους περιορισμούς ($xyz = 1$):

$$\implies x^2 + y + z \geq \frac{5}{4^{\frac{2}{5}}} \implies x^2 + y + z \geq \frac{5}{\sqrt[5]{4^2}} \quad (4)$$

Για ποιες τιμές του x, y, z παίρνουμε την ισότητα; Από την θεωρία της Α-Γ η ισότητα ισχύει όταν τα x_i είναι ίσα μεταξύ τους. Έτσι, η ισότητα στην (4) ισχύει για: $x^2 = \frac{y}{2} = \frac{z}{2}$. Όμως: $xyz = 1$ και $x, y, z > 0$. Έτσι, λύνοντας το σύστημα:

$$\begin{cases} xyz = 1 \\ x, y, z > 0 \\ x^2 = \frac{y}{2} = \frac{z}{2} \end{cases}$$

Βρίσκουμε το σημείο (x, y, z) για το οποίο η $x^2 + y + z$ ελαχιστοποιείται. Το (τοπικό) ελάχιστο της $x^2 + y + z$ είναι το: $\frac{5}{4^{2/5}}$.

Και επειδή η $x^2 + y + z$ είναι γνήσια κυρτή, το $\frac{5}{4^{2/5}}$ είναι και το ολικό (και μάλιστα μοναδικό) ελάχιστο.

Άσκηση 3.6.10 -- Σελίδα 72

Μέγιστος Κύλινδρος σε Σφαίρα

Κάνοντας χρήση της (Α-Γ) ανισότητας λύστε τα παρακάτω προβλήματα:

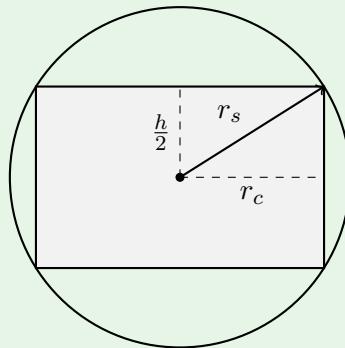
α) Βρείτε τον μεγαλύτερο κύλινδρο που μπορεί να εγγραφεί σε σφαίρα δοσμένης ακτίνας.

Λύση

Λύση:

- Έστω r_s ακτίνα σφαίρας.
- Ξέρουμε ότι ο όγκος του κυλίνδρου είναι: $V_c = \pi \cdot r_c^2 \cdot h$, όπου r_c η ακτίνα του κυλίνδρου και h το ύψος του κυλίνδρου. (Αυτή είναι η συνάρτηση προς μεγιστοποίηση).

Σχήμα: Παρακάτω βλέπουμε μια τομή του προβλήματος της σφαίρας:



Από πυθαγόρειο θεώρημα έχουμε ότι:

$$r_c^2 + \frac{h^2}{4} = r_s^2$$

Παρατηρώ ότι:

$$\begin{aligned} 3 \left(\frac{2}{3} \left(\frac{r_c^2}{2} \right) + \frac{1}{3} \left(\frac{h^2}{4} \right) \right) &= r_c^2 + \frac{h^2}{4} \xrightarrow{(A-\Gamma)} \\ \geq 3 \left(\left(\frac{r_c^2}{2} \right)^{\frac{2}{3}} \left(\frac{h^2}{4} \right)^{\frac{1}{3}} \right) &\implies r_s^2 \geq \frac{3}{2^{\frac{2}{3}} 4^{\frac{1}{3}}} (r_c^2)^{\frac{2}{3}} (h^2)^{\frac{1}{3}} \\ \implies r_s^2 &\geq \frac{3}{2^{\frac{2}{3}} 4^{\frac{1}{3}} \pi^{\frac{2}{3}}} (\pi r_c^2 h)^{\frac{2}{3}} \\ \implies r_s^2 &\geq \frac{3}{2^{\frac{2}{3}} 4^{\frac{1}{3}} \pi^{\frac{2}{3}}} (V_c)^{\frac{2}{3}} \end{aligned}$$

Λύνοντας ως προς V_c :

$$\implies V_c \leq \frac{4\pi r_s^3}{3\sqrt{3}} \quad (\text{ή } \frac{4\pi r_s^3}{3^{\frac{3}{2}}})$$

Για την ισότητα έχουμε τον μέγιστο όγκο του κυλίνδρου, που μπορεί να εγγραφεί στην σφαίρα δοσμένης ακτίνας.

- Η ισότητα ισχύει όταν: $\frac{r_c^2}{2} = \frac{h^2}{4}$ (1)

- Όμως $r_c^2 + \frac{h^2}{4} = r_s^2$ (2)

Ετσι, λύνοντας το σύστημα (1), (2) μπορούμε να βρούμε τα h, r_c συναρτήσει της r_s .

Άσκηση 3.6.8 -- Σελίδα 72

Ελαχιστοποίηση Συνάρτησης

Βρείτε τις τιμές του $x > 0$ που ελαχιστοποιούν τη συνάρτηση:

$$f(x) = c_1 x^3 + \frac{c_2}{x}, \quad \text{με } c_1, c_2 \text{ θετικές σταθερές.}$$

Λύση

Λύση:

$$f'(x) = 3c_1 x^2 - \frac{c_2}{x^2}$$

Επειδή $x > 0$ βρίσκουμε πότε $f'(x) = 0$:

$$\begin{aligned} f'(x) = 0 \implies 3c_1 x^2 - \frac{c_2}{x^2} = 0 &\xrightarrow{x \neq 0} 3c_1 x^4 - c_2 = 0 \implies \\ \implies 3c_1 x^4 = c_2 \implies x^4 = \frac{c_2}{3c_1} \implies x^* = \left(\frac{c_2}{3c_1} \right)^{\frac{1}{4}} \end{aligned}$$

(το κρίσιμο σημείο της f).

Ελέγχουμε την δεύτερη παράγωγο:

$$f''(x) = 6c_1 x + \frac{2c_2}{x^3} > 0, \quad \forall x > 0$$

(αφού $c_1, c_2 > 0$ και $x > 0$).

$\nabla^2 f > 0 \implies H f$ είναι γνήσια κυρτή $\forall x > 0$. Από το Θεώρημα 3.1.4 (ή Συμπέρασμα 3.1.1), επειδή η f είναι γνήσια κυρτή, το κρίσιμο σημείο x^* είναι μοναδικό ολικό ελάχιστο.

Άσκηση 3.6.14 -- Σελίδα 73**Αυστηρή Κυρτότητα**

Έστω $f(x)$ μια αυστηρά κυρτή συνάρτηση στο \mathbb{R}^n . Αν x, y είναι διακριτά σημεία στο \mathbb{R}^n τέτοια ώστε $f(x) = f(y) = 0$, δείξτε ότι υπάρχει $z \in \mathbb{R}^n$ τέτοιο ώστε $f(z) < 0$.

Λύση**Λύση:**

Θα εφαρμόσω τον ορισμό της κυρτότητας. Επειδή η f είναι γνήσια (αυστηρά) κυρτή, έχουμε από τον ορισμό της κυρτότητας ότι για $x \neq y$ και $\lambda \in (0, 1)$:

$$f(\lambda x + (1 - \lambda)y) < \lambda f(x) + (1 - \lambda)f(y) \quad (1)$$

Όμως $f(x) = f(y) = 0$. Οπότε η (1) γίνεται:

$$\begin{aligned} f(\lambda x + (1 - \lambda)y) &< \lambda \cdot 0 + (1 - \lambda) \cdot 0 = 0 \\ \implies \boxed{f(\lambda x + (1 - \lambda)y) < 0} \end{aligned}$$

Άρα, αν επιλέξουμε $z = \lambda x + (1 - \lambda)y$ για κάποιο $\lambda \in (0, 1)$, τότε $f(z) < 0$. (Το z είναι ένα σημείο στο ευθύγραμμο τμήμα που ενώνει τα x και y).

Άσκηση 3.6.21 -- Σελίδα 74

Συνθήκες KKT

Εφαρμόστε το Θεώρημα των Karush-Kuhn-Tucker (KKT) για να προσδιορίσετε όλες τις λύσεις του προβλήματος:

Βρείτε τα x_1, x_2 που ελαχιστοποιούν την:

$$f(x_1, x_2) = e^{-(x_1+x_2)}$$

και ικανοποιούν:

$$e^{x_1} + e^{x_2} \leq 20 \quad \text{και} \quad x_1 \geq 0$$

Είναι το πρόβλημα κυρτό;

Λύση

Λύση:

Ουσιαστικά προσπαθούμε να βρούμε τα x_1, x_2 που ελαχιστοποιούν την αντικειμενική συνάρτηση $f(x)$ και ταυτόχρονα ικανοποιούν τους περιορισμούς:

$$e^{x_1} + e^{x_2} \leq 20, \quad x_1 \geq 0$$

1) Ας ονοματίσουμε λίγο πρώτα κάποια πράγματα:

Για τους περιορισμούς: Θεωρώ $g_1(x_1, x_2) = e^{x_1} + e^{x_2} - 20$ (για να είναι ≤ 0) και $g_2(x_1, x_2) = -x_1$ (για να είναι ≤ 0).

Για την αντικειμενική συνάρτηση έχουμε:

$$f(x_1, x_2) = e^{-(x_1+x_2)}$$

2) Είναι το πρόβλημα κυρτό;

Θυμόμαστε ότι ένα πρόβλημα ονομάζεται κυρτό, όταν η αντικειμενική συνάρτηση και οι περιορισμοί είναι κυρτές συναρτήσεις.

Για την $f(x_1, x_2)$:

$$\nabla f(x) = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \end{bmatrix} = \begin{bmatrix} -e^{-(x_1+x_2)} \\ -e^{-(x_1+x_2)} \end{bmatrix}$$

$$\nabla^2 f(x) = \begin{bmatrix} e^{-(x_1+x_2)} & e^{-(x_1+x_2)} \\ e^{-(x_1+x_2)} & e^{-(x_1+x_2)} \end{bmatrix}$$

Ο $\nabla^2 f(x)$ είναι θετικά ημιορισμένος (Ορίζουσα = 0, ίχνος > 0). Έτσι, η $f(x)$ είναι κυρτή.

Για την $g_1(x)$:

- $\nabla g_1(x) = \begin{bmatrix} e^{x_1} \\ e^{x_2} \end{bmatrix}$
- $\nabla^2 g_1(x) = \begin{bmatrix} e^{x_1} & 0 \\ 0 & e^{x_2} \end{bmatrix}$

Ο $\nabla^2 g_1(x)$ είναι θετικά ορισμένος (διαγώνιοι > 0). Άρα, η $g_1(x)$ είναι γνήσια κυρτή συνάρτηση.

Η $g_2(x) = -x_1$ είναι γραμμική, άρα και κυρτή. Θα εφαρμόσουμε τώρα το Θεώρημα των KKT conditions...

KKT conditions:

Επειδή έχουμε δύο περιορισμούς θα έχουμε και δύο συντελεστές $\lambda_1, \lambda_2 \geq 0$. Δηλαδή θα έχουμε ότι:

- $\lambda_1, \lambda_2 \geq 0$
- $\lambda_1 g_1(x) = 0 \implies \lambda_1(e^{x_1} + e^{x_2} - 20) = 0 \quad (1)$
- $\lambda_2 g_2(x) = 0 \implies \lambda_2(-x_1) = 0 \quad (2)$
- $\nabla f(x) + \lambda_1 \nabla g_1(x) + \lambda_2 \nabla g_2(x) = 0$

Γράφοντας ξανά αυτές τις σχέσεις (συνοψίζοντας):

$$\begin{bmatrix} -e^{-(x_1+x_2)} \\ -e^{-(x_1+x_2)} \end{bmatrix} + \lambda_1 \begin{bmatrix} e^{x_1} \\ e^{x_2} \end{bmatrix} + \lambda_2 \begin{bmatrix} -1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Έχουμε δημιουργήσει το εξής σύστημα εξισώσεων:

1. $\lambda_1, \lambda_2 \geq 0$
2. $\lambda_1(e^{x_1} + e^{x_2} - 20) = 0$
3. $-\lambda_2 x_1 = 0$
4. $-e^{-(x_1+x_2)} + \lambda_1 e^{x_1} - \lambda_2 = 0 \quad (3)$
5. $-e^{-(x_1+x_2)} + \lambda_1 e^{x_2} = 0 \quad (4)$

Αν υπάρχουν λοιπόν τέτοια λ_i και x_i , ώστε να ικανοποιούν το παραπάνω σύστημα, τότε αυτά τα λ_i και x_i αποτελούν την λύση του προβλήματος...

Οπότε, θα εξετάσουμε το σύστημα ανά περιπτώσεις:

case i) $\lambda_1 = \lambda_2 = 0$: Αν $\lambda_1 = \lambda_2 = 0 \xrightarrow{(3)} -e^{-(x_1+x_2)} = 0$ (άτοπο, γιατί $e^z > 0$).

case ii) $\lambda_1 = 0$ και $\lambda_2 \neq 0$: Αν $\lambda_1 = 0$ και $\lambda_2 \neq 0 \xrightarrow{(4)} -e^{-(x_1+x_2)} = 0$ (άτοπο).

case iii) $\lambda_1 \neq 0$ και $\lambda_2 = 0$:

Για $\lambda_2 = 0$ από την (3) και (4) προκύπτει ότι:

$$-e^{-(x_1+x_2)} + \lambda_1 e^{x_1} = 0$$

$$-e^{-(x_1+x_2)} + \lambda_1 e^{x_2} = 0$$

$$\implies e^{x_1} = e^{x_2} \implies \boxed{x_1 = x_2}$$

Για $x_1 = x_2$ η (1) γίνεται:

$$\lambda_1(2e^{x_1} - 20) = 0 \xrightarrow{\lambda_1 \neq 0} 2e^{x_1} - 20 = 0 \implies e^{x_1} = 10 \implies \boxed{x_1 = \ln 10}$$

Όμως $x_1 = x_2$, άρα $x_2 = \ln 10$.

Για την τιμή του λ_1 έχουμε:

$$\begin{aligned} -e^{-(x_1+x_2)} + \lambda_1 e^{x_2} = 0 &\implies -e^{-2\ln 10} + \lambda_1 e^{\ln 10} = 0 \implies -e^{\ln(10^{-2})} + \lambda_1 \cdot 10 = 0 \\ &\implies -0.01 + 10\lambda_1 = 0 \implies \lambda_1 = 0.001 > 0 \quad (\deltaεκτή) \end{aligned}$$

case iv) $\lambda_1 \neq 0$ και $\lambda_2 \neq 0$: Αν $\lambda_1 \neq 0, \lambda_2 \neq 0$, τότε από την (1) έχουμε ότι $e^{x_1} + e^{x_2} - 20 = 0$ και από την (2) έχουμε ότι $x_1 = 0$. Από την (1) προκύπτει: $e^0 + e^{x_2} = 20 \implies 1 + e^{x_2} = 20 \implies \boxed{x_2 = \ln 19}$

Αντικαθιστώντας τα x_1, x_2 στην (4) έχουμε:

$$-e^{-(0+\ln 19)} + \lambda_1 e^{\ln 19} = 0 \implies -\frac{1}{19} + 19\lambda_1 = 0 \implies \lambda_1 = \frac{1}{19^2} > 0 \quad (\deltaεκτή)$$

Αντικαθιστώντας τα x_1, x_2, λ_1 στην (3) έχουμε:

$$-e^{-(0+\ln 19)} + \lambda_1 e^0 - \lambda_2 = 0 \implies -\frac{1}{19} + \frac{1}{19^2} - \lambda_2 = 0 \implies \lambda_2 = \frac{1-19}{19^2} = -\frac{18}{19^2} < 0 \quad (\alphaπορρίπτεται)$$

Συνεπώς η μόνη λύση είναι η $x_1^* = x_2^* = \ln 10$.

Άσκηση 3.6.17 -- Σελίδα 73

Ελαχιστοποίηση $\sum \frac{c_j}{x_j}$

Για το πρόβλημα: Βρείτε το $x \in \mathbb{R}^n$ που ελαχιστοποιεί την:

$$f(x) = \sum_{j=1}^n \frac{c_j}{x_j}$$

και ικανοποιεί:

$$\sum_{j=1}^n a_j x_j - b = 0$$

$x_j \geq 0, j = 1, \dots, n$ με $a_j > 0, c_j > 0, b > 0$.

α) Να γραφούν οι συνθήκες Karush-Kuhn-Tucker. β) Να βρεθεί το σημείο x^* που τις ικανοποιεί.

Λύση**Λύση:**

Θεωρώ: $h(x) = \sum_{j=1}^n a_j x_j - b$ και $g_j(x) = -x_j, j = 1, \dots, n$.

- $\nabla f(x) = \begin{bmatrix} -\frac{c_1}{x_1^2} \\ \vdots \\ -\frac{c_n}{x_n^2} \end{bmatrix}$
- $\nabla h(x) = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}, \nabla g_1(x) = \begin{bmatrix} -1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \dots, \nabla g_n(x) = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ -1 \end{bmatrix}$

KKT conditions:

1. $\lambda_j \geq 0$ για $j = 1, \dots, n$
2. $\lambda_j \cdot g_j(x) = 0 \implies \lambda_j(-x_j) = 0$
3. $\nabla f(x) + \mu \nabla h(x) + \sum_{j=1}^n \lambda_j \nabla g_j(x) = 0$

Κάνοντας αντικαταστάσεις στις παραπάνω σχέσεις έχουμε ότι:

- $\lambda_j \geq 0, j = 1, \dots, n \quad (1)$
- $-\lambda_j x_j = 0, j = 1, \dots, n \quad (2)$
- $\begin{bmatrix} -\frac{c_1}{x_1^2} \\ \vdots \\ -\frac{c_n}{x_n^2} \end{bmatrix} + \mu \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} + \begin{bmatrix} -\lambda_1 \\ \vdots \\ -\lambda_n \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} \quad (3)$

Από την (2) και το γεγονός ότι η λύση μας είναι κάποια $x_j \neq 0, j = 1, \dots, n$, καταλήγουμε στο συμπέρασμα ότι:

$$\lambda_j = 0, \quad j = 1, \dots, n \quad (4)$$

Αντικαθιστώντας το (4) στην σχέση (3) θα καταλήξουμε στα εξής:

$$-\frac{c_j}{x_j^2} + \mu a_j = 0 \implies x_j^2 = \frac{c_j}{\mu a_j} \implies x_j = \sqrt{\frac{c_j}{\mu a_j}}$$

Δηλαδή:

$$x_1 = \sqrt{\frac{c_1}{\mu a_1}}, \dots, x_n = \sqrt{\frac{c_n}{\mu a_n}}$$

Πώς βρίσκουμε το μ ; Παίρνοντας τον ισοτικό περιορισμό:

$$\begin{aligned} \sum_{j=1}^n a_j x_j - b = 0 &\implies \sum_{j=1}^n a_j \sqrt{\frac{c_j}{\mu a_j}} = b \\ &\implies \frac{1}{\sqrt{\mu}} \sum_{j=1}^n \sqrt{a_j c_j} = b \implies \sqrt{\mu} = \frac{\sum_{j=1}^n \sqrt{a_j c_j}}{b} \\ &\implies \boxed{\mu = \left(\frac{\sum_{j=1}^n \sqrt{a_j c_j}}{b} \right)^2} \end{aligned}$$

Αντικαθιστώντας το μ στο x_j , βρίσκουμε το σημείο x^* που ικανοποιεί τις συνθήκες KKT:

$$x_j^* = \frac{\sqrt{c_j/a_j}}{\sqrt{\mu}} \implies \boxed{x_j^* = \frac{b \sqrt{c_j/a_j}}{\sum_{k=1}^n \sqrt{a_k c_k}}}$$

Πρόβλημα Εφοπλιστικής Εταιρείας

Ελαχιστοποίηση Κόστους Μεταφοράς

Μια εφοπλιστική εταιρεία διαθέτει πλοία δύο τύπων Π_1 , Π_2 και θέλει να μεταφέρει 254 τόνους εμπορεύματος.

- Το Π_1 μπορεί να μεταφέρει 22 τόνους ανά ταξίδι.
- Το Π_2 μπορεί να μεταφέρει 12 τόνους ανά ταξίδι.

Για κάθε ταξίδι:

- Το Π_1 κοστίζει 12.000 και καίει 4.000 λίτρα καύσιμα.
- Το Π_2 κοστίζει 10.000 και καίει 900 λίτρα καύσιμα.

Η εταιρεία διαθέτει 30.800 λίτρα καύσιμα. Ποιο είναι το ελάχιστο κόστος μεταφοράς ώστε τα πλοία να ταξιδεύουν με γεμάτα αμπάρια (δηλαδή με γεμάτες αποθήκες);

Λύση

Λύση:

Πρέπει να βρούμε προφανώς την αντικειμενική συνάρτηση και τους περιορισμούς του προβλήματος...

Κόστος μεταφοράς:

$$f(x, y) = 12.000x + 10.000y$$

(Συνολικό Κόστος → αυτή είναι η αντικειμενική συνάρτηση).

Τώρα θα βρούμε τους περιορισμούς... Για να ταξιδεύουν τα πλοία με γεμάτες αποθήκες θα πρέπει:

$$22x + 12y = 254 \quad (1^{\circ} \text{ περιορισμός})$$

Για τα καύσιμα έχουμε ότι:

$$4.000x + 900y \leq 30.800$$

- Θεωρώ λοιπόν: $h(x, y) = 22x + 12y - 254$
- $g_1(x, y) = 4.000x + 900y - 30.800$

Προφανώς τα x, y δεν μπορούν να πάρουν αρνητικές τιμές! Έτσι, θα έχουμε δύο ακόμη περιορισμούς:

- $g_2(x, y) = -x$
- $g_3(x, y) = -y$

Κλίσεις (Gradients):

- $\nabla f(x, y) = \begin{bmatrix} 12.000 \\ 10.000 \end{bmatrix}$
- $\nabla h(x, y) = \begin{bmatrix} 22 \\ 12 \end{bmatrix}$
- $\nabla g_1(x, y) = \begin{bmatrix} 4.000 \\ 900 \end{bmatrix}$
- $\nabla g_2(x, y) = \begin{bmatrix} -1 \\ 0 \end{bmatrix}, \nabla g_3(x, y) = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$

KKT Conditions:

1. $\lambda_1, \lambda_2, \lambda_3 \geq 0$
2. $\lambda_1(4.000x + 900y - 30.800) = 0 \quad (1)$

3. $-\lambda_2 x = 0 \quad (2)$
4. $-\lambda_3 y = 0 \quad (3)$
5. $12.000 + 22\mu + 4.000\lambda_1 - \lambda_2 = 0 \quad (4)$
6. $10.000 + 12\mu + 900\lambda_1 - \lambda_3 = 0 \quad (5)$

Εξέταση περιπτώσεων:

case i: $x = 0$: Για $x = 0$ το y από την σχέση $22x + 12y = 254$ βγαίνει ίσο με: $y = 21,16$ (Απορρίπτεται, γιατί τα ταξίδια πρέπει να είναι ακέραιος αριθμός και επίσης για να είναι γεμάτα τα αμπάρια).

case ii: $y = 0$: Για $y = 0$ το x από την σχέση $22x + 12y = 254$ βγαίνει ίσο με: $x = 11,54$ (Απορρίπτεται).

case iii: $x, y > 0$: Για να ικανοποιούνται οι (2), (3) όταν $x, y > 0$ έχουμε ότι:

$$\lambda_2 = 0 \quad \text{και} \quad \lambda_3 = 0$$

Εφόσον έχω ότι $\lambda_1 \neq 0$, από την (1) προκύπτει ότι:

$$4.000x + 900y - 30.800 = 0 \quad (1)$$

Ομως, ισχύει επίσης ότι: $22x + 12y - 254 = 0 \quad (6)$.

