Διάλεξη 14

Προβλήματα με ανισοτικούς περιορισμούς, Τρίτη 19/4/18

Παράδειγμα 1

Θεωρείστε το πρόβλημα με έναν ισοτικό και έναν ανισοτικό περιορισμό:

$$\min x_1^2 - \log x_2 + x_3^2$$

έτσι ώστε $x_1 + x_2 \leq 1$

$$x_1 + x_3 = 2$$

$$x_2 > 0, x_1, x_3 \in \mathbb{R}$$

Θα το λύσουμε με τη μέθοδο Lagrange αφού το μετατρέψουμε σε πρόβλημα μόνο με ισοτικούς περιορισμούς με βάση την εξής ισοδυναμία:

$x_1 + x_2 \leq 1 \Leftrightarrow x_1 + x_2 + z_1 = 1$ για κάποιο $z_1 \geq 0$

Το πρόβλημα βελτιστοποίησης είναι ισοδύναμο με:

$$\min x_1^2 - \log x_2 + x_3^2$$

έτσι ώστε $x_1 + x_2 + z_1 = 1$

$$x_1 + x_3 = 2$$

$$z_1 \geq 0, x_2 > 0, x_1, x_3 \in \mathbb{R}$$

Απο αυτό το σημείο και μετά χρησιμοποιούμε τη μέθοδο Lagrange:
Η Λαγκραντζιανή είναι

$$L(x_1, x_2, x_3, z_1; \lambda_1, \lambda_2) = x_1^2 - \log x_2 + x_3^2 + \lambda_1(x_1 + x_2 + z_1 - 1) + \lambda_2(x_1 + x_3 - 2)$$

Παρατηρήστε ότι εάν $\lambda_1 \leq 0$ τότε $\min\limits_{x_1, x_2, x_3 \in \mathbb{R}, z_1 \geq 0} L(x_1, x_2, x_3, z_1; \lambda_1, \lambda_2) = -\infty$, συνεπώς θα πρέπει $\lambda_1 > 0$, δηλαδή $\Lambda = \{(\lambda_1, \lambda_2) \in \mathbb{R}^2 : \lambda_1 > 0\}$.

Θέτοντας $\frac{\partial}{\partial x_i} L(\vec{x}, z_1; \vec{\lambda}) = 0$ για $i = 1, 2, 3$, παίρνουμε $x_1 = -\frac{\lambda_1 + \lambda_2}{2}, x_2 = 1/\lambda_1, x_3 = -\frac{\lambda_2}{2}$. Αφού η Λαγκραντζιανή είναι κυρτή συνάρτηση των

x_1, x_2, x_3, από την ικανή συνθήκη η ελάχιστη τιμή λαμβάνεται για $x_1(\vec{\lambda}) = -\frac{\lambda_1 + \lambda_2}{2}, x_2(\vec{\lambda}) = 1/\lambda_1, x_3(\vec{\lambda}) = -\frac{\lambda_2}{2}$.

Για να ελαχιστοποιήσουμε τη Λαγκραντζιανή ως προς z_1, λύνουμε το πρόβλημα $\min_{z_1 \geq 0} \lambda_1 z_1$ όπου εφόσον $\lambda_1 > 0$ η βέλτιστη λύση είναι $z_1(\vec{\lambda}) = 0$.
Τέλος βρίσκουμε τις τιμές των $(\lambda_1, \lambda_2) \in \Lambda$ που δίνουν εφικτή (και άρα τη βέλτιστη -από το Θεώρημα…-) λύση:
$x_1(\vec{\lambda}) + x_2(\vec{\lambda}) + z_1(\vec{\lambda}) = 1, x_1(\vec{\lambda}) + x_3(\vec{\lambda}) = 2$. Αντικαθιστώντας τις πιο πάνω εκφράσεις, παίρνουμε $-\frac{\lambda_1 + \lambda_2}{2} + \frac{1}{\lambda_1} = 1, -\frac{\lambda_1 + \lambda_2}{2} - \frac{\lambda_2}{2} = 2$. Λύνοντας ως προς λ_1, λ_2 βρίσκουμε $(\lambda_1^*, \lambda_2^*) = (2, -3) \in \Lambda$.

Συνεπώς η βέλτιστη λύση είναι $x_1(\vec{\lambda}^*) = x_2(\vec{\lambda}^*) = 1/2, x_3(\vec{\lambda}^*) = 3/2$.