Από τις (1) και (6) βρίσκω ότι: $x = 5$ και $y = 12$.

Αντικαθιστώντας τώρα τα λ_2, λ_3 στις (4), (5) έχουμε ότι:

- $12.000 + 22\mu + 4.000\lambda_1 = 0 \quad (7)$
- $10.000 + 12\mu + 900\lambda_1 = 0 \quad (8)$

Λύνοντας το σύστημα των εξισώσεων (7), (8) βρίσκουμε ότι:

$$\boxed{\mu = -1035,5 \quad \text{και} \quad \lambda_1 = 2,7}$$

Άσκηση 3.6.22 -- Σελίδα 75

Επαναλήψτε την Άσκηση 3.6.21

Επαναλάβετε την Άσκηση 3.6.21 για το πρόβλημα: Βρείτε τα x_1, x_2 που ελαχιστοποιούν την:

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

και ικανοποιούν:

$$x_1^2 - x_2 \leq 0$$

$$x_1 + x_2 \leq 2$$

Είναι το πρόβλημα κυρτό;

Λύση

Λύση:

- 1) Ορίζω ως: $g_1(x_1, x_2) = x_1^2 - x_2 \leq 0$ και $g_2(x_1, x_2) = x_1 + x_2 - 2 \leq 0$.
 2) Υπολογίζω τις κλίσεις των $f(x_1, x_2), g_1(x_1, x_2)$ και $g_2(x_1, x_2)$.

$$\nabla f(x_1, x_2) = \begin{bmatrix} 2x_1 - 4 \\ 2x_2 - 4 \end{bmatrix}, \quad \nabla g_1(x_1, x_2) = \begin{bmatrix} 2x_1 \\ -1 \end{bmatrix}, \quad \nabla g_2(x_1, x_2) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

3) Υπολογίζω τους Εσοιανούς πίνακες:

$$\nabla^2 f = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \implies D_1 = 2 > 0, D_2 = 4 > 0 \implies \text{Θετικά ορισμένος} \implies f \text{ γνήσια κυρτή.}$$

$$\nabla^2 g_1 = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix} \implies D_1 = 2 > 0, D_2 = 0 \implies \text{Θετικά ημιορισμένος} \implies g_1 \text{ κυρτή.}$$

$$\nabla^2 g_2 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \implies \text{Θετικά ημιορισμένος} \implies g_2 \text{ κυρτή.}$$

Άρα, το πρόβλημα είναι κυρτό.

4) KKT conditions:

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

$$\implies 2x_1 - 4 + 2\lambda_1 x_1 + \lambda_2 = 0 \quad (3) \implies 2x_2 - 4 - \lambda_1 + \lambda_2 = 0 \quad (4)$$

case i) $\lambda_1 = \lambda_2 = 0$: Από (3) $\rightarrow 2x_1 = 4 \rightarrow x_1 = 2$. Από (4) $\rightarrow 2x_2 = 4 \rightarrow x_2 = 2$. Ελέγχω τους περιορισμούς: $g_1(2, 2) = 2^2 - 2 = 2 \not\leq 0$ (Απορρίπτεται).

case ii) $\lambda_1 = 0, \lambda_2 \neq 0$: Από (2) $\rightarrow x_1 + x_2 = 2 \rightarrow x_1 = 2 - x_2$. Από (3), (4) με $\lambda_1 = 0 \rightarrow 2x_1 - 4 + \lambda_2 = 0$ και $2x_2 - 4 + \lambda_2 = 0 \implies x_1 = x_2$. Άρα $2x_1 = 2 \rightarrow x_1 = 1, x_2 = 1$. $\lambda_2 = 4 - 2(1) = 2 > 0$ (Δεκτή). $g_1(1, 1) = 1^2 - 1 = 0 \leq 0$ (Ικανοποιείται). Τελικά $x_1^* = 1, x_2^* = 1$.

case iii) $\lambda_1 \neq 0, \lambda_2 = 0$: Από (1) $\rightarrow x_2 = x_1^2$. Από (3) $\rightarrow 2x_1 - 4 + 2\lambda_1 x_1 = 0 \rightarrow \lambda_1 = \frac{4-2x_1}{2x_1} = \frac{2-x_1}{x_1}$. Από (4) $\rightarrow 2x_2 - 4 - \lambda_1 = 0 \rightarrow 2x_1^2 - 4 - \frac{2-x_1}{x_1} = 0 \rightarrow 2x_1^3 - 4x_1 - 2 + x_1 = 0 \rightarrow 2x_1^3 - 3x_1 - 2 = 0$. Λύνοντας προσεγγιστικά $x_1 \approx 1.47, x_2 \approx 2.17$. $g_2(1.47, 2.17) = 1.47 + 2.17 = 3.64 \not\leq 2$ (Απορρίπτεται).

case iv) $\lambda_1 \neq 0, \lambda_2 \neq 0$: $x_2 = x_1^2$ και $x_1 + x_2 = 2 \rightarrow x_1^2 + x_1 - 2 = 0 \rightarrow (x_1 + 2)(x_1 - 1) = 0$. Για $x_1 = 1, x_2 = 1 \rightarrow$ βρίσκουμε $\lambda_1 = 0$, άτοπο. Για $x_1 = -2, x_2 = 4 \rightarrow \lambda_1 < 0$, άτοπο.

Άσκηση 3.6.24 -- Σελίδα 75

Κατανομή Φορτίου Γεννητριών

Δύο ηλεκτρικές γεννήτριες είναι συνδεδεμένες παράλληλα για να ικανοποιούν ένα φορτίο ισχύος 60 W. Το κόστος παραγωγής κάθε γεννήτριας είναι συνάρτηση της ισχύος εξόδου $P_i, i = 1, 2$ ως εξής:

- 1^η γεννήτρια: $c_1 = 1 - P_1 + P_1^2$ Ευρώ/W
- 2^η γεννήτρια: $c_2 = 1 + 0,6P_2 + P_2^2$ Ευρώ/W

α) Να καταστρώσετε το πρόβλημα ελαχιστοποίησης κόστους. β) Να λύσετε το πρόβλημα αναλυτικά με συνθήκες KKT.

Λύση

Λύση:

Ορίζουμε την αντικειμενική συνάρτηση κόστους:

$$f(P_1, P_2) = (1 - P_1 + P_1^2) + (1 + 0,6P_2 + P_2^2) = P_1^2 + P_2^2 - P_1 + 0,6P_2 + 2$$

Περιορισμοί:

- $h(P_1, P_2) = P_1 + P_2 - 60 = 0$
- $g_1(P_1, P_2) = -P_1 \leq 0$
- $g_2(P_1, P_2) = -P_2 \leq 0$

Κυρτότητα: $\nabla^2 f = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \rightarrow$ Θετικά ορισμένος (γνήσια κυρτή). Οι περιορισμοί είναι γραμμικοί \rightarrow κυρτοί. Άρα το πρόβλημα είναι κυρτό.

KKT Conditions:

1. $\lambda_1, \lambda_2 \geq 0$
2. $-\lambda_1 P_1 = 0 \quad (1)$
3. $-\lambda_2 P_2 = 0 \quad (2)$
4. $2P_1 - 1 + \mu - \lambda_1 = 0 \quad (3)$
5. $2P_2 + 0,6 + \mu - \lambda_2 = 0 \quad (4)$

case 4) $P_1 \neq 0, P_2 \neq 0$: $\lambda_1 = \lambda_2 = 0$. Από (3), (4) $\rightarrow 2P_1 - 1 + \mu = 0$ και $2P_2 + 0,6 + \mu = 0 \implies 2P_1 - 1 = 2P_2 + 0,6 \rightarrow P_1 - P_2 = 0,8$. Επίσης $P_1 + P_2 = 60$. Λύνοντας: $2P_1 = 60,8 \rightarrow \boxed{P_1^* = 30,4, P_2^* = 29,6}$. $\mu = 1 - 2(30,4) = -59,8$. Η λύση είναι δεκτή.

Κεφάλαιο 5 -- Θεωρητική Εισαγωγή

Θεώρημα 5.1.1 (Strictly quasi-convex)

Έστω $f(x)$ μια αυστηρά σχεδόν-κυρτή συνάρτηση στο διάστημα $[a, b]$. Έστω δύο σημεία $x_1, x_2 \in [a, b]$ με $x_1 < x_2$.

1. Αν $f(x_1) < f(x_2)$, τότε $f(x) \geq f(x_1), \forall x \in (x_2, b]$. (Αποκλείουμε το $(x_2, b]$).
2. Αν $f(x_1) \geq f(x_2)$, τότε $f(x) \geq f(x_2), \forall x \in [a, x_1]$. (Αποκλείουμε το $[a, x_1)$).

Ορισμός αυστηρά σχεδόν-κυρτής

Μια συνάρτηση f λέγεται αυστηρά σχεδόν-κυρτή όταν:

$$f(\lambda x_1 + (1 - \lambda)x_2) < \max\{f(x_1), f(x_2)\}, \quad \forall x_1, x_2 \in S, \lambda \in (0, 1)$$

Διαισθητικά, μια τέτοια συνάρτηση είναι μονότονα φθίνουσα έως ένα σημείο και μετά μονότονα αύξουσα (unimodal).

Ερμηνεία με Compact Sets

Μια συνάρτηση είναι αυστηρά σχεδόν κυρτή αν για οποιαδήποτε οριζόντια γραμμή, το σύνολο των σημείων κάτω από τη γραμμή ($S_\alpha = \{x : f(x) \leq \alpha\}$) είναι κυρτό (στο \mathbb{R} αυτό σημαίνει διάστημα, δηλαδή compact set αν η συνάρτηση "ανεβαίνει" και στις δύο πλευρές).

Αλγόριθμοι Μονοδιάστατης Αναζήτησης

1) Μέθοδος της Διχοτόμου (Dichotomous Search)

Στη μέθοδο αυτή επιλέγουμε μια σταθερά ανοχής $l > 0$ και μια πολύ μικρή τιμή $\varepsilon > 0$. Στόχος είναι να βρούμε ένα τελικό διάστημα $[a_n, b_n]$ εύρους $< l$.

Βήματα: Σε κάθε επανάληψη k , επιλέγουμε δύο σημεία x_1, x_2 γύρω από το μέσο του διαστήματος $[a_k, b_k]$:

$$x_{1,k} = \frac{a_k + b_k}{2} - \varepsilon, \quad x_{2,k} = \frac{a_k + b_k}{2} + \varepsilon$$

- Αν $f(x_1) < f(x_2)$, τότε το ελάχιστο βρίσκεται στο $[a_k, x_2]$. (Αποκλείουμε το $(x_2, b_k]$).
- Αν $f(x_1) > f(x_2)$, τότε το ελάχιστο βρίσκεται στο $[x_1, b_k]$. (Αποκλείουμε το $[a_k, x_1)$).

Η διαδικασία επαναλαμβάνεται μέχρι $b_n - a_n < l$.

2) Μέθοδος του Χρυσού Τομέα (Golden Section Search)

Η μέθοδος αυτή είναι πιο αποδοτική γιατί σε κάθε επανάληψη απαιτεί μόνο έναν νέο υπολογισμό της συνάρτησης (επαναχρησιμοποιεί το ένα από τα δύο σημεία της προηγούμενης επανάληψης).

Επιλέγουμε τα σημεία x_1, x_2 σύμφωνα με τον χρυσό λόγο $\gamma \approx 0,618$:

$$x_{1,k} = a_k + (1 - \gamma)(b_k - a_k)$$

$$x_{2,k} = a_k + \gamma(b_k - a_k)$$

όπου $\gamma = \frac{\sqrt{5}-1}{2} \approx 0,618$.

Ιδιότητα reuse: Αν αποκλείσουμε το δεξί κομμάτι, το νέο x_2 της επόμενης επανάληψης θα συμπίπτει με το παλιό x_1 . Έτσι γλυτώνουμε υπολογισμούς.

3) Μέθοδος της Διχοτόμου με χρήση Παραγώγων

Η μέθοδος αυτή προϋποθέτει ότι γνωρίζουμε την παράγωγο της f . Πηγαίνουμε στο μέσο $x_k = \frac{a_k+b_k}{2}$ και υπολογίζουμε την κλίση $f'(x_k)$.

1. Αν $f'(x_k) = 0$, τότε το x_k είναι το ελάχιστο.
2. Αν $f'(x_k) > 0$, τότε το ελάχιστο βρίσκεται στο $[a_k, x_k]$. (Αποκλείουμε το $(x_k, b_k]$).
3. Αν $f'(x_k) < 0$, τότε το ελάχιστο βρίσκεται στο $[x_k, b_k]$. (Αποκλείουμε το $[a_k, x_k)$).

Η μέθοδος αυτή είναι υπολογιστικά "βαριά" αν η παράγωγος είναι δύσκολη στον υπολογισμό, αλλά συγκλίνει γρήγορα.

Μέθοδος Παρεμβολής Τετραγωνικών Συναρτήσεων

Αλγόριθμος: Τετραγωνική Παρεμβολή

Έστω $f(x)$ μια αυστηρά σχεδόν-κυρτή συνάρτηση στο $[0, \infty)$. Υποθέτουμε ότι διαθέτουμε τρία σημεία x_1, x_2, x_3 τέτοια ώστε:

$$0 \leq x_1 < x_2 < x_3 \quad \text{και} \quad f(x_1) \geq f(x_2), \quad f(x_2) \leq f(x_3)$$

(Τα σημεία αυτά σχηματίζουν ένα "σχήμα V" που περικλείει το ελάχιστο).

Εκτέλεση: 1. Παρεμβάλλουμε μια τετραγωνική καμπύλη (παραβολή) που περνάει από τα $(x_1, f(x_1)), (x_2, f(x_2)), (x_3, f(x_3))$. 2. Βρίσκουμε το ελάχιστο \hat{x} της παραβολής. 3. Επιλέγουμε μια νέα τριάδα σημείων:

- Αν $\hat{x} > x_2$:
 - Αν $f(\hat{x}) \geq f(x_2)$, η νέα τριάδα είναι $\{x_1, x_2, \hat{x}\}$.
 - Αν $f(\hat{x}) < f(x_2)$, η νέα τριάδα είναι $\{x_2, \hat{x}, x_3\}$.
 - Αν $\hat{x} < x_2$:
 - Αν $f(\hat{x}) \geq f(x_2)$, η νέα τριάδα είναι $\{\hat{x}, x_2, x_3\}$.
 - Αν $f(\hat{x}) < f(x_2)$, η νέα τριάδα είναι $\{x_1, \hat{x}, x_2\}$.
4. Επαναλαμβάνουμε μέχρι το διάστημα $[x_3 - x_1] < l$.

Ανισότητα Minkowski

Θεώρημα Minkowski

Έστω $x, y \in \mathbb{R}^n$ και $p \geq 1$. Τότε ισχύει:

$$\left(\sum_{i=1}^n |x_i + y_i|^p \right)^{1/p} \leq \left(\sum_{i=1}^n |x_i|^p \right)^{1/p} + \left(\sum_{i=1}^n |y_i|^p \right)^{1/p}$$

Απόδειξη

1. Για $p = 1$, η ανισότητα ανάγεται στην τριγωνική ανισότητα $|x_i + y_i| \leq |x_i| + |y_i|$. 2. Για $p > 1$, θεωρούμε τη συνάρτηση $\phi(t) = t^p$ για $t > 0$. Η δεύτερη παράγωγος είναι $\phi''(t) = p(p-1)t^{p-2} > 0$, άρα η ϕ είναι γνήσια κυρτή. Ορίζουμε:

$$\|x\|_p = \left(\sum |x_i|^p \right)^{1/p}, \quad \|y\|_p = \left(\sum |y_i|^p \right)^{1/p}$$

Επιλέγουμε $\lambda_1 = \frac{\|x\|_p}{\|x\|_p + \|y\|_p}$ και $\lambda_2 = \frac{\|y\|_p}{\|x\|_p + \|y\|_p}$, ώστε $\lambda_1 + \lambda_2 = 1$. Εφαρμόζουμε τον ορισμό της κυρτότητας για τα σημεία $a_i = \frac{|x_i|}{\|x\|_p}$ και $b_i = \frac{|y_i|}{\|y\|_p}$:

$$\phi(\lambda_1 a_i + \lambda_2 b_i) \leq \lambda_1 \phi(a_i) + \lambda_2 \phi(b_i)$$

$$\begin{aligned} \left(\frac{\|x\|_p}{\|x\|_p + \|y\|_p} \frac{|x_i|}{\|x\|_p} + \frac{\|y\|_p}{\|x\|_p + \|y\|_p} \frac{|y_i|}{\|y\|_p} \right)^p &\leq \lambda_1 \left(\frac{|x_i|}{\|x\|_p} \right)^p + \lambda_2 \left(\frac{|y_i|}{\|y\|_p} \right)^p \\ \left(\frac{|x_i| + |y_i|}{\|x\|_p + \|y\|_p} \right)^p &\leq \frac{\|x\|_p}{\|x\|_p + \|y\|_p} \frac{|x_i|^p}{\|x\|_p^p} + \frac{\|y\|_p}{\|x\|_p + \|y\|_p} \frac{|y_i|^p}{\|y\|_p^p} \end{aligned}$$

Αθροίζοντας ως προς i :

$$\frac{\sum(|x_i| + |y_i|)^p}{(\|x\|_p + \|y\|_p)^p} \leq \frac{1}{\|x\|_p + \|y\|_p} \left(\frac{\sum |x_i|^p}{\|x\|_p^{p-1}} + \frac{\sum |y_i|^p}{\|y\|_p^{p-1}} \right)$$

Επειδή $\sum |x_i|^p = \|x\|_p^p$:

$$\frac{\sum(|x_i| + |y_i|)^p}{(\|x\|_p + \|y\|_p)^p} \leq \frac{1}{\|x\|_p + \|y\|_p} (\|x\|_p + \|y\|_p) = 1$$

Άρα:

$$\sum |x_i + y_i|^p \leq \sum (|x_i| + |y_i|)^p \leq (\|x\|_p + \|y\|_p)^p$$

Υψώνοντας στην $1/p$:

$$\|x + y\|_p \leq \|x\|_p + \|y\|_p$$

4) Μέθοδος Fibonacci

Η μέθοδος αυτή είναι παρόμοια με τη μέθοδο του χρυσού τομέα, αλλά χρησιμοποιεί την ακολουθία Fibonacci F_n .

$$F_0 = 1, \quad F_1 = 1, \quad F_{n+1} = F_n + F_{n-1}$$

Τα σημεία επιλέγονται ως:

$$x_{1,k} = a_k + \frac{F_{n-k-1}}{F_{n-k+1}}(b_k - a_k), \quad x_{2,k} = a_k + \frac{F_{n-k}}{F_{n-k+1}}(b_k - a_k)$$

Στην τελευταία επανάληψη ($k = n$), τα x_1, x_2 συμπίπτουν στο μέσο. Για να γίνει η σύγκριση, επιλέγουμε ένα σημείο x'_2 σε απόσταση ε από το x_1 . Ο αριθμός των επαναλήψεων n προκαθορίζεται από την ακρίβεια l :

$$\frac{1}{F_n}(b_1 - a_1) < l$$

Άσκηση 5.5.6 (σελ. 172)

Έστω $\{x_k\}$ η ακολουθία που παράγεται από τη μέθοδο της μέγιστης καθόδου με σφάλμα:

$$x_{k+1} = x_k - \gamma_k(\nabla f(x_k) + e_k), \quad |e_k| \leq \delta, \quad \forall k$$

Δείξτε ότι αν $|\nabla f(x_k)| > \delta$, τότε υπάρχει $\gamma_1 > 0$ τέτοιο ώστε $f(x_{k+1}) < f(x_k)$ για $\gamma_k \in (0, \gamma_1]$.

Λύση**Λύση:**

Θεωρούμε τη συνάρτηση $\phi_k(\gamma) = f(x_k - \gamma(\nabla f(x_k) + e_k))$. Η παράγωγος της ϕ_k ως προς γ είναι:

$$\phi'_k(\gamma) = -\nabla f(x_k - \gamma(\nabla f(x_k) + e_k))^T \cdot (\nabla f(x_k) + e_k)$$

Για $\gamma = 0$:

$$\phi'_k(0) = -\nabla f(x_k)^T \cdot (\nabla f(x_k) + e_k) = -|\nabla f(x_k)|^2 - \nabla f(x_k)^T e_k$$

Χρησιμοποιώντας την ανισότητα Cauchy-Schwarz και το γεγονός ότι $|e_k| \leq \delta$:

$$\phi'_k(0) \leq -|\nabla f(x_k)|^2 + |\nabla f(x_k)| \cdot |e_k| \leq -|\nabla f(x_k)| \cdot (|\nabla f(x_k)| - \delta)$$

Επειδή από την εκφώνηση $|\nabla f(x_k)| > \delta$, έχουμε:

$$\phi'_k(0) < 0$$

Εφόσον η παράγωγος στο 0 είναι αρνητική, η συνάρτηση $\phi_k(\gamma)$ είναι φθίνουσα για μικρά γ . Άρα υπάρχει $\gamma_1 > 0$ τέτοιο ώστε για κάθε $\gamma \in (0, \gamma_1]$ να ισχύει:

$$\phi_k(\gamma) < \phi_k(0) \implies f(x_{k+1}) < f(x_k)$$

Άσκηση 5.5.8 (σελ. 175)

Θεωρήστε το πρόβλημα ελαχιστοποίησης της $f(x) = x^4 - 1, x \in \mathbb{R}$. Λύστε το κάνοντας χρήση της μέθοδου Newton. Βρείτε τη συνθήκη σύγκλισης και τον ρυθμό σύγκλισης. Επιλέξτε το βήμα γ_k για βέλτιστη ταχύτητα.

Λύση**Λύση:**

Για την $f(x) = x^4 - 1$:

$$f'(x) = 4x^3, \quad f''(x) = 12x^2$$

Η μέθοδος Newton είναι:

$$x_{k+1} = x_k - \gamma_k [f''(x_k)]^{-1} f'(x_k)$$

$$x_{k+1} = x_k - \gamma_k \frac{4x_k^3}{12x_k^2} = x_k - \gamma_k \frac{1}{3} x_k = x_k \left(1 - \frac{\gamma_k}{3}\right)$$

Για σύγκλιση πρέπει $|x_{k+1}/x_k| < 1$, άρα:

$$|1 - \gamma_k/3| < 1 \implies -1 < 1 - \gamma_k/3 < 1 \implies 0 < \gamma_k < 6$$

Ο ρυθμός σύγκλισης είναι γραμμικός με λόγο $(1 - \gamma_k/3)$. Για βέλτιστη επιλογή γ_k , θέλουμε $x_{k+1} = 0$ σε ένα βήμα:

$$1 - \gamma_k/3 = 0 \implies \gamma_k = 3$$

Με $\gamma_k = 3$ η μέθοδος συγκλίνει ακαριαία (σε 1 βήμα).

Άσκηση 5.6.17 (σελ. 192)

Έστω $f_1(x) = -x_1^2 - x_2^2$ και $f_2(x) = x_1^2 - x_2^3$. Δείξτε ότι ο αλγόριθμος Levenberg-Marquardt με $\gamma_k = 1$ δεν

συγκλίνει στο μέγιστο $(0, 0)$ της f_1 αν $x_0 \neq 0$. Δείξτε επίσης ότι δεν συγκλίνει στο σαμαροειδές σημείο $(0, 0)$ της f_2 αν $x_2(0) \neq 0$.

Λύση

Λύση:

Για την $f_1(x) = -x_1^2 - x_2^2$:

$$\nabla f_1 = [-2x_1, -2x_2]^T, \quad \nabla^2 f_1 = \text{diag}(-2, -2)$$

Ο Levenberg-Marquardt προσθέτει $\mu_k I$ ώστε ο πίνακας $(\nabla^2 f + \mu_k I)$ να είναι θετικά ορισμένος. Επιλέγουμε $\mu_k > 2$, έστω $\mu_k = 3$. Τότε ο πίνακας γίνεται I . Η διεύθυνση είναι $d_k = -[I]^{-1} \nabla f_1 = 2x_k$. Με $\gamma_k = 1$: $x_{k+1} = x_k + 2x_k = 3x_k$. Η ακολουθία $x_k = 3^k x_0$ αποκλίνει για $x_0 \neq 0$.

Για την $f_2(x) = x_1^2 - x_2^3$:

$$\nabla f_2 = [2x_1, -3x_2^2]^T, \quad \nabla^2 f_2 = \text{diag}(2, -6x_2)$$

Ο Εσσιανός στη διεύθυνση x_2 μπορεί να είναι αρνητικός. Προσθέτοντας μ_k ώστε ο πίνακας να είναι θετικά ορισμένος, η κίνηση στον άξονα x_2 θα είναι της μορφής $d_{2,k} = \frac{3x_2^2}{\mu_k - 6x_2}$. Αν $x_{2,0} \neq 0$, η ακολουθία στον άξονα x_2 δεν θα μηδενιστεί απαραίτητα (ανάλογα με το μ_k). Σημειώνεται ότι για $x_2 \rightarrow 0$, η παράγωγος $-3x_2^2$ μηδενίζεται πιο γρήγορα από τον παρονομαστή, αλλά ο αλγόριθμος Levenberg-Marquardt έχει σχεδιαστεί για ελαχιστοποίηση, ενώ εδώ το x_2 "θέλει" να αυξηθεί για να μειώσει την $f_2 = -x_2^3$.

Τελευταία 15 Θέματα (Pages 101-115)

Άσκηση 5.5.3 (Σελίδα 170)

Πρόβλημα: Θεωρείστε τη συνάρτηση $f(x) = \frac{1}{3}|x|^3$, $x \in \mathbb{R}^n$. Θέλουμε να την ελαχιστοποιήσουμε χρησιμοποιώντας τη μέθοδο της μέγιστης καθόδου με σταθερό βήμα $\gamma_k = \gamma, \forall k$. Βρείτε τη σχέση που πρέπει να ικανοποιούν οι αρχικές τιμές εκκίνησης του αλγορίθμου ώστε η μέθοδος να συγκλίνει.

Λύση

Λύση:

- Μέθοδος Μέγιστης Καθόδου: $x_{k+1} = x_k + \gamma_k d_k$, όπου $d_k = -\nabla f(x_k)$.
- Υπολογισμός Κλίσης: $\nabla f(x) = \frac{1}{3} \cdot 3|x|^2 \nabla(|x|) = |x|^2 \frac{x}{|x|} = x|x|$.
- Αναδρομή: $x_{k+1} = x_k - \gamma x_k |x_k| = x_k(1 - \gamma |x_k|)$.
- Μέτρο: $|x_{k+1}| = |x_k| |1 - \gamma |x_k||$.

Συνθήκη Σύγκλισης: $|1 - \gamma |x_k|| < 1$

- $-1 < 1 - \gamma |x_k| < 1 \implies -2 < -\gamma |x_k| < 0 \implies 2 > \gamma |x_k| > 0$.
- Άρα: $|x_k| < \frac{2}{\gamma}$. (Το γ πρέπει να επιλεγεί έτσι ώστε να ισχύει αυτή η σχέση).

Επιλογή γ_k (Optimal Step):

- Ελαχιστοποίηση της $f(x_k - \gamma \nabla f(x_k))$ ως προς γ .
- $\frac{df(x_k - \gamma \nabla f(x_k))}{d\gamma} = \nabla f(x_k - \gamma \nabla f(x_k))^T (-\nabla f(x_k)) = 0$.
- Καταλήγει στο: $1 - \gamma_k |x_k| = 0 \implies \gamma_k = \frac{1}{|x_k|}$.
- Για αυτή την επιλογή: $x_{k+1} = x_k - \frac{1}{|x_k|} x_k |x_k| = 0$. (Σύγκλιση σε 1 βήμα).

Άσκηση 5.5.9 (Σελίδα 176)

Πρόβλημα: Έστω ότι $|x - x^*| = \epsilon$ όπου x^* είναι τοπικό ελάχιστο. Βρείτε φράγματα στις ποσότητες $|f(x) - f(x^*)|$ και $|\nabla f(x) - \nabla f(x^*)|$ συναρτήσει του $\nabla^2 f(x)$.

Λύση

Λύση:

- Taylor 2ης τάξης γύρω από το x^* : $f(x) = f(x^*) + \nabla f(x^*)^T (x - x^*) + \frac{1}{2}(x - x^*)^T \nabla^2 f(x^*)(x - x^*)$.
- Επειδή x^* ελάχιστο: $\nabla f(x^*) = 0$.
- $|f(x) - f(x^*)| \leq \frac{1}{2}|x - x^*|^2 |\nabla^2 f(x^*)| = \frac{1}{2}\epsilon^2 |\nabla^2 f(x^*)|$.
- Taylor 1ης τάξης για ∇f : $\nabla f(x) = \nabla f(x^*) + \nabla^2 f(x^*)(x - x^*) \implies |\nabla f(x) - \nabla f(x^*)| \leq |\nabla^2 f(x^*)| \epsilon$.

Άσκηση 5.5.12 (Σελίδα 181)

Πρόβλημα: Θεωρείστε τη μέθοδο μέγιστης καθόδου $x_{k+1} = x_k - \gamma_k \nabla f(x_k)$. Υποθέτοντας ότι η f είναι κυρτή, δείξτε ότι $\forall y \in \mathbb{R}^n$: $|x_{k+1} - y|^2 \leq |x_k - y|^2 - 2\gamma_k(f(x_k) - f(y)) + (\gamma_k |\nabla f(x_k)|)^2$.

Ερώτημα β): Υποθέστε $\sum_{k=0}^{\infty} \gamma_k = \infty$ και $\gamma_k |\nabla f(x_k)|^2 \rightarrow 0$. Δείξτε ότι $\liminf_{k \rightarrow \infty} f(x_k) = \inf_{x \in \mathbb{R}^n} f(x)$.

Λύση**Λύση:****Απόδειξη (α):**

- $|x_{k+1} - y|^2 = |x_k - \gamma_k \nabla f(x_k) - y|^2 = |x_k - y|^2 + (\gamma_k |\nabla f(x_k)|)^2 - 2\gamma_k \nabla f(x_k)^T (x_k - y)$.
- Από κυρτότητα (Θ. 3.1.5): $f(x_k) + \nabla f(x_k)^T (y - x_k) \leq f(y) \implies \nabla f(x_k)^T (x_k - y) \geq f(x_k) - f(y)$.
- Αντικατάσταση στην 1η σχέση δίνει το ζητούμενο.

Λύση (β) - Εις άτοπον απαγωγή:

- Έστω $\liminf f(x_k) \neq \inf f(x)$.
- Άν $\liminf f(x_k) > \inf f(x)$, τότε $\exists y, \delta > 0$ τέτοια ώστε $f(y) < f(x_k) - \delta, \forall k \geq \bar{k}$.
- Από τη σχέση στο (α): $|x_{k+1} - y|^2 \leq |x_k - y|^2 + \gamma_k^2 |\nabla f(x_k)|^2 - 2\gamma_k \delta$.
- Άθροιση από \bar{k} έως ∞ : $0 \leq |x_{\bar{k}} - y|^2 - \delta \sum \gamma_k + \sum (\gamma_k^2 |\nabla f(x_k)|^2 \text{ terms})$.
- Επειδή $\sum \gamma_k = \infty$, καταλήγουμε σε $0 \leq -\infty$ (Άτοπο).
- Άρα η υπόθεση ήταν λανθασμένη.

Μέσος Όρος ως Πρόβλημα Βελτιστοποίησης

Πρόβλημα: Ο μέσος όρος $x^* = \frac{\sum y_i}{m}$ ελαχιστοποιεί τη συνάρτηση $f(x) = \sum_{i=1}^m \frac{1}{2}(y_i - x)^2$. Εφαρμόστε τη μέθοδο μέγιστης κλίσης.

Λύση**Λύση:**

- $\nabla f(x) = -\sum_{i=1}^m (y_i - x) = mx - \sum y_i$.
- $x_{k+1} = x_k - \gamma_k \nabla f(x_k) = x_k - \gamma_k (mx_k - \sum y_i) = (1 - \gamma_k m)x_k + \gamma_k \sum y_i$.
- Επιλέγοντας $\gamma_k = \frac{1}{m}$: $x_{k+1} = \frac{1}{m} \sum y_i = x^*$. (Σύγκλιση σε 1 βήμα).

Μέθοδος Σχεδόν Newton (Quasi-Newton Update)

Πρόβλημα: $f(x) = \frac{1}{2}x^T Q x - b^T x$. Ο πινακας Δ_k ενημερώνεται ως: $\Delta_{k+1} = \Delta_k + \frac{y_k y_k^T}{q_k^T y_k}$, όπου $y_k = p_k - \Delta_k q_k$.

1. Δείξτε $\Delta_{k+1} q_i = \rho_i, \forall i \leq k$.
2. Δείξτε $\Delta_n = Q^{-1}$.

Λύση**Λύση:****Απόδειξη (Επαγωγή):**

- Για $i = k$: $\Delta_{k+1} q_k = \Delta_k q_k + y_k = \Delta_k q_k + \rho_k - \Delta_k q_k = \rho_k$. (Ισχύει).
- Για $i < k$: $\Delta_{k+1} q_i = \Delta_k q_i + \frac{y_k (y_k^T q_i)}{q_k^T y_k} = \dots$ (χρήση ιδιοτήτων Q -συζυγών κατευθύνσεων).
- Τελικά $\Delta_n q_i = \rho_i \implies \Delta_n Q \rho_i = \rho_i \implies \Delta_n = Q^{-1}$.

Άσκηση 6.7.2 (Σελίδα 230) - Προβολή Κλίσης

Πρόβλημα: $f(x) = \frac{1}{2}x^2$ στο $X = [-1, 2]$ με σταθερό βήμα s .

Λύση**Λύση:**

- $\bar{x}_k = Pr_X\{x_k - s\nabla f(x_k)\} = Pr_X\{x_k(1 - s)\}.$
- Προβολή:
 - -1 , αν $x_k(1 - s) \leq -1$
 - 2 , αν $x_k(1 - s) \geq 2$
 - $x_k(1 - s)$, αν $-1 \leq x_k(1 - s) \leq 2$
- $x_{k+1} = x_k + \gamma_k(\bar{x}_k - x_k)$. Για $\gamma_k = 1$: $x_{k+1} = \dots$ (δίνει τις 3 περιπτώσεις).

Άσκηση 6.8.14 (Σελίδα 236)

Πρόβλημα: $f(x) = \frac{1}{2}x_1^2x_2$ με $x_1 \in [-1, 2]$ και $x_2 \in \mathbb{R}$.

Λύση**Λύση:**

- $\nabla f(x) = [x_1x_2, \frac{1}{2}x_1^2]^T$.
- $\bar{x}_k = Pr_X\{[x_{1,k} - sx_{1,k}x_{2,k}, x_{2,k} - \frac{s}{2}x_{1,k}^2]^T\}.$
- Επειδή x_2 ελεύθερο: $\bar{x}_{2,k} = x_{2,k} - \frac{s}{2}x_{1,k}^2$.
- Για το x_1 εφαρμόζεται η προβολή στο $[-1, 2]$.

Θέματα Εξετάσεων

Ιούνιος 2014 -- Θέματα & Λύσεις

Θέμα 1

Να βρεθεί το ορθογώνιο παραλληλεπίπεδο με τη μέγιστη δυνατή χωρητικότητα (όγκο) αν το άθροισμα των τριών ακμών του ισούται με a .

Λύση

Λύση:

Διατύπωση: Μεγιστοποίηση $V = xyz$ υπό τον περιορισμό $x + y + z = a$ και $x, y, z > 0$. Ισοδύναμα: $\min f(x, y, z) = -xyz$.

KKT: Lagrangian: $L = -xyz + \lambda(x + y + z - a)$. Εξισώσεις: 1. $-yz + \lambda = 0 \implies \lambda = yz$ 2. $-xz + \lambda = 0 \implies \lambda = xz$ 3. $-xy + \lambda = 0 \implies \lambda = xy$ 4. $x + y + z = a$

Από (1),(2): $yz = xz \implies y = x$ (για $z \neq 0$). Από (2),(3): $xz = xy \implies z = y$ (για $x \neq 0$). Άρα $x = y = z$. Αντικαθιστώντας στον περιορισμό:

$$3x = a \implies x = \frac{a}{3}$$

Άρα το βέλτιστο σχήμα είναι **κύβος** με πλευρά $a/3$.

Ιούλιος 2018 -- Θέματα & Λύσεις

Θέμα 1

Δίνονται οι παρατηρήσεις y_1, y_2, \dots, y_m . **α)** Να δειχθεί ότι για την ελαχιστοποίηση της συνάρτησης $f(x) = \frac{1}{2} \sum_{i=1}^m (y_i - x)^2$ με τη μέθοδο της μέγιστης καθόδου, υπάρχει βήμα γ_k τέτοιο ώστε ο αλγόριθμος να συγκλίνει στη βέλτιστη λύση \bar{x} σε μία επανάληψη. **β)** Να δειχθεί ότι ο αναδρομικός τύπος του αριθμητικού μέσου είναι $x_{k+1} = x_k + \frac{1}{k+1}(y_{k+1} - x_k)$.

Λύση

Λύση:

α) Η βέλτιστη λύση είναι ο αριθμητικός μέσος $\bar{x} = \frac{1}{m} \sum y_i$. Η παράγωγος είναι:

$$\nabla f(x) = \sum_{i=1}^m -(y_i - x) = \sum y_i - mx$$

Ο αλγόριθμος μέγιστης καθόδου:

$$x_{k+1} = x_k - \gamma_k \nabla f(x_k) = x_k + \gamma_k (\sum y_i - mx_k)$$

$$x_{k+1} = x_k(1 - \gamma_k m) + \gamma_k \sum y_i$$

Αν επιλέξουμε $\gamma_k = \frac{1}{m}$:

$$x_{k+1} = x_k(1 - 1) + \frac{1}{m} \sum y_i = \bar{x}$$

Άρα συγκλίνει σε 1 βήμα.

β) Για τον αναδρομικό τύπο του μέσου όρου x_k (των k παρατηρήσεων):

$$x_k = \frac{1}{k} \sum_{i=1}^k y_i \implies \sum_{i=1}^k y_i = kx_k$$

Για $k+1$:

$$x_{k+1} = \frac{1}{k+1} \sum_{i=1}^{k+1} y_i = \frac{1}{k+1} (\sum_{i=1}^k y_i + y_{k+1})$$

$$x_{k+1} = \frac{1}{k+1} (kx_k + y_{k+1}) = \frac{kx_k + x_k - x_k + y_{k+1}}{k+1} = \frac{(k+1)x_k + y_{k+1} - x_k}{k+1}$$

$$x_{k+1} = x_k + \frac{1}{k+1} (y_{k+1} - x_k)$$

Θέμα 2

Πρόβλημα Network Utility Maximization (NUM) σε δίκτυο με 3 κόμβους. Πηγή → Κόμβος 1 (x_{r1}) → Κόμβος 2 (x_{r2}). Χωρητικότητες C_1, C_2 . Συνάρτηση χρησιμότητας $U_i(x_i) = \ln x_i$ (Proportional Fairness). Να διατυπωθεί το πρόβλημα και να δειχθεί ότι η συνθήκη βελτιστότητας οδηγεί σε δίκαιη κατανομή.

Λύση

Λύση:

Διατύπωση:

$$\max \sum_i \ln x_i \iff \min \sum_i -\ln x_i$$

Περιορισμοί:

$$\sum_{i \in L_j} x_i \leq C_j \quad (\forall \text{ link } j)$$

Λόγω κυρτότητας των συναρτήσεων $-\ln x_i$, το πρόβλημα έχει μοναδικό ολικό ελάχιστο. Η συνθήκη 1ης

τάξης για κυρτή συνάρτηση:

$$\nabla f(x^*)^T(x - x^*) \geq 0$$

Για $f(x) = -\sum \ln x_i$, $\nabla f_i = -1/x_i$. Αυτό οδηγεί στο κριτήριο:

$$\sum_i \frac{x_i - x_i^*}{x_i^*} \leq 0$$

που αποτελεί τον ορισμό της αναλογικής δικαιοσύνης (Proportional Fairness). Αν προσπαθήσουμε να αυξήσουμε τον ρυθμό μιας ροής κατά ποσοστό p , το άθροισμα των ποσοστιαίων μειώσεων των άλλων ροών θα είναι μεγαλύτερο.

Φεβρουάριος 2020 -- Θέματα & Λύσεις

Θέμα 2

Δύο διυλιστήρια Δ_1, Δ_2 τροφοδοτούν δύο κέντρα κατανάλωσης K_1, K_2 . Υπάρχουν οι εξής συνδέσεις με τα αντίστοιχα κόστη: $\Delta_1 \rightarrow K_1$ (ροή x_1 , μοναδιαίο κόστος l_1) $\Delta_1 \rightarrow K_2$ (ροή x_2 , μοναδιαίο κόστος l_2) $\Delta_2 \rightarrow K_2$ (ροή x_4 , μοναδιαίο κόστος l_4) $K_1 \rightarrow K_2$ (ροή x_3 , μοναδιαίο κόστος l_3)

Δίνονται οι μέγιστες παραγωγές των διυλιστηρίων D_1, D_2 και οι απαιτήσεις των κέντρων K_1, K_2 . Να διατυπωθεί το πρόβλημα ελαχιστοποίησης του τετραγωνικού κόστους:

$$f(x) = l_1x_1^2 + l_2x_2^2 + l_3x_3^2 + l_4x_4^2$$

Λύση

Λύση:

A) Διατύπωση Προβλήματος Ελαχιστοποίηση κόστους:

$$\min f(x) = l_1x_1^2 + l_2x_2^2 + l_3x_3^2 + l_4x_4^2$$

Υπό τους περιορισμούς: 1. Ισοζύγιο Ροών (Νόμοι Kirchoff): Βάσει της διόρθωσης ότι το K_1 έχει μόνο εισορόές, η ροή x_3 έχει φορά $K_2 \rightarrow K_1$. Για το K_1 : Εισροές x_1, x_3 , Ζήτηση K_1 .

$$x_1 + x_3 = K_1 \implies x_1 + x_3 - K_1 = 0 \quad (h_1)$$

Για το K_2 : Εισροές x_2, x_4 , Εκροή x_3 , Ζήτηση K_2 .

$$x_2 + x_4 = K_2 + x_3 \implies x_2 - x_3 + x_4 - K_2 = 0 \quad (h_2)$$

2. Περιορισμοί Παραγωγής: Μέγιστη παραγωγή Δ_1 :

$$x_1 + x_2 \leq D_1 \implies x_1 + x_2 - D_1 \leq 0 \quad (g_1)$$

Μέγιστη παραγωγή Δ_2 :

$$x_4 \leq D_2 \implies x_4 - D_2 \leq 0 \quad (g_2)$$

3. Περιορισμοί Χωρητικότητας Αγωγών: Έστω C_i η μέγιστη χωρητικότητα του αγωγού i .

$$x_i \leq C_i \implies x_i - C_i \leq 0, \quad i = 1, 2, 3, 4 \quad (k_i)$$

4. Μη Αρνητικότητα:

$$x_i \geq 0 \iff -x_i \leq 0, \quad i = 1, 2, 3, 4$$

B) Συνθήκες KKT Σχηματίζουμε τη συνάρτηση Lagrangian. Σημείωση: Τα ν_i (νι) είναι οι πολλαπλασιαστές Lagrange που αντιστοιχούν στους περιορισμούς μη αρνητικότητας ($-x_i \leq 0$). Τα ρ_i είναι οι πολλαπλασιαστές για τις χωρητικότητες αγωγών.

$$L(x, \lambda, \mu, \rho, \nu) = \sum_{i=1}^4 l_i x_i^2 + \mu_1 h_1 + \mu_2 h_2 + \lambda_1 g_1 + \lambda_2 g_2 + \sum_{i=1}^4 \rho_i (x_i - C_i) - \sum_{i=1}^4 \nu_i x_i$$

Οι συνθήκες στασιμότητας ($\frac{\partial L}{\partial x_i} = 0$): 1. $\frac{\partial L}{\partial x_1} : 2l_1 x_1 + \mu_1 + \lambda_1 + \rho_1 - \nu_1 = 0$ 2. $\frac{\partial L}{\partial x_2} : 2l_2 x_2 + \mu_2 + \lambda_2 + \rho_2 - \nu_2 = 0$

3. $\frac{\partial L}{\partial x_3} : 2l_3 x_3 + \mu_1 - \mu_2 + \rho_3 - \nu_3 = 0$ 4. $\frac{\partial L}{\partial x_4} : 2l_4 x_4 + \mu_2 + \lambda_2 + \rho_4 - \nu_4 = 0$

Επιπλέον ισχύουν οι συνθήκες συμπληρωματικότητας:

- $\lambda_1(x_1 + x_2 - D_1) = 0, \quad \lambda_2(x_4 - D_2) = 0$
- $\rho_i(x_i - C_i) = 0, \quad \forall i$
- $\nu_i x_i = 0, \quad \forall i$
- $\lambda, \rho, \nu \geq 0$

Φεβρουάριος 2021 -- Θέματα & Λύσεις

Λύση

Λύση:

A) Δυνατότητα Χρήσης Μεθόδων Προβολής Για να χρησιμοποιηθούν μέθοδοι προβολής gradient πρέπει: 1. Η συνάρτηση f να είναι κυρτή. $H_f = \begin{bmatrix} 2 & 2 \\ 2 & 4 \end{bmatrix}$. Οι ιδιοτιμές είναι $\lambda = 3 \pm \sqrt{5} > 0$. Άρα H_f θετικά ορισμένος $\implies f$ αυστηρά κυρτή. 2. Το σύνολο των περιορισμών Ω να είναι κυρτό. Οι περιορισμοί ορίζουν ένα ορθογώνιο $[-1, 1] \times [-2, 0.5]$, το οποίο είναι κυρτό σύνολο. Άρα **μπορούν** να χρησιμοποιηθούν μέθοδοι προβολής.