Παράδειγμα 2

Ας θεωρήσουμε το εξής πρόβλημα

$$\min x_1^2 - \log x_2 + x_3^2$$

έτσι ώστε $x_1 + x_2 \leq 1$

$$x_1 + x_3 \leq 2$$

$$x_2 > 0, x_1, x_3 \in \mathbb{R}$$

με δύο ανισοτικούς περιορισμούς.
Όπως στο προηγούμενο παράδειγμα μετασχηματίζουμε σε πρόβλημα μόνο με ισοτικούς περιορισμούς, εισάγωντας μια μεταβλητή χαλαρότητας για κάθε περιορισμό:

$$\min x_1^2 - \log x_2 + x_3^2$$

έτσι ώστε $x_1 + x_2 + z_1 = 1$

$$x_1 + x_3 + z_2 = 2$$

$$z_1, z_2 \geq 0, x_2 > 0, x_1, x_3 \in \mathbb{R}$$

Η Λαγκραντζιανή είναι

$L(x_1, x_2, x_3, z_1; \lambda_1, \lambda_2) = x_1^2 - \log x_2 + x_3^2 + \lambda_1(x_1 + x_2 + z_1 - 1) + \lambda_2(x_1 + x_3 + z_2 - 2)$

Όπως στο προηγούμενο παράδειγμα, πρέπει $\lambda_1 > 0$. Εδώ επίσης θα πρέπει $\lambda_2 \geq 0$, γιατί σε διαφορετική περίπτωση $L \to -\infty$ καθώς $z_2 \to +\infty$. Συνεπώς, $\Lambda = \{(\lambda_1, \lambda_2) \in \mathbb{R}^2 : \lambda_1 > 0, \lambda_2 \geq 0\}$.

Η εξίσωση των μερικών παραγώγων της Λαγκραντζιανής ως προς x_1, x_2, x_3 δίνει $x_1(\vec{\lambda}) = -\dfrac{\lambda_1 + \lambda_2}{2}, x_2(\vec{\lambda}) = 1/\lambda_1, x_3(\vec{\lambda}) = -\dfrac{\lambda_2}{2}$. Από το προηγούμενο παράδειγμα, λαμβάνουμε πάλι $z_1(\vec{\lambda}) = 0$ αφού $\lambda_1 > 0$.

Η ελαχιστοποίηση της Λαγκραντζιανής ως προς z_2 μας οδηγεί στο πρόβλημα $\min\limits_{z_2 \geq 0} \lambda_2 z_2$.
Για να βρούμε τη βέλτιστη λύση διακρίνουμε δύο περιπτώσεις:

1. Εάν $\lambda_2 > 0$ η βέλτιστη λύση είναι $z_2(\vec{\lambda}) = 0$.
2. Εάν $\lambda_2 = 0$ *τότε κάθε* $z_2(\vec{\lambda}) \geq 0$ *είναι βέλτιστο.*

Με λόγια: εάν $\lambda_2 > 0$ τότε η βέλτιστη λύση της Λαγκραντζιανής πρέπει να ικανοποιεί τον 2ο περιορισμό με ισότητα. Αν $\lambda_2 = 0$ τότε ο 2ος περιορισμός δεν επηρεάζει την ελαχιστοποίηση της Λαγκραντζιανής.

Μπορούμε να εκφράσουμε πιο συνοπτικά τις δύο περιπτώσεις γράφοντας $\lambda_2 z_2(\vec{\lambda}) = 0$. (Παρατηρήστε ότι και για τον 1ο περιορισμό ισχύει $\lambda_1 z_1(\vec{\lambda}) = 0$.)

Παρατηρήστε ότι ενώ για τις $x_1(\vec{\lambda}), x_2(\vec{\lambda}), x_3(\vec{\lambda}), z_1(\vec{\lambda})$ προσδιορίσαμε ακριβώς πως εξαρτώνται από το $\vec{\lambda} = (\lambda_1, \lambda_2)$, αυτό δεν ήταν δυνατό για τη $z_2(\vec{\lambda})$ αφού η τιμή της είναι απροσδιόριστη όταν $\lambda_2 = 0$.

Για τη λύση των εξισώσεων

$$x_1(\vec{\lambda}) + x_2(\vec{\lambda}) + z_1(\vec{\lambda}) = 1$$

$$x_1(\vec{\lambda}) + x_3(\vec{\lambda}) + z_2(\vec{\lambda}) = 2$$

θεωρούμε δύο περιπτώσεις ανάλογα εάν $\lambda_2 > 0$ ή $\lambda_2 = 0$:

1. Εάν $\lambda_2 > 0$ τότε $z_2(\vec{\lambda}) = 0$ κα λύνοντας το σύστημα

$$-\frac{\lambda_1 + \lambda_2}{2} + \frac{1}{\lambda_1} = 1$$

$$-\frac{\lambda_1 + \lambda_2}{2} + \frac{\lambda_2}{2} = 2$$

βρίσκουμε $(\lambda, \lambda_2) = (2, -3)$ το οποίο όμως έρχεται σε αντίθεση με την υπόθεση $\lambda_2 > 0$. Άρα δεν υπάρχει $\vec{\lambda} \in \Lambda$ όπου η βέλτιστη λύση της Λαγκραντζιανής είναι εφικτή, όταν $\lambda_2 > 0$.