B) Μέθοδος Newton με Προβολή Ο αλγόριθμος είναι:

$$x_{k+1} = P_\Omega(x_k - \gamma_k H^{-1} \nabla f(x_k))$$

Υπολογίζουμε $H^{-1} \nabla f$: $\nabla f = \begin{bmatrix} 2x + 2y \\ 2x + 4y \end{bmatrix}$. $H^{-1} = \frac{1}{4} \begin{bmatrix} 4 & -2 \\ -2 & 2 \end{bmatrix} = \begin{bmatrix} 1 & -0.5 \\ -0.5 & 0.5 \end{bmatrix}$. $H^{-1} \nabla f = \begin{bmatrix} 1 & -0.5 \\ -0.5 & 0.5 \end{bmatrix} \begin{bmatrix} 2x + 2y \\ 2x + 4y \end{bmatrix} = \begin{bmatrix} 2x + 2y - x - 2y \\ -x - y + x + 2y \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix}$. Άρα το βήμα Newton μας πάει στο: $x_{new} = x_k - 1 \cdot \begin{bmatrix} x_k \\ y_k \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$. Το σημείο $(0, 0)$ ανήκει στο σύνολο περιορισμών (εφικτό). Άρα η μέθοδος συγκλίνει στο $(0, 0)$ σε **μία επανάληψη**. Η προβολή είναι ταυτοτική αφού το σημείο είναι εντός.

Σεπτέμβριος 2021 -- Θέματα & Λύσεις

Θέμα 1 (6 μονάδες)

Να λυθεί αναλυτικά το πρόβλημα:

$$\min \left(- \sum_{i=1}^3 \ln x_i \right)$$

υπό τους περιορισμούς:

$$\begin{aligned} x_1 + x_2 &\leq 2 \\ x_1 + x_3 &\leq 1 \\ x_i &\geq 0, i = 1, 2, 3 \end{aligned}$$

Λύση

Λύση:

Το πρόβλημα είναι ισοδύναμο με τη μεγιστοποίηση του γινομένου $x_1 x_2 x_3$. Οι περιορισμοί $x_i \geq 0$ ικανοποιούνται αυστηρά ($x_i > 0$) λόγω του λογαρίθμου. Εφαρμόζουμε συνθήκες KKT με Lagrangian:

$$L = -\ln x_1 - \ln x_2 - \ln x_3 + \lambda_1(x_1 + x_2 - 2) + \lambda_2(x_1 + x_3 - 1)$$

Συνθήκες στασιμότητας: 1) $\frac{\partial L}{\partial x_1} = -\frac{1}{x_1} + \lambda_1 + \lambda_2 = 0 \implies x_1 = \frac{1}{\lambda_1 + \lambda_2}$ 2) $\frac{\partial L}{\partial x_2} = -\frac{1}{x_2} + \lambda_1 = 0 \implies x_2 = \frac{1}{\lambda_1}$
 3) $\frac{\partial L}{\partial x_3} = -\frac{1}{x_3} + \lambda_2 = 0 \implies x_3 = \frac{1}{\lambda_2}$

Υποθέτουμε ότι οι περιορισμοί είναι ενεργοί ($\lambda_1, \lambda_2 > 0$): $x_1 + x_2 = 2 \implies \frac{1}{\lambda_1 + \lambda_2} + \frac{1}{\lambda_1} = 2$ $x_1 + x_3 = 1 \implies \frac{1}{\lambda_1 + \lambda_2} + \frac{1}{\lambda_2} = 1$

Λύνοντας το σύστημα (από μεταγραφή χειρόγραφου): $3x_1^2 - 6x_1 + 2 = 0 \implies x_1 = \frac{6 \pm \sqrt{36-24}}{6} = 1 \pm \frac{\sqrt{3}}{3}$.

Επειδή $x_1 + x_3 = 1 \implies x_1 < 1$ (αφού $x_3 > 0$). Άρα $x_1 = 1 - \frac{\sqrt{3}}{3}$.

Τότε $x_3 = 1 - x_1 = \frac{\sqrt{3}}{3}$. Και $x_2 = 2 - x_1 = 1 + \frac{\sqrt{3}}{3}$.

Επαλήθευση: $x_1 \approx 0.42$, $x_2 \approx 1.57$, $x_3 \approx 0.57$. $\lambda_2 = 1/x_3 = \sqrt{3} > 0$. $\lambda_1 = 1/x_2 > 0$. Οι συνθήκες πληρούνται.

Θέμα 2 (4 μονάδες)

Για την ελαχιστοποίηση της συνάρτησης

$$f(x, y) = x^2 + y^2 - 2x - 4y + 5$$

ως προς x, y , εκκινώντας από οποιοδήποτε σημείο, έχετε στη διάθεσή σας τους αλγορίθμους: α) μέγιστης κλίσης, β) Newton, γ) Levenberg-Marquardt. Ποιον από τους τρεις θα επιλέξετε αν σας ενδιαφέρει η εύρεση του ελαχίστου να πραγματοποιηθεί μ' όσο το δυνατόν λιγότερα βήματα; Να αιτιολογήσετε την απάντησή σας.

Λύση

Λύση:

Η συνάρτηση $f(x, y)$ είναι τετραγωνική: $f(x, y) = (x-1)^2 + (y-2)^2$. Ο πίνακας Hessian είναι $H = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$.

Είναι σταθερός και θετικά ορισμένος (ιδιοτιμές 2, 2).

Επιλογή: Μέθοδος Newton. Αιτιολόγηση: Η μέθοδος Newton, όταν εφαρμόζεται σε τετραγωνική συνάρτηση με θετικά ορισμένο Hessian, συγκλίνει στο ακριβές ελάχιστο σε **μία μόνο επανάληψη**, ανεξαρτήτως του σημείου εκκίνησης. Η Levenberg-Marquardt θα συνέκλινε επίσης γρήγορα (ισοδυναμεί με Newton αν επιλεγεί παράμετρος $\mu = 0$ καθώς H θετικά ορισμένος), αλλά η Newton είναι η πιο άμεση απάντηση για "λιγότερα βήματα" (1 βήμα). Η μέγιστη κλίση θα χρειαστεί περισσότερα βήματα (εκτός αν ξεκινήσουμε από ειδικό σημείο).

Φεβρουάριος 2022 -- Θέματα & Λύσεις

Θέμα 1 (6 μονάδες)

Έστω ένα σύστημα μίας εισόδου x και μίας εξόδου y για το οποίο γνωρίζουμε ότι η μαθηματική σχέση που συνδέει την είσοδο με την έξοδο είναι:

$$y = a_0 + a_1x + a_2x^2$$

Για την είσοδο και την έξοδο έχουμε συλλέξει τρία ζεύγη μετρήσεων όπως στον πίνακα που ακολουθεί:

i	1	2	3
x(i)	1	2	3
y(i)	4	9	16

A) (2 μονάδες) Να προσδιοριστούν αναλυτικά οι σταθερές αλλά άγνωστες παράμετροι a_0, a_1, a_2 έτσι ώστε να ελαχιστοποιείται το άθροισμα των τετραγώνων των σφαλμάτων μεταξύ της εκτιμώμενης και της πραγματικής εξόδου.

B) (2 μονάδες) Να υλοποιήσετε τη μέθοδο της μέγιστης κλίσης με σταθερό βήμα για την αναδρομική εκτίμηση των a_0, a_1, a_2 .

Γ) (2 μονάδες) Πως πρέπει να επιλεγούν οι αρχικές εκτιμήσεις $\hat{a}_0(1), \hat{a}_1(1), \hat{a}_2(1)$ ώστε ο αλγόριθμος του ερωτήματος (B) να συγκλίνει στις βέλτιστες τιμές σε μία επανάληψη;

Λύση

Λύση:

A) Αναλυτική Επίλυση Το μοντέλο είναι $y = a_0 + a_1x + a_2x^2$. Θέλουμε να ελαχιστοποιήσουμε το άθροισμα σφαλμάτων:

$$E(a) = \sum_{i=1}^3 (y_i - (a_0 + a_1x_i + a_2x_i^2))^2$$

Αντικαθιστώντας τα δεδομένα $(1, 4), (2, 9), (3, 16)$:

$$E = (4 - (a_0 + a_1 + a_2))^2 + (9 - (a_0 + 2a_1 + 4a_2))^2 + (16 - (a_0 + 3a_1 + 9a_2))^2$$

Παρατηρούμε ότι $4 = 2^2, 9 = 3^2, 16 = 4^2$, δηλαδή $y = (x+1)^2 = x^2 + 2x + 1$. Άρα η ακριβής λύση είναι:

$$a_0 = 1, \quad a_1 = 2, \quad a_2 = 1$$

B) Μέθοδος Μέγιστης Κλίσης Η παράγωγος ∇E ως προς $a = [a_0, a_1, a_2]^T$ είναι:

$$\nabla E = -2 \sum (y_i - \hat{y}_i) \begin{bmatrix} 1 \\ x_i \\ x_i^2 \end{bmatrix}$$

Ο αλγόριθμος ανανέωσης είναι:

$$a_{k+1} = a_k - \gamma \nabla E(a_k)$$

Γ) Σύγκλιση σε μία επανάληψη Για να συγκλίνει ο αλγόριθμος μέγιστης κλίσης σε μία επανάληψη, πρέπει να ξεκινήσουμε από το βέλτιστο! Δηλαδή, αν επιλέξουμε αρχικές τιμές:

$$\hat{a}_0(1) = 1, \quad \hat{a}_1(1) = 2, \quad \hat{a}_2(1) = 1$$

τότε $\nabla E(a_{start}) = 0$, και ο αλγόριθμος τερματίζει αμέσως.

Θέμα 2 (4 μονάδες)

A) (1 μονάδα) Να μελετηθεί ως προς την κυρτότητα η συνάρτηση:

$$f(x) = \begin{cases} x^2, & 0 \leq x \leq 1 \\ 2x - 1, & 1 < x \end{cases}$$

B) (1 μονάδα) Δίνονται οι περιορισμοί:

$$\begin{aligned} x_1 &\geq 0 \\ x_2 &\geq 0 \\ -x_1 + (x_2 - 1)^2 &\leq 0 \end{aligned}$$

Είναι το $(x_1^*, x_2^*) = (0, 1)$ εφικτό και κανονικό; Να αιτιολογήσετε την απάντησή σας.

Γ) (1 μονάδα) Δίνεται το πρόβλημα:

$$\min(x_1^2 + x_2^2 - 4x_1 - 4x_2)$$

υπό τους περιορισμούς:

$$\begin{aligned} -x_1^2 + x_2 &\leq 0 \\ x_1 + x_2 &\geq 0 \end{aligned}$$

α) Να γραφούν οι συνθήκες Karush-Kuhn-Tucker.

β) Είναι οι συνθήκες αυτές ικανές και αναγκαίες για το δοσμένο πρόβλημα; Να αιτιολογήσετε την απάντησή σας.

Δ) (1 μονάδα) Στους γενετικούς αλγορίθμους το μέγεθος του πληθυσμού είναι κρίσιμη λειτουργική παράμετρος. Τι θα συμβεί πιθανότητα στη λειτουργία του αλγορίθμου αν το μέγεθος του πληθυσμού αυξηθεί και τι αν μειωθεί;

Λύση**Λύση:**

A) (1 μονάδα) Για $0 \leq x \leq 1$: $f(x) = x^2$, $f''(x) = 2 > 0$ (κυρτή). Για $x > 1$: $f(x) = 2x - 1$, $f''(x) = 0$ (κυρτή). Ελέγχουμε την κυρτότητα στο σημείο αλλαγής $x = 1$ και γενικά τον ορισμό: $f(\lambda x_1 + (1 - \lambda)x_2) \leq \lambda f(x_1) + (1 - \lambda)f(x_2)$. Αποδεικνύεται ότι η συνάρτηση είναι **κυρτή** (convex) σε όλο το πεδίο ορισμού, καθώς η εφαπτομένη (παράγωγος) είναι αύξουσα ($2x$ για $x \leq 1$ που φτάνει το 2, και 2 για $x > 1$). Η κλίση δεν μειώνεται.

B) (1 μονάδα) Σημείο $(0, 1)$. Περιορισμοί: $g_1 = -x_1 \leq 0$, $g_2 = -x_2 \leq 0$, $g_3 = -x_1 + (x_2 - 1)^2 \leq 0$. Έλεγχος εφικτότητας: $0 \geq 0$, $1 \geq 0$, $0 + (1 - 1)^2 = 0 \leq 0$. **Εφικτό.** Ενεργοί περιορισμοί: g_1 και g_3 . Κλίσεις ενεργών περιορισμών: $\nabla g_1 = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$, $\nabla g_3 = \begin{bmatrix} -1 \\ 2(x_2 - 1) \end{bmatrix}_{(0,1)} = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$. Παρατηρούμε ότι $\nabla g_1 = \nabla g_3$. Τα διανύσματα είναι γραμμικά εξαρτημένα. Άρα το σημείο **δεν είναι κανονικό**.

Γ) (1 μονάδα) Πρόβλημα: $\min(x_1^2 + x_2^2 - 4x_1 - 4x_2)$ με $g_1(x) = x_1^2 - x_2 \leq 0$ και $g_2(x) = -x_1 - x_2 \leq 0$. (Διορθώθηκε βάσει σημειώσεων για να βγάζει νόημα η κυρτότητα). **α) Συνθήκες KKT:** $L = f + \lambda_1 g_1 + \lambda_2 g_2$.

1. Στασιμότητα: $\nabla f + \lambda_1 \nabla g_1 + \lambda_2 \nabla g_2 = 0$. $\begin{bmatrix} 2x_1 - 4 \\ 2x_2 - 4 \end{bmatrix} + \lambda_1 \begin{bmatrix} 2x_1 \\ -1 \end{bmatrix} + \lambda_2 \begin{bmatrix} -1 \\ -1 \end{bmatrix} = \vec{0}$ 2. Εφικτότητα: $g_1(x) \leq 0$, $g_2(x) \leq 0$. 3. Συμπληρωματικότητα: $\lambda_1 g_1(x) = 0$, $\lambda_2 g_2(x) = 0$. 4. Μη αρνητικότητα: $\lambda_1, \lambda_2 \geq 0$.

β) Ικανές και Αναγκαίες; Είναι **ικανές και αναγκαίες** διότι: 1. Η f είναι κυρτή (θετικά ορισμένη τετραγωνική). 2. Οι περιορισμοί g_1, g_2 ορίζουν κυρτό σύνολο (αν θεωρήσουμε την τυπική μορφή που δίνει κυρτότητα).

Δ) (1 μονάδα) Αν το μέγεθος του πληθυσμού **αυξηθεί**: Ο αλγόριθμος εξερευνά μεγαλύτερο μέρος του χώρου αναζήτησης (καλύτερο exploration), μειώνοντας την πιθανότητα εγκλωβισμού σε τοπικά ελάχιστα, αλλά αυξάνεται ο υπολογιστικός χρόνος ανά γενιά (πιο αργή σύγκλιση σε χρόνο). Αν το μέγεθος του πληθυσμού **μειωθεί**: Ο αλγόριθμος τρέχει πιο γρήγορα ανά γενιά, αλλά υπάρχει μεγαλύτερος κίνδυνος πρόωρης σύγκλισης σε τοπικό ελάχιστο (χειρότερο exploration).

Ιούνιος 2022 -- Θέματα & Λύσεις

Θέμα 1 (6 μονάδες)

Δίνεται η συνάρτηση $f(x, y) = x^2 + xy + y^2 - x + y$, $x, y \in \mathbb{R}$.

Α) (3 μονάδες) Να υπολογιστούν αναλυτικά τα βέλτιστα σημεία και να σχεδιαστούν ποιοτικά οι ισοσταθμικές καμπύλες.

Β) (3 μονάδες) Για τον αριθμητικό υπολογισμό των βέλτιστων σημείων του προηγούμενου ερωτήματος χρησιμοποιούμε: α) τη μέθοδο της μέγιστης κλίσης, β) τη μέθοδο Newton, γ) τη μέθοδο εσωτερικής βελτιστοποίησης. Για σημείο εκκίνησης το $(x, y) = (-3, 1)$ σχεδιάστε ποιοτικά, αλλά με αιτιολόγηση, πάνω στις ισοσταθμικές καμπύλες του προηγούμενου ερωτήματος, την εξέλιξη των αλγορίθμων. Να δώσετε διαφορετικό διάγραμμα για κάθε αλγόριθμο.

Λύση

Λύση:

Α) Αναλυτικός Υπολογισμός Υπολογίζουμε την κλίση της f :

$$\nabla f = \begin{bmatrix} 2x + y - 1 \\ x + 2y + 1 \end{bmatrix}$$

Θέτουμε $\nabla f = 0$:

$$\begin{cases} 2x + y = 1 \implies y = 1 - 2x \\ x + 2y = -1 \end{cases}$$

Αντικαθιστώντας το y στη δεύτερη:

$$x + 2(1 - 2x) = -1 \implies x + 2 - 4x = -1 \implies -3x = -3 \implies x = 1$$

Άρα $y = 1 - 2(1) = -1$. Το κρίσιμο σημείο είναι το $(1, -1)$.

Για να χαρακτηρίσουμε το σημείο, υπολογίζουμε τον πίνακα Hessian:

$$H = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

Η ορίζουσα είναι $\det(H) = 4 - 1 = 3 > 0$ και η πρώτη ελάσσονα $2 > 0$. Άρα ο πίνακας είναι θετικά ορισμένος και το σημείο $(1, -1)$ είναι **ολικό ελάχιστο**.

Ισοσταθμικές καμπύλες: Οι καμπύλες $f(x, y) = c$ είναι ελλείψεις με κέντρο το $(1, -1)$ και κύριους άξονες στραμμένους κατά 45 μοίρες (λόγω του όρου xy).

Β) Εξέλιξη Αλγορίθμων από $(-3, 1)$ Το σημείο εκκίνησης είναι το $A(-3, 1)$. Ο στόχος είναι το $O(1, -1)$.

- Μέθοδος Μέγιστης Κλίσης:** Η κατεύθυνση είναι $d_k = -\nabla f(x_k)$. Οι διαδοχικές κατευθύνσεις είναι κάθετες μεταξύ τους (Zig-zag). Αφού οι ισοσταθμικές είναι ελλείψεις (όχι κύκλοι), η μέθοδος θα συγκλίνει αργά με πολλές επαναλήψεις, κάνοντας ζιγκ-ζαγκ προς το ελάχιστο.
- Μέθοδος Newton:** Δεδομένου ότι η συνάρτηση είναι αυστηρά τετραγωνική (quadratic), η μέθοδος Newton συγκλίνει σε **μία επανάληψη**. Η τροχιά θα είναι ένα ευθύγραμμο τμήμα απευθείας από το $(-3, 1)$ στο $(1, -1)$.
- Μέθοδος Συζυγών Κλίσεων (Conjugate Gradients):** Για τετραγωνική συνάρτηση n μεταβλητών, συγκλίνει σε το πολύ n βήματα. Εδώ $n = 2$. Άρα θα φτάσει στο ελάχιστο σε το πολύ 2 βήματα. Το πρώτο βήμα είναι ίδιο με τη μέγιστη κλίση, και το δεύτερο διορθώνει την κατεύθυνση για να πέσει κατευθείαν στο κέντρο.

Θέμα 2 (4 μονάδες)

Έστω ο κύκλος κέντρου $(0, 0)$ και ακτίνας $\rho > 0$. Χρησιμοποιώντας τη μέθοδο ποινής να αποδείξετε πως απ' όλα τα εγγεγραμμένα στον κύκλο ορθογώνια παραλληλόγραμμα το τετράγωνο έχει το μεγαλύτερο εμβαδό. Ποιο είναι το μήκος της πλευράς αυτού του τετραγώνου;

Λύση**Λύση:**

Λύση με Μέθοδο Ποινής Θέλουμε να μεγιστοποιήσουμε το εμβαδόν $E = ab$ (όπου a, b οι πλευρές), υπό τον περιορισμό ότι το ορθογώνιο είναι εγγεγραμμένο στον κύκλο. Η διαγώνιος του ορθογωνίου είναι ίση με τη διάμετρο του κύκλου, άρα: $a^2 + b^2 = (2\rho)^2 = 4\rho^2$. Το πρόβλημα μετασχηματίζεται σε:

$$\min(-ab) \quad \text{υπό τον περιορισμό} \quad h(a, b) = a^2 + b^2 - 4\rho^2 = 0$$

Ορίζουμε τη συνάρτηση ποινής:

$$P(a, b, r_k) = -ab + \frac{r_k}{2}(a^2 + b^2 - 4\rho^2)^2$$

Υπολογίζουμε τις μερικές παραγώγους και τις εξισώνουμε με το μηδέν:

$$\frac{\partial P}{\partial a} = -b + r_k(a^2 + b^2 - 4\rho^2) \cdot 2a = 0$$

$$\frac{\partial P}{\partial b} = -a + r_k(a^2 + b^2 - 4\rho^2) \cdot 2b = 0$$

Από την πρώτη εξίσωση: $2ar_k(a^2 + b^2 - 4\rho^2) = b$. Από τη δεύτερη εξίσωση: $2br_k(a^2 + b^2 - 4\rho^2) = a$.

Διαιρώντας κατά μέλη (υποθέτοντας $a, b \neq 0$):

$$\frac{a}{b} = \frac{b}{a} \implies a^2 = b^2 \implies a = b$$

Άρα το βέλτιστο ορθογώνιο είναι **τετράγωνο**.

Για να βρούμε την πλευρά, αντικαθιστούμε στον περιορισμό (καθώς $r_k \rightarrow \infty$, ο περιορισμός ικανοποιείται):

$$a^2 + b^2 = 4\rho^2 \implies 2a^2 = 4\rho^2 \implies a^2 = 2\rho^2 \implies a = \sqrt{2}\rho$$

Επομένως, το μήκος της πλευράς είναι $\boxed{a = \rho\sqrt{2}}$.

Σεπτέμβριος 2022 -- Θέματα & Λύσεις

Θέμα 1 (4 μονάδες)

Μια αεροπορική εταιρία επιτρέπει στους επιβάτες της τη μεταφορά μιας μόνο αποσκευής σχήματος ορθογωνίου παραλληλεπιπέδου της οποίας οι διαστάσεις x, y, z πρέπει να ικανοποιούν:

$$x + y + z \leq a$$

όπου $a > 0$ δεδομένη σταθερά. Να υπολογιστούν αναλυτικά οι βέλτιστες διαστάσεις μιας αποσκευής που πληροί την παραπάνω συνθήκη και έχει μέγιστο όγκο.

Λύση

Λύση:

Θέλουμε να μεγιστοποιήσουμε τον όγκο $V(x, y, z) = xyz$ υπό τον περιορισμό $x + y + z \leq a$ και $x, y, z > 0$.

Ανισότητα Αριθμητικού-Γεωμετρικού Μέσου (AM-GM)

Για θετικούς αριθμούς x_1, x_2, \dots, x_n ισχύει:

$$\frac{x_1 + x_2 + \dots + x_n}{n} \geq \sqrt[n]{x_1 x_2 \dots x_n}$$

Η ισότητα ισχύει αν και μόνο αν $x_1 = x_2 = \dots = x_n$.

Εφαρμόζουμε την ανισότητα AM-GM για τους τρεις θετικούς αριθμούς x, y, z :

$$\frac{x + y + z}{3} \geq \sqrt[3]{xyz}$$

Υψώνουμε στην τρίτη δύναμη:

$$\left(\frac{x + y + z}{3} \right)^3 \geq xyz = V$$

Από τον περιορισμό έχουμε $x + y + z \leq a$, οπότε:

$$V \leq \left(\frac{a}{3} \right)^3 = \frac{a^3}{27}$$

Για να μεγιστοποιηθεί ο όγκος, πρέπει να ισχύει η ισότητα στην ανισότητα AM-GM. Αυτό συμβαίνει όταν:

$$x = y = z$$

Αντικαθιστώντας στον ενεργό περιορισμό $x + y + z = a$:

$$3x = a \implies x = \frac{a}{3}$$

Άρα, οι βέλτιστες διαστάσεις είναι:

$$x = y = z = \frac{a}{3}$$

και ο μέγιστος όγκος είναι $V_{max} = \frac{a^3}{27}$.

Θέμα 2 (3 μονάδες)

Για την ελαχιστοποίηση της συνάρτησης

$$f(x, y) = x^2 + y^2 - 2x - 4y + 5$$

ως προς x, y εκκινώντας από οποιοδήποτε σημείο έχετε στη διάθεσή σας τους αλγορίθμους: α) μέγιστης κλίσης, β) Newton, γ) Levenberg-Marquardt. Ποιον από τους τρεις θα επιλέξετε αν σας ενδιαφέρει η εύρεση του ελαχίστου να πραγματοποιηθεί μ' όσο το δυνατό λιγότερες επαναλήψεις, εκτελώντας σε

κάθε επανάληψη τις λιγότερες δυνατές πράξεις; Να αιτιολογήσετε την απάντησή σας.

Λύση

Λύση:

Η δοσμένη συνάρτηση είναι:

$$f(x, y) = x^2 + y^2 - 2x - 4y + 5$$

Παρατηρούμε ότι μπορεί να γραφτεί στη μορφή (συμπλήρωση τετραγώνου):

$$f(x, y) = (x^2 - 2x + 1) + (y^2 - 4y + 4) = (x - 1)^2 + (y - 2)^2$$

Πρόκειται για μια αυστηρά κυρτή τετραγωνική συνάρτηση.

Υπολογίζουμε την κλίση (∇f) και τον Εσσιανό πίνακα ($\nabla^2 f$):

$$\nabla f(x, y) = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix} = \begin{bmatrix} 2x - 2 \\ 2y - 4 \end{bmatrix}$$

$$H(x, y) = \nabla^2 f(x, y) = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

Παρατηρούμε ότι ο πίνακας H είναι σταθερός, διαγώνιος και θετικά ορισμένος (ιδιοτιμές $\lambda_1 = \lambda_2 = 2 > 0$).

Θεώρημα 5.2.4 (Σύγκλιση Newton)

Αν η $f(x)$ είναι τετραγωνική συνάρτηση με θετικά ορισμένο Εσσιανό πίνακα, τότε η μέθοδος του Newton συγκλίνει στο ακριβές ελάχιστο x^* σε **μία** μόνο επανάληψη, ξεκινώντας από οποιοδήποτε αρχικό σημείο x_0 .

Συνεπώς:

- Μέθοδος Μέγιστης Κλίσης:** Έχει γραμμικό ρυθμό σύγκλισης και ενδέχεται να χρειαστεί πολλές επαναλήψεις, ειδικά αν οι ισοσταθμικές καμπύλες είναι επιμήκεις (αν και εδώ είναι κύκλοι, οπότε θα σύγλινε γρήγορα, αλλά όχι σε 1 βήμα γενικά).
- Μέθοδος Newton:** Θα βρει το ελάχιστο ακριβώς σε **1 επανάληψη** ανεξαρτήτως αρχικού σημείου.
- Μέθοδος Levenberg-Marquardt:** Είναι μια παραλλαγή της μεθόδου Newton ($H + \mu I$). Αν επιλεχθεί $\mu = 0$, ταυτίζεται με τη Newton. Αν $\mu > 0$, θα χρειαστεί περισσότερες επαναλήψεις.

Επιλογή: Θα επιλέξουμε τη **μέθοδο Newton**, διότι για τετραγωνική συνάρτηση εντοπίζει το ολικό ελάχιστο σε **μία μόλις επανάληψη**, ελαχιστοποιώντας έτσι το συνολικό υπολογιστικό κόστος.

Θέμα 3 (3 μονάδες)

Θέλουμε να ελαχιστοποιήσουμε τη συνάρτηση

$$f(x, y) = x^2 + 2y^2 + 2xy$$

ως προς x, y με περιορισμούς:

$$\begin{aligned} -1 &\leq x \leq 1 \\ -2 &\leq y \leq \frac{1}{2} \end{aligned}$$

Μπορούν να χρησιμοποιηθούν μέθοδοι προβολής για να λυθεί το πρόβλημα; Να αιτιολογηθεί η απάντησή σας.

Λύση**Λύση:**

Για να εφαρμοστεί η μέθοδος της Προβολής της Κλίσης (Gradient Projection) και να εγγυηθούμε τη σύγκλισή της στο ολικό ελάχιστο, πρέπει να πληρούνται δύο προϋποθέσεις:

1. Η αντικειμενική συνάρτηση $f(x)$ να είναι κυρτή.
2. Το σύνολο των περιορισμών Ω να είναι κυρτό σύνολο.

Βήμα 1: Έλεγχος κυρτότητας της $f(x, y)$ Υπολογίζουμε τον Εσσιανό πίνακα της $f(x, y) = x^2 + 2y^2 + 2xy$:

$$\nabla f = \begin{bmatrix} 2x + 2y \\ 4y + 2x \end{bmatrix}, \quad H = \nabla^2 f = \begin{bmatrix} 2 & 2 \\ 2 & 4 \end{bmatrix}$$

Βρίσκουμε τις κύριες ελάσσονες ορίζουσες ή τις ιδιοτιμές. Ιδιοτιμές ($\det(H - \lambda I) = 0$):

$$(2 - \lambda)(4 - \lambda) - 4 = 0 \implies \lambda^2 - 6\lambda + 8 - 4 = 0 \implies \lambda^2 - 6\lambda + 4 = 0$$

$$\Delta = 36 - 16 = 20 \implies \lambda_{1,2} = \frac{6 \pm \sqrt{20}}{2} = 3 \pm \sqrt{5}$$

Και οι δύο ιδιοτιμές είναι θετικές ($3 + \sqrt{5} > 0$ και $3 - \sqrt{5} > 0$). Άρα, ο πίνακας H είναι **θετικά ορισμένος**, οπότε η f είναι **αυστηρά κυρτή** (και άρα κυρτή) στο \mathbb{R}^2 .

Βήμα 2: Έλεγχος κυρτότητας του Ω Οι περιορισμοί είναι:

$$\Omega = \{(x, y) \in \mathbb{R}^2 \mid -1 \leq x \leq 1, -2 \leq y \leq \frac{1}{2}\}$$

Το σύνολο Ω είναι ένα ορθογώνιο και ως τομή ημιεπιπέδων (γραμμικοί περιορισμοί), είναι **κυρτό σύνολο**.

Συμπέρασμα: Ναι, μπορούν να χρησιμοποιηθούν μέθοδοι προβολής (όπως η μέθοδος Προβολής της Κλίσης) διότι το πρόβλημα είναι κυρτό (ελαχιστοποίηση κυρτής συνάρτησης σε κυρτό σύνολο). Επιπλέον, επειδή οι περιορισμοί είναι απλά όρια (box constraints), η προβολή ενός σημείου (x, y) στο Ω είναι υπολογιστικά πολύ εύκολη:

$$P_\Omega(x, y) = (\min(1, \max(-1, x)), \min(1/2, \max(-2, y)))$$

Φεβρουάριος 2023 -- Θέματα & Λύσεις

Θέμα 1 (5 μονάδες)

A) (2.5 μονάδες)

Έστω ο κύκλος κέντρου $(0, 0)$ και ακτίνας $\rho > 0$. Με χρήση της ανισότητας Αριθμητικού Γεωμετρικού Μέσου να δείξετε ότι, από όλα τα εγγεγραμμένα στον κύκλο ορθογώνια παραλληλόγραμμα, το τετράγωνο έχει το μέγιστο εμβαδόν. Προσδιορίστε και το μέγιστο εμβαδόν και την πλευρά του τετραγώνου.

B) (2.5 μονάδες) Να λυθεί το πρόβλημα του Θέματος 1A) με την μέθοδο ποινής.

Λύση

Λύση:

Κύκλος κέντρου $(0, 0)$ ακτίνας ρ . Εγγεγραμμένο ορθογώνιο με πλευρές a, b .

Λύση με AM-GM: Η διαγώνιος του ορθογώνιου ισούται με τη διάμετρο: $a^2 + b^2 = (2\rho)^2 = 4\rho^2$. Εμβαδόν: $E = ab$.

Από AM-GM: $\frac{a^2+b^2}{2} \geq \sqrt{a^2b^2} = ab \cdot \frac{4\rho^2}{2} \geq ab \implies E \leq 2\rho^2$.

Η ισότητα επιτυγχάνεται όταν $a^2 = b^2 \implies a = b$ (τετράγωνο). Για τετράγωνο: $2a^2 = 4\rho^2 \implies a = \sqrt{2}\rho$.

$E_{max} = 2\rho^2$, πλευρά τετραγώνου $a = \sqrt{2}\rho$.

Μέθοδος ποινής:

$\max E = ab$ υπό $a^2 + b^2 = 4\rho^2 \iff \min -ab$ υπό $h(a, b) = a^2 + b^2 - 4\rho^2 = 0$.

Συνάρτηση ποινής: $P(a, b, \mu) = -ab + \frac{\mu}{2}(a^2 + b^2 - 4\rho^2)^2$ $\nabla P = 0: -b + 2\mu(a^2 + b^2 - 4\rho^2) \cdot a = 0 -a + 2\mu(a^2 + b^2 - 4\rho^2) \cdot b = 0$

Διαιρώντας: $\frac{b}{a} = \frac{a}{b} \implies a = b$. Αντικαθιστώντας: $2a^2 = 4\rho^2 \implies a = \sqrt{2}\rho$. Για $\mu \rightarrow \infty$: $(a, b) \rightarrow (\sqrt{2}\rho, \sqrt{2}\rho)$.

Θέμα 2 (2 μονάδες)

Θέλουμε να ελαχιστοποιήσουμε τη συνάρτηση

$$f(x, y) = x^2 + 2y^2 + 2xy$$

ως προς x, y με περιορισμούς:

$$-1 \leq x \leq 1$$

$$-2 \leq y \leq 1/2$$

Να δείξετε ότι μπορούν να χρησιμοποιηθούν μέθοδοι προβολής για να λυθεί το πρόβλημα.

Λύση

Λύση:

Για να μπορούν να χρησιμοποιηθούν μέθοδοι προβολής, πρέπει το πρόβλημα βελτιστοποίησης να είναι κυρτό (Convex Programming). Αυτό απαιτεί δύο συνθήκες:

1. Η αντικειμενική συνάρτηση $f(x, y)$ να είναι κυρτή.
2. Το σύνολο των περιορισμών S να είναι κυρτό.

1. Έλεγχος Κυρτότητας της $f(x, y)$: Η συνάρτηση είναι $f(x, y) = x^2 + 2y^2 + 2xy$. Υπολογίζουμε την Εσσιανή μήτρα (Hessian Matrix):

$$\nabla f = \begin{bmatrix} 2x + 2y \\ 4y + 2x \end{bmatrix}$$

$$H_f = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix} = \begin{bmatrix} 2 & 2 \\ 2 & 4 \end{bmatrix}$$

Οι κύριες ελάσσοσες ορίζουσες είναι:

- $\Delta_1 = 2 > 0$
- $\Delta_2 = \det(H) = 2 \cdot 4 - 2 \cdot 2 = 8 - 4 = 4 > 0$

Επειδή όλες οι κύριες ελάσσοες ορίζουσες είναι θετικές, ο πίνακας H_f είναι θετικά ορισμένος παντού. Άρα η συνάρτηση f είναι **αυστηρά κυρτή**.

2. Έλεγχος Κυρτότητας των Περιορισμών: Οι περιορισμοί $S = \{(x, y) \in \mathbb{R}^2 \mid -1 \leq x \leq 1, -2 \leq y \leq 1/2\}$ ορίζουν ένα ορθογώνιο. Κάθε ορθογώνιο είναι κυρτό σύνολο (ως τομή ημιεπιπέδων που είναι κυρτά σύνολα).

Συμπέρασμα: Εφόσον έχουμε ελαχιστοποίηση κυρτής συνάρτησης σε κυρτό σύνολο περιορισμών, μπορούν να χρησιμοποιηθούν επαναληπτικές μέθοδοι προβολής (όπως η Projected Gradient Descent), οι οποίες εγγυώνται σύγκλιση στο ολικό ελάχιστο.

Θέμα 3 (3 μονάδες)

Χαρακτηρίστε τις παρακάτω προτάσεις ως Σωστές ή Λανθασμένες

1) (1 μονάδα) Η $f(x) = |x|$, $x \in \mathbb{R}$ είναι γνήσια κυρτή διότι

$$f(x_1) + \nabla f^T(x_1)(x_2 - x_1) < f(x_2), \quad \forall x_1, x_2 \in \mathbb{R}.$$

2) (1 μονάδα) Η συνάρτηση

$$f(x) = \begin{cases} x^2, & 0 \leq x \leq 1 \\ x + 1, & 1 < x \end{cases}$$

είναι γνήσια κυρτή.

3) (1 μονάδα) Δίνονται οι περιορισμοί

$$\begin{aligned} x_1 &\geq 0 \\ x_2 &\geq 0 \\ -x_1 + (x_2 - 1)^2 &\geq 0 \end{aligned}$$

Το $(x_1^*, x_2^*) = (0, 1)$ είναι εφικτό και κανονικό.

Λύση

Λύση:

1) (1 μονάδα) $f(x) = |x|$ γνήσια κυρτή; **ΛΑΘΟΣ.** Η $f(x) = |x|$ είναι κυρτή αλλά **όχι γνήσια κυρτή**. Για $x_1, x_2 > 0$: $f(x_1) + f'(x_1)(x_2 - x_1) = x_1 + (x_2 - x_1) = x_2 = f(x_2)$ (ισότητα, όχι αυστηρή ανισότητα). Επιπλέον, η f δεν είναι διαφορίσιμη στο $x = 0$.

2) (1 μονάδα) $f(x) = \begin{cases} x^2, & 0 \leq x \leq 1 \\ x + 1, & x > 1 \end{cases}$ γνήσια κυρτή; **ΛΑΘΟΣ.** Για $x > 1$, $f''(x) = 0$ (γραμμική), άρα

δεν είναι γνήσια κυρτή εκεί. Επίσης, στο $x = 1$: $f(1) = 1$ (από αριστερά), $f(1) = 2$ (από δεξιά) - ασυνέχεια! Άρα η συνάρτηση δεν είναι καν συνεχής.

3) (1 μονάδα) $(0, 1)$ εφικτό και κανονικό; Περιορισμοί: $x_1 \geq 0$, $x_2 \geq 0$, $-x_1 + (x_2 - 1)^2 \geq 0$. Για $(0, 1)$: $0 \geq 0 \checkmark$, $1 \geq 0 \checkmark$, $0 + 0 = 0 \geq 0 \checkmark$. **Εφικτό.** Ενεργοί περιορισμοί: $g_1 = -x_1 = 0$, $g_3 = -x_1 + (x_2 - 1)^2 = 0$. $\nabla g_1 = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$, $\nabla g_3 = \begin{bmatrix} -1 \\ 2(x_2 - 1) \end{bmatrix}|_{(0,1)} = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$. Οι κλίσεις είναι γραμμικά εξαρτημένες! **Δεν είναι κανονικό (LICQ αποτυγχάνει).** **ΛΑΘΟΣ.**

Ιούνιος 2023 -- Θέματα & Λύσεις

Θέμα 2 (7.5 μονάδες)

Έστω το σύστημα μίας εισόδου x και μίας εξόδου y για το οποίο γνωρίζουμε ότι η μαθηματική σχέση που συνδέει την είσοδο με την έξοδο είναι:

$$y = a_0 + a_1x + a_2x^2$$

όπου a_0, a_1, a_2 είναι κάποιες σταθερές αλλά άγνωστες παράμετροι. Με σκοπό την εκτίμηση των άγνωστων παραμέτρων συλλέγουμε τρία ζεύγη μετρήσεων όπως στον πίνακα που ακολουθεί:

i	1	2	3
$x(i)$	1	2	3
$y(i)$	4	9	16

A) (5 μονάδες) Να προσδιοριστούν αναλυτικά (χωρίς την εφαρμογή κάποιας μεθόδου βελτιστοποίησης) οι εκτιμήσεις $\hat{a}_0, \hat{a}_1, \hat{a}_2$ των άγνωστων παραμέτρων, έτσι ώστε να ελαχιστοποιείται το άθροισμα των τετραγώνων των σφαλμάτων μεταξύ της πραγματικής εξόδου και της εκτίμησης αυτής \hat{y} . Να θεωρήσετε ότι:

$$\hat{y} = \hat{a}_0 + \hat{a}_1x + \hat{a}_2x^2$$

B) (2.5 μονάδες) Να σχεδιάσετε τη μέθοδο της μέγιστης κλίσης με σταθερό βήμα για την αναδρομική εκτίμηση των άγνωστων παραμέτρων.

Λύση

Λύση:

A) Αναλυτική Επίλυση (Ελάχιστα Τετράγωνα) Το μοντέλο είναι $y = a_0 + a_1x + a_2x^2$. Θέλουμε να ελαχιστοποιήσουμε το άθροισμα σφαλμάτων:

$$E(a) = \sum_{i=1}^3 (y_i - (a_0 + a_1x_i + a_2x_i^2))^2$$

Αντικαθιστώντας τα δεδομένα $(1, 4), (2, 9), (3, 16)$:

$$E = (4 - (a_0 + a_1 + a_2))^2 + (9 - (a_0 + 2a_1 + 4a_2))^2 + (16 - (a_0 + 3a_1 + 9a_2))^2$$

Για να ελαχιστοποιηθεί το E , μηδενίζουμε τις μερικές παραγώγους ή λύνουμε το σύστημα εξισώσεων:

$$4 = a_0 + a_1 + a_2$$

$$9 = a_0 + 2a_1 + 4a_2$$

$$16 = a_0 + 3a_1 + 9a_2$$

(Παρατηρούμε ότι $4 = 2^2, 9 = 3^2, 16 = 4^2$, δηλαδή $y = (x+1)^2 = x^2 + 2x + 1$). Άρα η ακριβής λύση είναι:

$$a_0 = 1, \quad a_1 = 2, \quad a_2 = 1$$

B) Μέθοδος Μέγιστης Κλίσης Η παράγωγος ∇E ως προς $a = [a_0, a_1, a_2]^T$ είναι:

$$\nabla E = \begin{bmatrix} \frac{\partial E}{\partial a_0} \\ \frac{\partial E}{\partial a_1} \\ \frac{\partial E}{\partial a_2} \end{bmatrix} = -2 \sum (y_i - \hat{y}_i) \begin{bmatrix} 1 \\ x_i \\ x_i^2 \end{bmatrix}$$

Ο αλγόριθμος ανανέωσης είναι:

$$a_{k+1} = a_k - \gamma \nabla E(a_k)$$

Γ) Σύγκλιση σε μία επανάληψη Για να συγκλίνει ο αλγόριθμος μέγιστης κλίσης σε μία επανάληψη, πρέπει να ξεκινήσουμε από το βέλτιστο! Δηλαδή, αν επιλέξουμε αρχικές τιμές:

$$\hat{a}_0(1) = 1, \quad \hat{a}_1(1) = 2, \quad \hat{a}_2(1) = 1$$

τότε $\nabla E(a_{start}) = 0$, οπότε $a_{next} = a_{start} - 0 = a_{start}$. Η λύση βρίσκεται ήδη εκεί.