2. Εάν $\lambda_2 = 0$ λαμβάνουμε το σύστημα εξισώσεων

$$-\frac{\lambda_1}{2} + \frac{1}{\lambda_1} = 1$$

$$-\frac{\lambda_1}{2} + z_2 = 2$$

όπου $z_2 \geq 0$ η απροσδιόριστη τιμή $z_2(\vec{\lambda})$. Λύνοντας το σύστημα των 2 εξισώσεων και 2 αγνώστων βρίσκουμε $\lambda_1 = 1 + \sqrt{3}, z_2 = \dfrac{5 + \sqrt{3}}{2}$. Παρατηρήστε ότι $(\lambda_1, \lambda_2) = (1 + \sqrt{3}, 0) \in \Lambda$ καθώς και η λύση $x_1(\vec{\lambda}), x_2(\vec{\lambda}), x_3(\vec{\lambda}), z_1(\vec{\lambda}), z_2(\vec{\lambda})$ είναι εφικτή, άρα δίνουν τη βέλτιστη λύση του προβλήματος: $x_1^* = -\dfrac{1 + \sqrt{3}}{2}, x_2^* = \dfrac{1}{1 + \sqrt{3}}, x_3^* = 0$.

Συνθήκες Karush-Kuhn-Tucker (KKT)

Στα παραδείγματα παραπάνω χρησιμοποιήσαμε τη μέθοδο Lagrange αφού μετατρέψαμε τους ανισοτικούς περιορισμούς σε ισοτικούς με την εισαγωγή μεταβλητών χαλαρότητας. Ικανή συνθήκη για την ελαχιστοποίηση κάθε μεταβλητής χαλαρότητας ήταν $\lambda_i z_i = 0$ για τον i-οστό ανισοτικό περιορισμό. Αυτή η συνθήκη μαζί με αυτές για τις μεταβλητές απόφασης αποτελούν επίσης αναγκαίες συνθήκες που ικανοποιούν οι βέλτιστες λύσεις και ονομάζονται συνθήκες Karush-Kuhn-Tucker (KKT).

Για το πρόβλημα

$$\begin{aligned} \min \quad & f(x_1, \ldots, x_n) \\ \text{έτσι ώστε} \quad & g_i(x_1, \ldots, x_n) \leq b_i \text{ για κάθε } i = 1, \ldots, m \qquad (1) \\ & (x_1, \ldots, x_n) \in A \end{aligned}$$

η Λαγκραντζιανή (μετά την εισαγωγή των μεταβλητών χαλαρότητας $\vec{z} = (z_1, \ldots, z_m)$) είναι:

$$L(x_1, \ldots, x_n, z_1, \ldots, z_m; \lambda_1, \ldots, \lambda_m) = f(x_1, \ldots, x_n) + \sum_{i=1}^{m} \lambda_i (g_i(x_1, \ldots, x_n) + z_i - b_i)$$

ΘΕΩΡΗΜΑ ((Αναγκαίες) συνθήκες Karush-Kuhn-Tucker (KKT)):
Εάν $\vec{x}^* = (x_1^*, \ldots, x_n^*)$ *η βέλτιστη λύση του προβλήματος (1), υπάρχουν τιμές* $\vec{\lambda}^* = (\lambda_1^*, \ldots, \lambda_m^*)$ *που ικανοποιούν:*

1) ***Συνθήκες βελτίστου:*** *Το* $\vec{x}^*$ *είναι η βέλτιστη λύση του προβλήματος ελαχιστοποίησης της Λαγκραντζιανής* $\min_{\vec{x} \in A} L(\vec{x}, \vec{z}^*; \vec{\lambda}^*)$.
2) ***Συνθήκες συμπληρωματικής χαλαρότητας****:* $\lambda_i^* z_i^* = 0$ *για κάθε* $i = 1, \ldots, m$.
3) ***Συνθήκες εφικτότητας****:* $\vec{x}^* \in A, \vec{z}^* \geq 0, \vec{\lambda}^* \geq 0$ *και* $\vec{\lambda}^* \in \Lambda$ *(όπου* $\Lambda = \{\vec{\lambda} \in \mathbb{R}^m : \min_{\vec{x} \in A} L(\vec{x}, \vec{z}; \vec{\lambda}) > -\infty\}$*.)*

Σημειώστε ότι προβλήματα με ανισοτικούς και ισοτικούς περιορισμούς μπορούν να έρθουν στη μορφή (1), αφού πχ, $x - 2y^2 = 3$ είναι ισοδύναμο με τις ανισότητες $x - 2y^2 \leq 3$, $x - 2y^2 \geq 3$.

Η στρατηγική λύση προβλημάτων με ανισότητες ή/και ισότητες είναι η εξής:

1. Μετατρέπουμε τους ανισοτικούς περιορισμούς σε ισοτικούς με την εισαγωγή μεταβλητών χαλαρότητας.
2. Ακολουθούμε τη μέθοδο Lagrange, με τη διαφορά ότι για τις μεταβλητές χαλαρότητας χρησιμοποιούμε τις *συνθήκες συμπληρωματικής χαλαρότητας* και διακρίνουμε περιπτώσεις εάν $\lambda_i > 0$ ή $\lambda_i = 0$, όπως κάναμε στο Παράδειγμα 2.