Σεπτέμβριος 2023 -- Θέματα & Λύσεις

Θέμα 1 (2 μονάδες)

Έστω $f(x)$ μια κυρτή συνάρτηση στο \mathbb{R} . Αν x_1, x_2, x_3 τρεις πραγματικοί αριθμοί που ικανοποιούν $x_2 < x_3 < x_1$, να δειχθεί ότι:

$$\frac{f(x_2) - f(x_1)}{x_2 - x_1} \leq \frac{f(x_3) - f(x_1)}{x_3 - x_1}$$

Λύση

Λύση:

Έστω $f(x)$ μια κυρτή συνάρτηση στο \mathbb{R} και $x_1 < x_2 < x_3$. Μπορούμε να εκφράσουμε το x_2 ως κυρτό συνδυασμό των x_1 και x_3 :

$$x_2 = \frac{x_3 - x_2}{x_3 - x_1}x_1 + \frac{x_2 - x_1}{x_3 - x_1}x_3$$

Οι συντελεστές $\lambda_1 = \frac{x_3 - x_2}{x_3 - x_1}$ και $\lambda_2 = \frac{x_2 - x_1}{x_3 - x_1}$ είναι θετικοί και το άθροισμά τους είναι 1.

Λόγω της κυρτότητας της f , ισχύει:

$$f(x_2) \leq \frac{x_3 - x_2}{x_3 - x_1}f(x_1) + \frac{x_2 - x_1}{x_3 - x_1}f(x_3) \quad (1)$$

Επιπλέον, μπορούμε να γράψουμε το $f(x_2)$ ως:

$$f(x_2) = 1 \cdot f(x_2) = \left(\frac{x_3 - x_2}{x_3 - x_1} + \frac{x_2 - x_1}{x_3 - x_1} \right) f(x_2) = \frac{x_3 - x_2}{x_3 - x_1}f(x_2) + \frac{x_2 - x_1}{x_3 - x_1}f(x_2) \quad (2)$$

Από (1) και (2):

$$\frac{x_3 - x_2}{x_3 - x_1}f(x_2) + \frac{x_2 - x_1}{x_3 - x_1}f(x_2) \leq \frac{x_3 - x_2}{x_3 - x_1}f(x_1) + \frac{x_2 - x_1}{x_3 - x_1}f(x_3)$$

Αναδιατάσσοντας τους όρους:

$$\frac{x_3 - x_2}{x_3 - x_1}(f(x_2) - f(x_1)) \leq \frac{x_2 - x_1}{x_3 - x_1}(f(x_3) - f(x_2))$$

ή ισοδύναμα (πολλαπλασιάζοντας με $x_3 - x_1 > 0$):

$$(x_3 - x_2)(f(x_2) - f(x_1)) \leq (x_2 - x_1)(f(x_3) - f(x_2))$$

Διαιρώντας με $(x_3 - x_2)(x_2 - x_1)$ (που είναι θετικό) καταλήγουμε στο ζητούμενο:

$$\frac{f(x_2) - f(x_1)}{x_2 - x_1} \leq \frac{f(x_3) - f(x_2)}{x_3 - x_2}$$

Η γεωμετρική ερμηνεία είναι ότι η κλίση της χορδής (x_1, x_2) είναι μικρότερη ή ίση από την κλίση της χορδής (x_2, x_3) για μια κυρτή συνάρτηση.

Θέμα 2 (4 μονάδες)

Μια εταιρεία κατασκευάζει ηχεία σε δύο εργοστάσια που βρίσκονται σε διαφορετικές πόλεις, το E_1 και το E_2 . Το E_1 παράγει 350 ηχεία ανά ημέρα, ενώ το E_2 250 ηχεία ανά ημέρα. Τα ηχεία αποστέλλονται σε δύο αποθήκες την A_1 και την A_2 . Οι αποθήκες θεωρούνται ότι είναι αρχικά γεμάτες. Επίσης, ηχεία αποστέλλονται από τις δύο αποθήκες στα κέντρα διανομής τριών πόλεων $\{K_1, K_2, K_3\}$. Μεταξύ των αποθηκών υπάρχει δυνατότητα αποστολής περιορισμένου αριθμού ηχείων.

Έστω ότι τα κέντρα διανομής K_i , $i = 1, 2, 3$ ζητούν 150, 200 και 220 ηχεία ανά ημέρα αντίστοιχα. Η μετοχή των αποθηκών αποστολή ηχείων περιορίζεται σε 35 ηχεία ανά ημέρα, με κόστος μεταφοράς 2 χρηματικές μονάδες ανά ηχείο. Το κόστος μεταφοράς (σε χρηματικές μονάδες) ανά ηχείο για τους υπόλοιπους προορισμούς είναι: $(E_1 \rightarrow A_1) = 36$, $(E_1 \rightarrow A_2) = 27$, $(E_2 \rightarrow A_1) = 33$, $(E_2 \rightarrow A_2) = 29$, $(A_1 \rightarrow K_1) = 21$, $(A_1 \rightarrow K_2) = 15$, $(A_1 \rightarrow K_3) = 13$, $(A_2 \rightarrow K_1) = 16$, $(A_2 \rightarrow K_2) = 16$, $(A_2 \rightarrow K_3) = 17$.

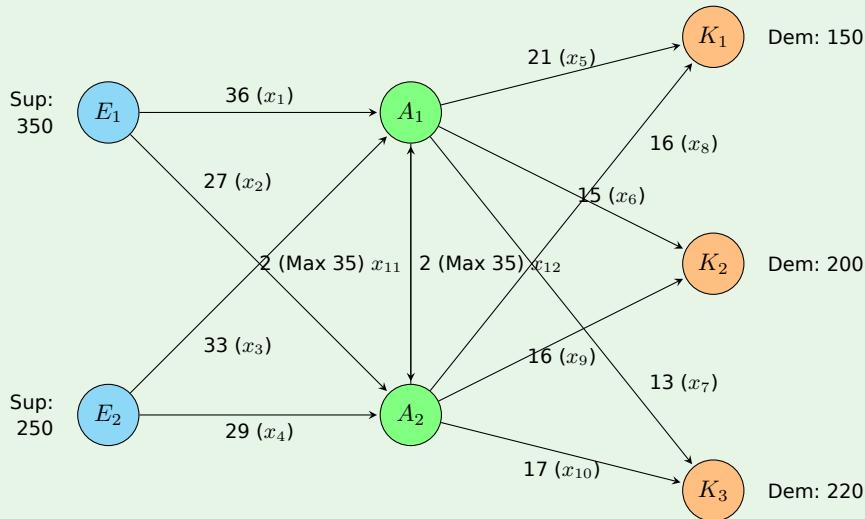
Να διατυπωθεί το πρόβλημα βελτιστοποίησης, η επίλυση του οποίου οδηγεί στην ελαχιστοποίηση του

συνολικού κόστους διανομής ηχείων ανά ημέρα.

Λύση

Λύση:

Δικτυακή Αναπαράσταση:



Ορίζουμε τις μεταβλητές απόφασης x_i για τις ροές των ηχείων:

- $x_1: E_1 \rightarrow A_1, x_2: E_1 \rightarrow A_2$ (Παραγωγή από E_1)
- $x_3: E_2 \rightarrow A_1, x_4: E_2 \rightarrow A_2$ (Παραγωγή από E_2)
- x_5, x_6, x_7 : Ροές από A_1 προς K_1, K_2, K_3
- x_8, x_9, x_{10} : Ροές από A_2 προς K_1, K_2, K_3
- $x_{11}: A_1 \rightarrow A_2, x_{12}: A_2 \rightarrow A_1$ (Μεταφορά μεταξύ αποθηκών με όριο 35)

Μαθηματική Διατύπωση (LP):

$$\min Z = 36x_1 + 27x_2 + 33x_3 + 29x_4 + 21x_5 + 15x_6 + 13x_7 + 16x_8 + 16x_9 + 17x_{10} + 2x_{11} + 2x_{12}$$

Περιορισμοί:

1. Περιορισμοί Προσφοράς:

$$x_1 + x_2 \leq 350 \quad (E_1)$$

$$x_3 + x_4 \leq 250 \quad (E_2)$$

2. Περιορισμοί Ζήτησης:

$$x_5 + x_8 = 150 \quad (K_1)$$

$$x_6 + x_9 = 200 \quad (K_2)$$

$$x_7 + x_{10} = 220 \quad (K_3)$$

3. Ισοζύγιο Ροών στις Αποθήκες (Εισροές = Εκροές):

$$x_1 + x_3 + x_{12} = x_5 + x_6 + x_7 + x_{11} \quad (A_1)$$

$$x_2 + x_4 + x_{11} = x_8 + x_9 + x_{10} + x_{12} \quad (A_2)$$

4. Περιορισμοί Χωρητικότητας Συνδέσμων:

$$x_{11} \leq 35 \quad (A_1 \rightarrow A_2)$$

$$x_{12} \leq 35 \quad (A_2 \rightarrow A_1)$$

5. Μη αρνητικότητα:

$$x_i \geq 0, \quad \forall i$$

Σημείωση (Βέλτιστη Λύση): Αν επιλύσουμε το πρόβλημα, προκύπτει ότι συμφέρει να διακινήσουμε το φορτίο κυρίως μέσω της αποθήκης A_2 (λόγω χαμηλότερου κόστους εισόδου 27, 29 έναντι 36, 33) και να χρησιμοποιήσουμε τον σύνδεσμο $A_2 \rightarrow A_1$ στο μέγιστο (35 μονάδες) για να εκμεταλλευτούμε το εξαιρετικά χαμηλό κόστος του $A_1 \rightarrow K_3$ (13). Οι βέλτιστες ροές είναι:

- $E_1 \rightarrow A_2$: 350 (Όλη η παραγωγή)
- $E_2 \rightarrow A_2$: 220 (Μέρος της παραγωγής)
- $A_2 \rightarrow A_1$: 35 (Μέγιστη χωρητικότητα συνδέσμου)
- $A_1 \rightarrow K_3$: 35 (Τροφοδοσία μέρους του K_3)
- $A_2 \rightarrow K_3$: 185, $A_2 \rightarrow K_1$: 150, $A_2 \rightarrow K_2$: 200

Σύνολο κόστους: $Z_{min} = 24735$.

Θέμα 3 (4 μονάδες)

Η υλοποίηση της μεθόδου της μέγιστης καθόδου στηρίζεται στον αριθμητικό υπολογισμό του διανύσματος κλίσης της αντικειμενικής συνάρτησης, διαδικασία που εμπεριέχει σφάλματα στρογγυλοποίησης, διακριτοποίησης κλπ. Έστω η μέθοδος της μέγιστης καθόδου με σφάλμα, που χρησιμοποιεί βήμα με εξασθένηση

$$\left(\lim_{k \rightarrow \infty} \gamma_k = 0 \right).$$

$$x_{k+1} = x_k - \gamma_k (\nabla f(x_k) + \omega_k), \quad \gamma_k > 0.$$

Το ω_k παριστάνει σε κάθε επανάληψη της μεθόδου ένα μη-ελέγχιμο διάνυσμα σφάλματος. Το ω_k είναι μεν άγνωστο αλλά, γνωρίζουμε ότι το μέτρο του είναι φραγμένο άνω με φράγμα ανάλογο του βήματος γ_k . Στην αναδρομική εξίσωση της μεθόδου το $\nabla f(x_k)$ συμβολίζει την πραγματική αλλά άγνωστη τιμή του διανύσματος κλίσης και το $\nabla f(x_k) + \omega_k$ την εκτίμηση της άγνωστης αυτής τιμής. Συνεπώς, το $\nabla f(x_k) + \omega_k$ είναι γνωστό.

Υποθέτοντας ότι $\nabla f(x_k) \neq 0$, να δείξετε ότι υπάρχει πεπερασμένος αριθμός επαναλήψεων $\bar{k} > 0$ τέτοιος ώστε για κάθε $k \geq \bar{k}$ η κατεύθυνση αναζήτησης $-(\nabla f(x_k) + \omega_k)$ εξακολουθεί να ικανοποιεί την ιδιότητα της επαναληπτικής καθόδου, για αρκούντως μικρό γ_k .

Λύση

Λύση:

Δίνεται η αναδρομική σχέση:

$$x_{k+1} = x_k + d_k = x_k - \gamma_k (\nabla f(x_k) + \omega_k)$$

με $\lim_{k \rightarrow \infty} \gamma_k = 0$ και $\|\omega_k\| \leq c\gamma_k$ (το σφάλμα είναι φραγμένο και φθίνει μαζί με το βήμα). Υποθέτουμε $\nabla f(x_k) \neq 0$.

Για να έχουμε κατεύθυνση καθόδου $p_k = -(\nabla f(x_k) + \omega_k)$, πρέπει να ισχύει:

$$\begin{aligned} \nabla f(x_k)^T p_k < 0 &\iff \nabla f(x_k)^T [-(\nabla f(x_k) + \omega_k)] < 0 \\ &\iff \nabla f(x_k)^T (\nabla f(x_k) + \omega_k) > 0 \\ &\iff \|\nabla f(x_k)\|^2 + \nabla f(x_k)^T \omega_k > 0 \end{aligned}$$

Ο πρώτος όρος $\|\nabla f(x_k)\|^2$ είναι θετικός (αφού $\nabla f(x_k) \neq 0$). Πρέπει λοιπόν να ελέγξουμε αν το άθροισμα παραμένει θετικό. Από την ανισότητα Cauchy-Schwarz:

$$|\nabla f(x_k)^T \omega_k| \leq \|\nabla f(x_k)\| \|\omega_k\|$$

Άρα:

$$\nabla f(x_k)^T \omega_k \geq -\|\nabla f(x_k)\| \|\omega_k\| \geq -\|\nabla f(x_k)\| c \gamma_k$$

Αντικαθιστώντας στην ανισότητα καθόδου:

$$\|\nabla f(x_k)\|^2 + \nabla f(x_k)^T \omega_k \geq \|\nabla f(x_k)\|^2 - c \gamma_k \|\nabla f(x_k)\| = \|\nabla f(x_k)\| (\|\nabla f(x_k)\| - c \gamma_k)$$

Για να ισχύει η ιδιότητα της καθόδου, αρκεί:

$$\|\nabla f(x_k)\| - c \gamma_k > 0 \iff \gamma_k < \frac{\|\nabla f(x_k)\|}{c}$$

Επειδή $\lim_{k \rightarrow \infty} \gamma_k = 0$ και $\|\nabla f(x_k)\|$ είναι μη μηδενικό (και υποθέτουμε φραγμένο μακριά από το 0 σε μια περιοχή μη στάσιμου σημείου), υπάρχει ένα \bar{k} τέτοιο ώστε για κάθε $k \geq \bar{k}$, το γ_k να είναι αρκετά μικρό ώστε να ικανοποιείται η συνθήκη. Επομένως, μετά από πεπερασμένο αριθμό επαναλήψεων, η κατεύθυνση θα είναι πάντα κατεύθυνση καθόδου.

Ιούνιος 2024 -- Θέματα & Λύσεις

Θέμα 1 (3 μονάδες)

Να χαρακτηριστούν ως Σωστές ή Λανθασμένες οι παρακάτω προτάσεις και να αιτιολογήσετε την απάντησή σας.

A) (1 μονάδα) Η $f(x) = |x|$, $x \in \mathbb{R}$ είναι γνήσια κυρτή διότι:

$$f(x_1) + \nabla f^T(x_1)(x_2 - x_1) < f(x_2), \quad \forall x_1, x_2 \in \mathbb{R}.$$

B) (1 μονάδα) Το $z = \begin{bmatrix} 3 & 1 \\ 2 & 2 \end{bmatrix}^T$ είναι κυρτός συνδυασμός των $x = [1 \ 1]^T$, $y = [2 \ 0]^T$.

Γ) (1 μονάδα) Για την ελαχιστοποίηση χωρίς περιορισμούς της συνάρτησης $f(x)$, $x \in \mathbb{R}$ η οποία είναι τουλάχιστον δύο φορές παραγωγίσιμη, ο αλγόριθμος:

$$x_{k+1} = x_k - \gamma d_k, \quad \gamma > 0$$

επιβάλει την ιδιότητα της επαναληπτικής καθόδου αρκεί:

$$\nabla f^T(x_k)d_k < 0$$

σε κάθε επανάληψη k του αλγορίθμου.

Λύση

Λύση:

A) Η $f(x) = |x|$ δεν είναι γνήσια κυρτή γιατί δεν ικανοποιείται η ανισότητα

$$f(x_1) + \nabla f(x_1)(x_2 - x_1) < f(x_2)$$

για όλες τις περιπτώσεις $x_1, x_2 \in \mathbb{R}$. Συγκεκριμένα για τις περιπτώσεις $x_1, x_2 > 0$ και $x_1, x_2 < 0$, η ανισότητα γίνεται ισότητα, άρα η συνάρτηση δεν είναι αυστηρά κυρτή (είναι απλώς κυρτή). **Απάντηση:**

ΛΑΘΟΣ

B) Για να είναι το z κυρτός γραμμικός συνδυασμός των x, y πρέπει να υπάρχουν $a, b \in [0, 1]$ τέτοια ώστε:

$$z = ax + by$$

$$a + b = 1$$

Αντικαθιστώντας τα διανύσματα:

$$\begin{bmatrix} 3 \\ 1 \end{bmatrix} = a \begin{bmatrix} 1 \\ 1 \end{bmatrix} + b \begin{bmatrix} 2 \\ 0 \end{bmatrix} = \begin{bmatrix} a + 2b \\ a \end{bmatrix}$$

Επιλύουμε το σύστημα:

$$\begin{cases} a + 2b = 3 \\ a = 1 \end{cases} \implies \begin{cases} a = 1 \\ 1 + 2b = 3 \implies 2b = 2 \implies b = 1 \end{cases}$$

Το άθροισμα των συντελεστών είναι $a + b = 1 + 1 = 2 \neq 1$. Επομένως, το z **δεν** είναι κυρτός γραμμικός συνδυασμός (είναι απλώς γραμμικός συνδυασμός). **Απάντηση: ΛΑΘΟΣ** (Οι συντελεστές που ικανοποιούν τη γραμμική σχέση αθροίζουν στο 2, όχι στο 1).

Γ) Αυτή η συνθήκη ($\nabla f^T(x_k)d_k < 0$) είναι απαραίτητη για να διασφαλιστεί ότι η κατεύθυνση αναζήτησης d_k είναι κατεύθυνση καθόδου (descent direction). Αν ισχύει, τότε για μικρό βήμα γ , η τιμή της συνάρτησης μειώνεται. **Απάντηση: ΣΩΣΤΟ**

Θέμα 2 (3 μονάδες)

Να ελαχιστοποιηθεί αναλυτικά η συνάρτηση

$$f(x, y) = \frac{4x^2y^2 + x^2 + 4y^2}{xy^2}, \quad x > 0, y > 0.$$

Λύση**Λύση:**

Η συνάρτηση γράφεται ως:

$$f(x, y) = \frac{4x^2y^2}{xy^2} + \frac{x^2}{xy^2} + \frac{4y^2}{xy^2} = 4x + \frac{x}{y^2} + \frac{4}{x}, \quad x > 0, y > 0.$$

Επειδή $\frac{x}{y^2} > 0$, έχουμε για όλα τα $x > 0, y > 0$:

$$f(x, y) > 4x + \frac{4}{x}.$$

Με AM-GM:

$$4x + \frac{4}{x} \geq 2\sqrt{4x \cdot \frac{4}{x}} = 8,$$

και ισότητα μόνο όταν $4x = \frac{4}{x} \iff x = 1$. Άρα για κάθε $x > 0, y > 0$:

$$f(x, y) > 8.$$

Επιπλέον, αν πάρουμε $x_n = 1$ και $y_n \rightarrow +\infty$, τότε

$$f(1, y_n) = 4 + \frac{1}{y_n^2} + 4 = 8 + \frac{1}{y_n^2} \rightarrow 8.$$

Συνεπώς, $\inf\{f(x, y) : x > 0, y > 0\} = 8$, αλλά **δεν υπάρχει ελάχιστο** (η τιμή 8 δεν επιτυγχάνεται για κανένα πεπερασμένο y).

Θέμα 3 (4 μονάδες)

Θέλουμε να ελαχιστοποιήσουμε τη συνάρτηση

$$f(x, y) = x^2 + 2y^2 + 2xy$$

ως προς x, y με περιορισμούς:

$$\begin{aligned} -1 &\leq x \leq 1 \\ -2 &\leq y \leq 1/2 \end{aligned}$$

Να δείξετε ότι μπορούν να χρησιμοποιηθούν μέθοδοι προβολής για να λυθεί το πρόβλημα.

Λύση**Λύση:**

Για να μπορούν να χρησιμοποιηθούν μέθοδοι προβολής (Gradient Projection Methods) για την επίλυση του προβλήματος, πρέπει να ικανοποιούνται δύο βασικές συνθήκες:

1. Η αντικειμενική συνάρτηση $f(x, y)$ να είναι κυρτή.
 2. Το σύνολο των περιορισμών Ω να είναι κυρτό σύνολο.
- 1. Έλεγχος Κυρτότητας Συνάρτησης** Η συνάρτηση είναι $f(x, y) = x^2 + 2y^2 + 2xy$. Υπολογίζουμε την

Εσσιανή μήτρα (Hessian Matrix):

$$\nabla f = \begin{bmatrix} 2x + 2y \\ 4y + 2x \end{bmatrix}$$

$$H_f = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix} = \begin{bmatrix} 2 & 2 \\ 2 & 4 \end{bmatrix}$$

Οι κύριες ελάσσοσες ορίζουσες είναι:

- $\Delta_1 = 2 > 0$
- $\Delta_2 = \det(H) = 2 \cdot 4 - 2 \cdot 2 = 8 - 4 = 4 > 0$

Επειδή όλες οι κύριες ελάσσοσες ορίζουσες είναι θετικές, ο πίνακας H_f είναι θετικά ορισμένος για κάθε (x, y) . Άρα η συνάρτηση f είναι **αυστηρά κυρτή** (strictly convex).

2. Έλεγχος Κυρτότητας Περιορισμών Οι περιορισμοί είναι:

$$\begin{aligned} -1 \leq x \leq 1 \\ -2 \leq y \leq 1/2 \end{aligned}$$

Αυτοί ορίζουν ένα ορθογώνιο στο επίπεδο, δηλαδή το καρτεσιανό γινόμενο δύο κλειστών διαστημάτων: $\Omega = [-1, 1] \times [-2, 0.5]$. Κάθε ορθογώνιο (και γενικότερα κάθε πολύεδρο) είναι **κυρτό σύνολο**.

Συμπέρασμα: Εφόσον ελαχιστοποιούμε μια κυρτή συνάρτηση σε ένα κυρτό σύνολο, το πρόβλημα είναι Κυρτού Προγραμματισμού (Convex Programming). Επομένως, μπορούν να χρησιμοποιηθούν μέθοδοι προβολής (όπως η Projected Gradient Descent), οι οποίες εγγυώνται σύγκλιση στο ολικό ελάχιστο.

Σεπτέμβριος 2024 -- Θέματα & Λύσεις

Θέμα 1 (3 μονάδες)

Να αποδείξετε ότι αν διαφορίσιμη και $x^* \in \mathbb{R}^n$ είναι ένα τοπικό ελάχιστο της συνάρτησης f , τότε $\nabla f(x^*) = \vec{0}$.

Λύση

Λύση:

Απόδειξη: Έστω $f : \mathbb{R}^n \rightarrow \mathbb{R}$ διαφορίσιμη και x^* τοπικό ελάχιστο. Εφόσον το x^* είναι τοπικό ελάχιστο, υπάρχει $\delta > 0$ τέτοιο ώστε $f(x^*) \leq f(x)$ για κάθε $x \in B(x^*, \delta)$.

Έστω $\epsilon > 0$ τέτοιο ώστε $\delta > \epsilon$. Από το λήμμα 2.3.1 του βιβλίου (σελ. 15), ισχύει ότι για ένα διάστημα $(x^*, x^* + \epsilon p)$ υπάρχει $z \in (x^*, x^* + \epsilon p)$ τέτοιο ώστε:

$$f(x^* + \epsilon p) = f(x^*) + \epsilon \nabla f^T(z)p$$

Επειδή το x^* είναι τοπικό ελάχιστο, ισχύει $f(x^* + \epsilon p) - f(x^*) \geq 0$. Συνεπώς:

$$\epsilon \nabla f^T(z)p \geq 0$$

Για $\epsilon \rightarrow 0$, έχουμε $z \rightarrow x^*$, άρα:

$$\nabla f^T(x^*)p \geq 0$$

Αν ακολουθήσουμε την ίδια στρατηγική για το διάστημα $(x^*, x^* - \epsilon p)$, θα δείξουμε αντίστοιχα ότι:

$$-\epsilon \nabla f^T(z')p \geq 0 \implies \nabla f^T(z')p \leq 0$$

Για $\epsilon \rightarrow 0$, προκύπτει $\nabla f^T(x^*)p \leq 0$.

Από τις δύο ανισότητες συμπεραίνουμε ότι $\nabla f^T(x^*)p = 0$. Καθώς το διάνυσμα p είναι τυχαίο, έπειτα ότι:

$$\nabla f(x^*) = 0$$

Θέμα 2 (4 μονάδες)

Δίνεται η συνεχώς διαφορίσιμη συνάρτηση $h : \mathbb{R}^n \rightarrow \mathbb{R}$ και δύο σταθερά σημεία $x_1, x_2 \in \mathbb{R}^n$. Ζητείται να βρεθεί σημείο x πάνω στην επιφάνεια $S = \{x \in \mathbb{R}^n \mid h(x) = 0\}$ που ελαχιστοποιεί το άθροισμα των ευκλείδειων αποστάσεων από τα x_1 και x_2 .

α) (2 μονάδες) Να διατυπώσετε μαθηματικά το πρόβλημα βελτιστοποίησης. **β) (2 μονάδες)** Να γράψετε τις αναγκαίες συνθήκες βελτιστότητας (Karush-Kuhn-Tucker) και να δείξετε ότι στο βέλτιστο σημείο x^* , τα διανύσματα διεύθυνσης από το x^* προς τα x_1, x_2 σχηματίζουν ίσες γωνίες με την κάθετο της επιφάνειας (Νόμος της Ανάκλασης).

Λύση

Λύση:

Το πρόβλημα είναι:

$$\min_{h(x)=0} (|x_1 - x| + |x_2 - x|)$$

όπου $|y| = (y^T y)^{1/2}$ είναι η Ευκλείδεια νόρμα.

α) Συνθήκες ΚΚΤ Για προβλήματα με περιορισμούς ισότητας, οι συνθήκες ΚΚΤ είναι:

$$\nabla_x(|x_1 - x|) + \nabla_x(|x_2 - x|) + \lambda \nabla h(x^*) = 0$$

Γνωρίζουμε ότι η κλίση της απόστασης είναι:

$$\nabla(|x_i - x|) = \frac{-(x_i - x)}{|x_i - x|}$$

Άρα η εξίσωση γίνεται:

$$\frac{-(x_1 - x^*)}{|x_1 - x^*|} + \frac{-(x_2 - x^*)}{|x_2 - x^*|} + \lambda \nabla h(x^*) = 0 \quad (1)$$

β) Σχέση γωνιών Πολλαπλασιάζουμε την (1) εσωτερικά με το μοναδιαίο διάνυσμα $\frac{(x_1 - x^*)^T}{|x_1 - x^*|}$:

$$\begin{aligned} \frac{-(x_1 - x^*)^T(x_1 - x^*)}{|x_1 - x^*|^2} + \frac{-(x_1 - x^*)^T(x_2 - x^*)}{|x_1 - x^*||x_2 - x^*|} + \lambda \frac{(x_1 - x^*)^T \nabla h(x^*)}{|x_1 - x^*|} &= 0 \\ -1 - \cos \theta + \lambda \frac{(x_1 - x^*)^T \nabla h(x^*)}{|x_1 - x^*|} &= 0 \\ \Rightarrow \lambda \frac{(x_1 - x^*)^T \nabla h(x^*)}{|x_1 - x^*|} &= 1 + \frac{(x_1 - x^*)^T(x_2 - x^*)}{|x_1 - x^*||x_2 - x^*|} \quad (2) \end{aligned}$$

Πολλαπλασιάζουμε την (1) εσωτερικά με το μοναδιαίο διάνυσμα $\frac{(x_2 - x^*)^T}{|x_2 - x^*|}$:

$$\begin{aligned} \frac{-(x_2 - x^*)^T(x_1 - x^*)}{|x_2 - x^*||x_1 - x^*|} + \frac{-(x_2 - x^*)^T(x_2 - x^*)}{|x_2 - x^*|^2} + \lambda \frac{(x_2 - x^*)^T \nabla h(x^*)}{|x_2 - x^*|} &= 0 \\ \Rightarrow \lambda \frac{(x_2 - x^*)^T \nabla h(x^*)}{|x_2 - x^*|} &= 1 + \frac{(x_2 - x^*)^T(x_1 - x^*)}{|x_2 - x^*||x_1 - x^*|} \quad (3) \end{aligned}$$

Παρατηρούμε ότι το δεξί μέλος των (2) και (3) είναι ίσο, διότι $(x_2 - x^*)^T(x_1 - x^*) = (x_1 - x^*)^T(x_2 - x^*)$ (εσωτερικό γινόμενο στον \mathbb{R}^n). Επομένως:

$$\frac{(x_2 - x^*)^T \nabla h(x^*)}{|x_2 - x^*|} = \frac{(x_1 - x^*)^T \nabla h(x^*)}{|x_1 - x^*|}$$

Διαιρώντας και τα δύο μέλη με $|\nabla h(x^*)|$ (υποθέτοντας ότι $\nabla h(x^*) \neq 0$):

$$\boxed{\frac{(x_1 - x^*)^T \nabla h(x^*)}{|x_1 - x^*||\nabla h(x^*)|} = \frac{(x_2 - x^*)^T \nabla h(x^*)}{|x_2 - x^*||\nabla h(x^*)|}}$$

Αυτό σημαίνει ότι οι γωνίες που σχηματίζουν τα διανύσματα από το x^* προς τα x_1, x_2 με το εφαπτόμενο επίπεδο (ή το κάθετο διάνυσμα ∇h) είναι ίσες (Νόμος της Ανάκλασης).

Θέμα 3 (3 μονάδες)

Επιθυμούμε να ελαχιστοποιήσουμε τη συνάρτηση:

$$f(x_1, x_2) = \frac{1}{2}(x_1^2 - x_2^2) - 3x_2$$

υπό τον περιορισμό:

$$-x_1 - 2x_2 = 6$$

α) (1.5 μονάδες) Να επιλύσετε αναλυτικά το πρόβλημα.

β) (1.5 μονάδες) Να επιλύσετε το πρόβλημα χρησιμοποιώντας τη μέθοδο ποινής.

Λύση

Λύση:

α) Αναλυτική Λύση με KKT Η συνάρτηση είναι $f(x_1, x_2) = \frac{1}{2}(x_1^2 - x_2^2) - 3x_2$ και ο περιορισμός ισότητας $g(x_1, x_2) = -x_1 - 2x_2 - 6 = 0$. Εφόσον έχουμε μόνο ισότιμους περιορισμούς, η προϋπόθεση κυρτότητας της f δεν είναι απαραίτητη για την εφαρμογή των συνθηκών KKT (Βιβλίο σελ. 42-47).

Οι συνθήκες KKT είναι:

$$\nabla f(x^*) + \lambda \nabla g(x^*) = 0$$

Υπολογίζουμε τις κλίσεις:

$$\nabla f(x_1, x_2) = \begin{bmatrix} x_1 \\ -x_2 - 3 \end{bmatrix}, \quad \nabla g(x_1, x_2) = \begin{bmatrix} -1 \\ -2 \end{bmatrix}$$

Άρα το σύστημα εξισώσεων είναι:

$$x_1 - \lambda = 0 \implies x_1 = \lambda \quad (1)$$

$$-x_2 - 3 - 2\lambda = 0 \implies x_2 = -2\lambda - 3 \quad (2)$$

$$-x_1 - 2x_2 = 6 \quad (3)$$

Αντικαθιστούμε τις (1) και (2) στην (3):

$$-\lambda - 2(-2\lambda - 3) = 6 \implies -\lambda + 4\lambda + 6 = 6 \implies 3\lambda = 0 \implies \lambda = 0$$

Για $\lambda = 0$:

$$x_1 = 0, \quad x_2 = -3$$

Άρα η λύση είναι $x^* = (0, -3)$.

β) Μέθοδος Ποινής (Penalty Method) Χρησιμοποιούμε την Τετραγωνική Συνάρτηση Ποινής:

$$P(x, \mu) = f(x) + \frac{\mu}{2}(g(x))^2 = \frac{1}{2}(x_1^2 - x_2^2) - 3x_2 + \frac{\mu}{2}(-x_1 - 2x_2 - 6)^2$$

όπου $\mu > 0$ η παράμετρος ποινής. Για να βρούμε το ελάχιστο της $P(x, \mu)$ για δεδομένο μ , μηδενίζουμε τις μερικές παραγώγους:

$$\begin{aligned} \frac{\partial P}{\partial x_1} &= x_1 + \mu(-x_1 - 2x_2 - 6)(-1) = x_1 + \mu(x_1 + 2x_2 + 6) = 0 \\ &\implies x_1(1 + \mu) + 2\mu x_2 + 6\mu = 0 \quad (1) \end{aligned}$$

$$\begin{aligned} \frac{\partial P}{\partial x_2} &= -x_2 - 3 + \mu(-x_1 - 2x_2 - 6)(-2) = -x_2 - 3 + 2\mu(x_1 + 2x_2 + 6) = 0 \\ &\implies 2\mu x_1 + (4\mu - 1)x_2 + 12\mu - 3 = 0 \quad (2) \end{aligned}$$

Επιλύουμε το σύστημα των (1) και (2) για τα x_1, x_2 . Από την (1): $x_1 = -\frac{2\mu x_2 + 6\mu}{1 + \mu}$. Για $\mu \rightarrow \infty$, το κλάσμα $\frac{\mu}{1 + \mu} \rightarrow 1$. Άρα για μεγάλα μ :

$$x_1 \approx -(2x_2 + 6) = -2x_2 - 6 \quad (3)$$

Αντικαθιστούμε την (3) στην (2):

$$2\mu(-2x_2 - 6) + 4\mu x_2 - x_2 + 12\mu - 3 = 0$$

$$-4\mu x_2 - 12\mu + 4\mu x_2 - x_2 + 12\mu - 3 = 0$$

$$-x_2 - 3 = 0 \implies x_2 = -3$$

Και από την (3):

$$x_1 = -2(-3) - 6 = 6 - 6 \implies x_1 = 0$$

Άρα, το όριο της ακολουθίας των σημείων ελαχίστου της $P(x, \mu)$ καθώς $\mu \rightarrow \infty$ είναι το σημείο $x^* = (0, -3)$, που ταυτίζεται με την αναλυτική λύση.

Σημείωση Θεωρίας (Μέθοδος Ποινής vs Επαυξημένη Λαγκρανζιανή): Γενικά, για την επίλυση τέτοιων προβλημάτων μπορούμε να χρησιμοποιήσουμε και τις δύο μεθόδους. Ωστόσο:

- Η **Μέθοδος της Επαυξημένης Λαγκρανζιανής** σχεδιάστηκε αρχικά μόνο για **ισοτικούς περιορισμούς**. Δεν περιλαμβάνει απευθείας όρο για ανισοτικούς περιορισμούς. Αν υπάρχουν τέτοιοι, πρέπει θεωρητικά να μετατραπούν σε ισοτικούς (π.χ. προσθέτοντας μεταβλητές "slack").
- Η **Μέθοδος της Ποινής** διαθέτει όρους τόσο για ανισοτικούς όσο και για ισοτικούς περιορισμούς. Όταν έχουμε **μόνο ισοτικούς περιορισμούς** (όπως στη συγκεκριμένη άσκηση), απλά χρησιμοποιούμε τον όρο των ισοτικών και παραλείπουμε τον όρο που αφορά τους ανισοτικούς.

Φεβρουάριος 2025 -- Θέματα & Λύσεις

Θέμα 1 (3 μονάδες)

Έστω $f : \mathbb{R}^n \rightarrow \mathbb{R}$ μια συνεχώς διαφορίσιμη συνάρτηση. Στους αλγορίθμους τοπικής αναζήτησης του ελαχίστου της f η διαδικασία τερματίζει στο σημείο x_* για το οποίο:

$$\nabla f(x_*) = 0.$$

Με δεδομένο το παραπάνω, θεωρείστε τη συνάρτηση:

$$f(x_1, x_2) = (x_1 - x_2)^3 + (x_1 + x_2)^3$$

και το πρόβλημα:

$$\min_{x_1, x_2} f(x_1, x_2).$$

Έστω το σημείο $x^* = [x_1^* \ x_2^*]^T$ για το οποίο

$$\nabla f(x^*) = 0.$$

Να δείξετε ότι για τη δοσμένη συνάρτηση f , υπάρχουν άπειρα σημεία z , αρκούντως κοντά στο x^* για τα οποία:

$$f(z) < f(x^*)$$

και επομένως το x^* δεν είναι τοπικό ελάχιστο της f .

Λύση

Λύση:

Βήμα 1: Εύρεση στάσιμων σημείων

$$\nabla f(x) = \begin{bmatrix} 3(x_1 - x_2)^2 + 3(x_1 + x_2)^2 \\ -3(x_1 - x_2)^2 + 3(x_1 + x_2)^2 \end{bmatrix}$$

Εξισώνοντας με το μηδέν $\nabla f(x^*) = 0$:

$$3(x_1 - x_2)^2 + 3(x_1 + x_2)^2 = 0 \quad (1)$$

$$-3(x_1 - x_2)^2 + 3(x_1 + x_2)^2 = 0 \quad (2)$$

Αθροίζοντας τις (1) και (2) έχουμε $6(x_1 + x_2)^2 = 0 \implies x_1 = -x_2$. Αφαιρώντας τες έχουμε $6(x_1 - x_2)^2 = 0 \implies x_1 = x_2$. Συνεπώς, το μοναδικό κρίσιμο σημείο είναι το $x^* = (0, 0)$.

Βήμα 2: Έλεγχος Βελτιστότητας (2ης Τάξης) Υπολογίζουμε τον πίνακα Hessian:

$$\nabla^2 f(x) = \begin{bmatrix} 6(x_1 - x_2) + 6(x_1 + x_2) & -6(x_1 - x_2) + 6(x_1 + x_2) \\ -6(x_1 - x_2) + 6(x_1 + x_2) & 6(x_1 - x_2) + 6(x_1 + x_2) \end{bmatrix}$$

Στο σημείο $x^* = (0, 0)$, έχουμε $\nabla^2 f(0, 0) = \mathbb{O}$ (μηδενικός πίνακας). Επειδή ο Hessian είναι μηδενικός (και άρα έχει ιδιοτιμές 0), το κριτήριο 2ης τάξης είναι **ατελέσφορο**. Δεν μπορούμε να αποφανθούμε αν είναι ελάχιστο, μέγιστο ή σαγματικό σημείο μόνο από τις παραγώγους. **Υποχρεωτικά**, καταφεύγουμε στον ορισμό του τοπικού ελαχίστου.

Βήμα 3: Απόδειξη μέσω Ορισμού Για να είναι το x^* τοπικό ελάχιστο, πρέπει να ισχύει $f(z) \geq f(x^*)$ για όλα τα z σε μια περιοχή του x^* . Εδώ $f(x^*) = f(0, 0) = 0$.

Εξετάζουμε τη συμπεριφορά της συνάρτησης κατά μήκος του άξονα x_1 ($x_2 = 0$). Έστω σημείο δοκιμής $z = (-\epsilon, 0)$ με $\epsilon > 0$ αρκούντως μικρό.

$$f(-\epsilon, 0) = (-\epsilon - 0)^3 + (-\epsilon + 0)^3 = -\epsilon^3 - \epsilon^3 = -2\epsilon^3$$

Για κάθε $\epsilon > 0$, ισχύει $-2\epsilon^3 < 0 \implies f(z) < f(x^*)$.

Άρα, σε οποιαδήποτε γειτονιά του $(0, 0)$, υπάρχουν σημεία με τιμή μικρότερη του $f(x^*)$. Επομένως, το $x^* = (0, 0)$ **δεν είναι τοπικό ελάχιστο**.

Θέμα 2 (7 μονάδες)

Θέλουμε να ελαχιστοποιήσουμε τη συνάρτηση

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

ως προς x_1, x_2 με περιορισμό:

$$x_1 + x_2 = 2.$$

α) (3.5 μονάδες) Να επιλύσετε αναλυτικά το πρόβλημα.

β) (3.5 μονάδες) Να επιλύσετε το πρόβλημα χρησιμοποιώντας τη μέθοδο ποινής.

Λύση**Λύση:****α) Αναλυτική λύση με Πολλαπλασιαστές Lagrange**

Lagrange vs KKT

Γιατί Lagrange και όχι KKT; Οι συνθήκες KKT είναι μια γενίκευση της μεθόδου Lagrange που καλύπτει και ανισοτικούς περιορισμούς. Όταν το πρόβλημα έχει **μόνο περιορισμούς ισότητας** (όπως εδώ το $x_1 + x_2 = 2$), οι συνθήκες KKT ταυτίζονται απόλυτα με τη μέθοδο των Πολλαπλασιαστών Lagrange. Γι' αυτό χρησιμοποιούμε τον όρο "Lagrange" για συντομία, αν και τεχνικά λύνουμε τις συνθήκες KKT για ισοτικούς περιορισμούς.

Το πρόβλημα είναι: $\min f(x)$ υπό τον περιορισμό $h(x) = x_1 + x_2 - 2 = 0$. Σχηματίζουμε τη συνάρτηση Lagrangian:

$$L(x_1, x_2, \lambda) = x_1^4 + x_2^2 + x_1^2 x_2 - 2x_2 + \lambda(x_1 + x_2 - 2)$$

Οι αναγκαίες συνθήκες πρώτης τάξης (KKT) είναι $\nabla L = 0$:

$$\frac{\partial L}{\partial x_1} = 4x_1^3 + 2x_1 x_2 + \lambda = 0 \implies \lambda = -(4x_1^3 + 2x_1 x_2) \quad (1)$$

$$\frac{\partial L}{\partial x_2} = 2x_2 + x_1^2 - 2 + \lambda = 0 \implies \lambda = -(2x_2 + x_1^2 - 2) \quad (2)$$

$$\frac{\partial L}{\partial \lambda} = x_1 + x_2 - 2 = 0 \implies x_2 = 2 - x_1 \quad (3)$$

Εξισώνοντας τις εκφράσεις για το λ από τις (1) και (2):

$$4x_1^3 + 2x_1 x_2 = 2x_2 + x_1^2 - 2$$

Αντικαθιστούμε το x_2 από την (3):

$$4x_1^3 + 2x_1(2 - x_1) = 2(2 - x_1) + x_1^2 - 2$$

$$4x_1^3 + 4x_1 - 2x_1^2 = 4 - 2x_1 + x_1^2 - 2$$

$$4x_1^3 - 3x_1^2 + 6x_1 - 2 = 0$$

Επιλύοντας την κυβική εξίσωση (προσεγγιστικά):

$$x_1 \approx 0.387$$

Και από την (3):

$$x_2 = 2 - 0.387 \implies x_2 \approx 1.613$$

β) Μέθοδος Ποινής (Penalty Method) Χρησιμοποιούμε την Τετραγωνική Συνάρτηση Ποινής (Quadratic Penalty Function):

$$P(x, \mu) = f(x) + \frac{\mu}{2}(h(x))^2 = x_1^4 + x_2^2 + x_1^2 x_2 - 2x_2 + \frac{\mu}{2}(x_1 + x_2 - 2)^2$$

όπου $\mu > 0$ η παράμετρος ποινής. Καθώς $\mu \rightarrow \infty$, η ελαχιστοποίηση της $P(x, \mu)$ αναγκάζει τον όρο $(x_1 + x_2 - 2)^2$ να μηδενιστεί, οδηγώντας στη λύση του δεσμευμένου προβλήματος.

Για κάθε δεδομένο μ , βρίσκουμε το $x(\mu)$ που μηδενίζει την κλίση $\nabla_x P(x, \mu) = 0$:

$$\begin{aligned}\frac{\partial P}{\partial x_1} &= 4x_1^3 + 2x_1x_2 + \mu(x_1 + x_2 - 2) = 0 \\ \frac{\partial P}{\partial x_2} &= 2x_2 + x_1^2 - 2 + \mu(x_1 + x_2 - 2) = 0\end{aligned}$$

Από τις δύο εξισώσεις, απομονώνουμε τον όρο ποινής $\mu(x_1 + x_2 - 2)$:

$$-(4x_1^3 + 2x_1x_2) = \mu(x_1 + x_2 - 2) = -(2x_2 + x_1^2 - 2)$$

Αυτό οδηγεί στην ίδια σχέση που βρήκαμε με τη μέθοδο Lagrange:

$$4x_1^3 + 2x_1x_2 = 2x_2 + x_1^2 - 2$$

Στο όριο $\mu \rightarrow \infty$, η λύση πρέπει να ικανοποιεί τον περιορισμό $x_1 + x_2 - 2 = 0$. Συνδυάζοντας τη σχέση αυτή με τον περιορισμό, καταλήγουμε πάλι στην κυβική εξίσωση:

$$4x_1^3 - 3x_1^2 + 6x_1 - 2 = 0$$

Η οποία δίνει την ίδια φορά λύση:

$x_1 \approx 0.387, \quad x_2 \approx 1.613$
--

Ιούλιος 2025 -- Θέματα & Λύσεις

Θέμα 1 (5 μονάδες)

α) (2.5 μονάδες) Για την ελαχιστοποίηση της συνάρτησης

$$f(x, y) = x^2 + y^2 - 2x - 4y + 5$$

ως προς x, y εκκινώντας από οποιοδήποτε σημείο, έχετε στη διάθεσή σας τους αλγορίθμους: α) μέγιστης κλίσης, β) Newton, γ) Levenberg-Marquardt. Ποιον από τους τρεις θα επιλέξετε, αν σας ενδιαφέρει η εύρεση του ελαχίστου να πραγματοποιηθεί μ' όσο το δυνατόν λιγότερες επαναλήψεις, εκτελώντας σε κάθε επανάληψη τις λιγότερες δυνατές πράξεις; Να αιτιολογήσετε την απάντησή σας.

β) (2.5 μονάδες) Δίνεται το πρόβλημα:

$$\min_{x,y} (x^2 + y^2 - 2x - 4y + 5)$$

$$\begin{aligned} -1 \leq x \leq 1 \\ -1 \leq y \leq 1 \end{aligned}$$

Μπορεί να χρησιμοποιηθεί μέθοδος προβολής για να λυθεί; Να αιτιολογήσετε την απάντησή σας.

Λύση

Λύση:

$$f(x, y) = x^2 + y^2 - 2x - 4y + 5 = (x - 1)^2 + (y - 2)^2$$

α) Επιλογή αλγορίθμου Η συνάρτηση είναι τετραγωνική με Hessian $H = 2I$. Αυτό προκύπτει διότι:

$$\nabla f(x, y) = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix} = \begin{bmatrix} 2x - 2 \\ 2y - 4 \end{bmatrix} \implies H_f = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} = 2I$$

- **Newton:** Συγκλίνει σε **1 βήμα** για τετραγωνικές συναρτήσεις.
- Μέγιστης κλίσης: Πολλά βήματα.
- Levenberg-Marquardt: Για μη-γραμμικά ελάχιστα τετραγώνων.

Απάντηση: Μέθοδος **Newton**, γιατί για τετραγωνική f συγκλίνει σε ακριβώς 1 επανάληψη (Θεώρημα 5.2.4).

β) Μέθοδος προβολής και Αιτιολόγηση Για να χρησιμοποιηθεί μια μέθοδος προβολής της κλίσης (Gradient Projection) με εγγυημένη σύγκλιση στο ολικό ελάχιστο, πρέπει να ικανοποιούνται δύο βασικές συνθήκες:

1. **Κυρτότητα Συνάρτησης:** Η συνάρτηση $f(x, y)$ πρέπει να είναι κυρτή. Όπως δείξαμε στο (α), ο πίνακας Hessian είναι θετικά ορισμένος ($H = 2I$), άρα η f είναι αυστηρά κυρτή.
2. **Κυρτότητα Περιορισμών:** Το σύνολο των περιορισμών Ω πρέπει να είναι κυρτό. Εδώ, το σύνολο $\Omega = [-1, 1] \times [-1, 1]$ είναι ένα τετράγωνο (ορθογώνιο παραλληλόγραμμο), το οποίο είναι εξ ορισμού κυρτό σύνολο.

Επιπλέον, η πρακτική εφαρμοσμότητα της μεθόδου εξαρτάται από την ευκολία υπολογισμού της προβολής. Στην περίπτωσή μας, οι περιορισμοί ορίζουν ένα ορθογώνιο και η προβολή $P_\Omega(x, y)$ ενός σημείου στο σύνολο αυτό αντιστοιχεί στο πλησιέστερο σημείο του ορθογωνίου (γεωμετρικά προφανής υπολογισμός). Το ίδιο ισχύει και για το y .

Συμπέρασμα: Ναι, η μέθοδος προβολής είναι ιδανική για αυτό το πρόβλημα, καθώς πληρούνται οι συνθήκες κυρτότητας και η προβολή υπολογίζεται υπολογιστικά πολύ εύκολα χωρίς να απαιτείται η επίλυση κάποιου υπο-προβλήματος βελτιστοποίησης.

Θέμα 2 (5 μονάδες)

Θέλουμε να ελαχιστοποιήσουμε τη συνάρτηση

$$f(x, y) = x^2 + 2y^2 + 2xy$$

ως προς x, y με περιορισμούς:

$$\begin{aligned} -1 &\leq x \leq 1 \\ -2 &\leq y \leq 1/2 \end{aligned}$$

α) (2.5 μονάδες) Να δείξετε ότι μπορούν να χρησιμοποιηθούν μέθοδοι προβολής για να λυθεί το πρόβλημα.

β) (2.5 μονάδες) Να σχεδιαστεί η μέθοδος Newton με προβολή που να λύνει το πρόβλημα.

Λύση

Λύση:

$$f(x, y) = x^2 + 2y^2 + 2xy \text{ με } -1 \leq x \leq 1, -2 \leq y \leq 1/2.$$

α) Απόδειξη για μέθοδο προβολής Η συνάρτηση f είναι τετραγωνική. Η γενική μορφή είναι $f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T Q\mathbf{x} + \mathbf{c}^T \mathbf{x}$. Για να βρούμε τον πίνακα Q , υπολογίζουμε τον πίνακα Hessian:

$$\nabla f(x, y) = \begin{bmatrix} 2x + 2y \\ 4y + 2x \end{bmatrix} \implies H_f = \begin{bmatrix} 2 & 2 \\ 2 & 4 \end{bmatrix}$$

$$\text{Άρα } Q = H_f = \begin{bmatrix} 2 & 2 \\ 2 & 4 \end{bmatrix}.$$

Ιδιοτιμές: $\det(Q - \lambda I) = (2 - \lambda)(4 - \lambda) - 4 = \lambda^2 - 6\lambda + 4 = 0 \Rightarrow \lambda = 3 \pm \sqrt{5} > 0$. Άρα Q θετικά ορισμένος $\implies f$ αυστηρά κυρτή. Το σύνολο περιορισμών είναι ορθογώνιο (κυρτό). **Άρα** μπορούν να χρησιμοποιηθούν μέθοδοι προβολής.

β) Μέθοδος Newton με προβολή

$$\nabla f = \begin{bmatrix} 2x + 2y \\ 4y + 2x \end{bmatrix}, \quad H = \begin{bmatrix} 2 & 2 \\ 2 & 4 \end{bmatrix}$$

Αλγόριθμος:

1. Υπολογισμός κατεύθυνσης Newton: $d_k = -H^{-1}\nabla f(x_k)$
2. Βήμα Newton: $\tilde{x}_{k+1} = x_k + d_k$
3. Προβολή: $x_{k+1} = \text{proj}_\Omega(\tilde{x}_{k+1})$

Οπου $\text{proj}_\Omega(x, y)$ είναι η ευκλείδεια προβολή του σημείου στο ορθογώνιο Ω .

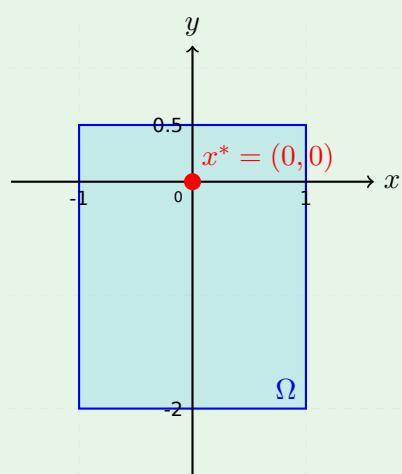
$$H^{-1} = \frac{1}{4} \begin{bmatrix} 4 & -2 \\ -2 & 2 \end{bmatrix} = \begin{bmatrix} 1 & -1/2 \\ -1/2 & 1/2 \end{bmatrix}$$

Εκτέλεση Αλγορίθμου: **Βήμα 1 (Ελαχιστοποίηση χωρίς περιορισμούς):** Λύνουμε το σύστημα $\nabla f(x, y) = 0$ για να βρούμε το σημείο που θα κατέληγε η μέθοδος Newton σε ένα βήμα, αγνοώντας τους περιορισμούς:

$$\begin{cases} 2x + 2y = 0 \\ 4y + 2x = 0 \end{cases} \implies (x, y) = (0, 0)$$

Βήμα 2 (Έλεγχος και Προβολή): Ελέγχουμε αν το σημείο $(0, 0)$ ανήκει στο σύνολο περιορισμών $\Omega = [-1, 1] \times [-2, 1/2]$. Ισχύει ότι $-1 \leq 0 \leq 1$ και $-2 \leq 0 \leq 1/2$. Άρα το σημείο είναι εφικτό: $(0, 0) \in \Omega$.

Επομένως, $x_{new} = \text{proj}_\Omega(0, 0) = (0, 0)$. Το σημείο $(0, 0)$ είναι το ολικό ελάχιστο του προβλήματος με τιμή $f(0, 0) = 0$.



Σεπτέμβριος 2025 -- Θέματα & Λύσεις

Θέμα 1 (2 μονάδες)

Δίνεται η συνάρτηση $f : \mathbb{R} \rightarrow \mathbb{R}$ με τύπο

$$f(x) = \begin{cases} x^2, & 0 \leq x \leq 1 \\ x + 1, & x > 1 \end{cases}$$

Να διερευνηθεί αν είναι κυρτή.

Λύση

Λύση:

Για να είναι η f κυρτή στο πεδίο ορισμού της, πρέπει για κάθε $x, x \in \text{Dom}(f)$ και για κάθε $\lambda \in [0, 1]$ να ισχύει:

$$f(\lambda x + (1 - \lambda)x) \leq \lambda f(x) + (1 - \lambda)f(x)$$

Παρατηρούμε τη γραφική παράσταση της συνάρτησης. Στο διάστημα $[0, 1]$, $f(x) = x^2$ (παραβολή που καταλήγει στο $(1, 1)$). Στο διάστημα $(1, +\infty)$, $f(x) = x + 1$ (ευθεία που ξεκινάει από το $(1, 2)$). **Παρατήρηση:** Υπάρχει ασυνέχεια (άλμα) στο $x = 1$, καθώς $\lim_{x \rightarrow 1^-} f(x) = 1$ ενώ $\lim_{x \rightarrow 1^+} f(x) = 2$.

Θα δοκιμάσουμε να βρούμε ένα αντιπαράδειγμα για τον ορισμό της κυρτότητας. Έστω τα σημεία $x_a = 1$ και $x_b = 2$.

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

Επιλέγουμε $\lambda = 0.5$ (το μέσο το ευθύγραμμου τμήματος που ενώνει τα $(1, 1)$ και $(2, 3)$). Το σημείο είναι $x_{mid} = 0.5(1) + 0.5(2) = 1.5$. Η τιμή της συνάρτησης στο 1.5 είναι:

$$f(1.5) = 1.5 + 1 = 2.5$$

Η τιμή της χορδής (ευθύγραμμου τμήματος) στο μέσο είναι:

$$0.5f(1) + 0.5f(2) = 0.5(1) + 0.5(3) = 0.5 + 1.5 = 2$$

Παρατηρούμε ότι:

$$f(1.5) = 2.5 > 2 = 0.5f(1) + 0.5f(2)$$

Δηλαδή $f(\lambda x + (1 - \lambda)x) > \lambda f(x) + (1 - \lambda)f(x)$. Η ανισότητα της κυρτότητας **δεν** ικανοποιείται.

Συμπέρασμα: Η συνάρτηση f **δεν είναι κυρτή**.

Εναλλακτικός Τρόπος (Άμεσος)

Η συνάρτηση f παρουσιάζει ασυνέχεια στο σημείο $x = 1$ (το οποίο είναι εσωτερικό σημείο του πεδίου ορισμού), καθώς:

$$\lim_{x \rightarrow 1^-} f(x) = 1 \neq \lim_{x \rightarrow 1^+} f(x) = 2$$

Γνωρίζουμε από τη θεωρία ότι κάθε κυρτή συνάρτηση ορισμένη σε ανοικτό διάστημα (ή γενικότερα σε κυρτό σύνολο) οφείλει να είναι **συνεχής** στο εσωτερικό του. Συνεπώς, η ύπαρξη ασυνέχειας αρκεί για να συμπεράνουμε άμεσα ότι η συνάρτηση **δεν** είναι κυρτή.

Θέμα 2 (4 μονάδες)

Εργοστάσιο παράγει ημερήσιως τα προϊόντα Α και Β σε ποσότητες x_1, x_2 αντίστοιχα. Η ημερήσια παραγωγή του Α δεν μπορεί να ξεπεράσει τους 4 τόνους, ενώ του Β τους $\sqrt{25 - x_1^2}$ τόνους. Το κάθε προϊόν Α που παράγεται δίνει κέρδος 9 χρηματικές μονάδες και του Β δίνει κέρδος 5 χρηματικές μονάδες. Επιθυμούμε να προσδιορίσουμε τον καταμερισμό της ημερήσιας παραγωγής του εργοστασίου ώστε να βελτιστοποιείται το κέρδος.

α) (2 μονάδες) Να διατυπωθεί ως πρόβλημα ελαχιστοποίησης.

β) (2 μονάδες) Χρησιμοποιώντας τις συνθήκες Karush-Kuhn-Tucker να λυθεί, αν είναι εφικτό, αναλυτικά το πρόβλημα του ερωτήματος (α).

Λύση

Λύση:

α) Μαθηματική Διατύπωση Έστω x_1, x_2 οι ποσότητες παραγωγής για τα προϊόντα A και B αντίστοιχα. Ο στόχος είναι η μεγιστοποίηση του κέρδους:

$$\max P(x_1, x_2) = 9x_1 + 5x_2$$

Περιορισμοί:

1. $x_1 \leq 4$
2. $x_2 \leq \sqrt{25 - x_1^2} \implies x_2^2 \leq 25 - x_1^2 \implies x_1^2 + x_2^2 \leq 25$ (παραδοχή $x_2 \geq 0$)
3. Φυσικοί περιορισμοί: $x_1 \geq 0, x_2 \geq 0$.

Για ελαχιστοποίηση, αντιστρέφουμε το πρόσημο της αντικειμενικής συνάρτησης και γράφουμε τους περιορισμούς στη μορφή $g(x) \leq 0$.

$$\min f(x_1, x_2) = -9x_1 - 5x_2$$

Υπό τους περιορισμούς:

$$\begin{aligned} g_1(x) &= x_1 - 4 \leq 0 \\ g_2(x) &= x_1^2 + x_2^2 - 25 \leq 0 \\ -x_1 &\leq 0, -x_2 \leq 0 \end{aligned}$$

β) Επίλυση με KKT Η συνάρτηση Lagrange είναι:

$$L(x, \lambda) = -9x_1 - 5x_2 + \lambda_1(x_1 - 4) + \lambda_2(x_1^2 + x_2^2 - 25)$$

Συνθήκες KKT:

1. **Στασιμότητα** ($\nabla_x L = 0$):

$$\frac{\partial L}{\partial x_1} = -9 + \lambda_1 + 2\lambda_2 x_1 = 0 \implies \lambda_1 + 2\lambda_2 x_1 = 9 \quad (1)$$

$$\frac{\partial L}{\partial x_2} = -5 + 2\lambda_2 x_2 = 0 \implies 2\lambda_2 x_2 = 5 \quad (2)$$

2. **Συμπληρωματικότητα:**

$$\lambda_1(x_1 - 4) = 0 \quad (3)$$

$$\lambda_2(x_1^2 + x_2^2 - 25) = 0 \quad (4)$$

3. **Εφικτότητα:**

$$x_1 \leq 4, \quad x_1^2 + x_2^2 \leq 25, \quad x \geq 0$$

4. **Δυϊκή Εφικτότητα:**

$$\lambda_1 \geq 0, \quad \lambda_2 \geq 0$$

Ανάλυση Περιπτώσεων: Από την (2): $2\lambda_2 x_2 = 5 \implies \lambda_2 \neq 0$ και $x_2 \neq 0$. Εφόσον $\lambda_2 > 0$, από την (4) ο περιορισμός είναι ενεργός: $x_1^2 + x_2^2 = 25$.

Περίπτωση 1: $\lambda_1 = 0$ (1) $\implies 2\lambda_2 x_1 = 9 \implies \lambda_2 = \frac{4.5}{x_1}$ (2) $\implies 2\lambda_2 x_2 = 5 \implies \lambda_2 = \frac{2.5}{x_2}$ Εξισώνοντας: $\frac{4.5}{x_1} = \frac{2.5}{x_2} \implies 9x_2 = 5x_1 \implies x_1 = 1.8x_2$. Στον κύκλο: $(1.8x_2)^2 + x_2^2 = 25 \implies 3.24x_2^2 + x_2^2 = 25 \implies 4.24x_2^2 = 25 \implies x_2^2 \approx 5.896$. $x_2 \approx 2.428$ και $x_1 \approx 4.37$. Όμως $x_1 \leq 4$. Το $4.37 > 4$, άρα **Απορρίπτεται**.

Περίπτωση 2: $\lambda_1 > 0$ Από (3) $\implies x_1 = 4$. Στον κύκλο: $16 + x_2^2 = 25 \implies x_2^2 = 9 \implies x_2 = 3$ (αφού $x_2 > 0$). Ελεγχος πολλαπλασιαστών: (2) $\implies 2\lambda_2(3) = 5 \implies \lambda_2 = 5/6 > 0$ (OK). (1) $\implies \lambda_1 + 2(5/6)(4) = 9 \implies \lambda_1 + 40/6 = 9 \implies \lambda_1 + 20/3 = 27/3 \implies \lambda_1 = 7/3 > 0$ (OK).

Άρα η βέλτιστη λύση είναι: $x_1^* = 4, \quad x_2^* = 3$. Μέγιστο Κέρδος: 51.

Θέμα 3 (4 μονάδες)

Να διερευνήσετε αν είναι εφικτή η ελαχιστοποίηση της συνάρτησης $f_1(x_1, x_2)$ ως προς $x = [x_1, x_2]^T \in S_1 \subset \mathbb{R}^2$ και της συνάρτησης $f_2(x_1, x_2)$ ως προς $x = [x_1, x_2]^T \in S_2 \subset \mathbb{R}^2$ όταν:

α) (2 μονάδες) $f_1(x_1, x_2) = x_1^2 + x_2^2 - 3x_1 - 3x_2 + \frac{9}{2}$ και

$$S_1 = \{x \in \mathbb{R}^2 : x_1 - x_2 \leq 0 \text{ και } x_1 + x_2 - 2 \geq 0 \text{ και } x_1 - x_2 + 2 \leq 0 \text{ και } x_1 - 2 \leq 0 \text{ και } x_2 \geq 0\}$$

β) (2 μονάδες) $f_2(x_1, x_2) = x_1^2 + x_2^2 - 4x_1 + 4$ και

$$S_2 = \{x \in \mathbb{R}^2 : x_1^2 + x_2^2 - 2x_1 - 4x_2 + 1 \leq 0\}$$

Λύση**Λύση:**

Εξετάζουμε την ύπαρξη ελαχίστου με βάση το **Θεώρημα Weierstrass** (Compactness) ή μέσω της ιδιότητας coercivity (για μη φραγμένα σύνολα).

α) Μελέτη για f_1, S_1

$$f_1(x_1, x_2) = (x_1 - 1.5)^2 + (x_2 - 1.5)^2 + \text{σταθ.}$$

Η f_1 είναι συνεχής και **coercive** ($\lim_{||x|| \rightarrow \infty} f_1(x) = \infty$). Το σύνολο S_1 ορίζεται από: $x_2 \geq x_1 + 2$, $x_1 \leq 2$, $x_2 \geq 0$. Περιέχει π.χ. την ημιευθεία $x_1 = 0, x_2 \geq 2$, άρα είναι **μη φραγμένο**. Είναι όμως **κλειστό** σύνολο (τομή κλειστών ημιεπιπέδων). Επειδή η f_1 είναι coercive και το S_1 κλειστό (και μη κενό), το ολικό ελάχιστο **υπάρχει**. **Απάντηση:** Η ελαχιστοποίηση είναι εφικτή.

β) Μελέτη για f_2, S_2

$$f_2(x_1, x_2) = (x_1 - 2)^2 + x_2^2$$

Το σύνολο S_2 είναι:

$$(x_1 - 1)^2 + (x_2 - 2)^2 \leq 4$$

Αυτό περιγράφει έναν δίσκο (με το σύνορο του). Ως δίσκος, το S_2 είναι σύνολο **κλειστό** και **φραγμένο**, άρα είναι **συμπαγές**. Η f_2 είναι συνεχής συνάρτηση. Σύμφωνα με το **Θεώρημα Weierstrass**, κάθε συνεχής συνάρτηση σε συμπαγές σύνολο λαμβάνει ελάχιστη και μέγιστη τιμή. **Απάντηση:** Η ελαχιστοποίηση είναι εφικτή.