B-14 Nemlineáris programozás

Szabó Viktor (L8ED48)

2009-06-01

1. Az általános nemlineáris optimalizálási feladat

Az általános nemlineáris optimalizálási feladat (Nonlinear Optimization – NLO) a következőképpen írható fel:

$$\begin{aligned} \min f(x) & \\ g_i(x) = 0 & \quad i \in I = \{1, \dots, m\} \\ h_j(x) \leq 0 & \quad j \in J = \{1, \dots, p\} \\ x \in \mathcal{C} & \end{aligned} \tag{1}$$

ahol $x \in \mathbb{R}^n$, $\mathcal{C} \subseteq \mathbb{R}^n$ egy adott halmaz és az $f, g_1, \dots, g_m, h_1, \dots, h_p$ függvények értelmezési tartománya $\mathcal{C}$ (vagy olyan nyílt halmaz, amely tartalmazza $\mathcal{C}$-t). A feltételeket kielégítő pontokat *megengedett megoldásoknak nevezzük*, halmazukat $\mathcal{F}$-fel jelöljük, vagyis

$$\mathcal{F} = \{x \in \mathcal{C} \,:\, g_i(x) = 0,\, h_j(x) \leq 0,\, \forall i = 1, \dots, m,\, j = 1, \dots, p\}.$$

2. Konvex programozás

1. Definíció. *Egy $\mathcal{C} \subseteq \mathbb{R}^n$ halmaz* konvex*, ha bármely $x^1, x^2 \in \mathcal{C}$ és $0 \leq \lambda \leq 1$ esetén az*

$$x = \lambda x^1 + (1 - \lambda) x^2$$

vektor is eleme $\mathcal{C}$-nek. A $\lambda x^1 + (1 - \lambda) x^2$ vektor az x^1, x^2 pontok konvex kombinációja.

Másként fogalmazva: egy konvex halmaz tetszőleges két pontját összekötő szakasz is a konvex halmazban van.

2. Definíció. *Egy konvex $\mathcal{C} \subseteq \mathbb{R}^n$ halmazon értelmezett $f : \mathcal{C} \to \mathbb{R}$ függvény* konvex*, ha minden $x^1, x^2 \in \mathcal{C}$ és $0 \leq \lambda \leq 1$ esetén*

$$f(\lambda x^1 + (1-\lambda)x^2) \leq \lambda f(x^1) + (1-\lambda) f(x^2).$$

3. Definíció. *Az f függvény* gradiense*, amit $\nabla f(x)$ jelöl, a parciális deriváltak vektora:*

$$(\nabla f(x))_j = \frac{\partial f(x)}{\partial x_j} \quad (j = 1, \ldots, n).$$

4. Definíció. *Az f függvény* Hesse-mátrixa *a másodrendű parciális deriváltakból áll:*

$$(\nabla^2 f(x))_{ij} = \frac{\partial^2 f(x)}{\partial x_i \partial x_j} \quad (i = 1, \ldots, n, j = 1, \ldots, n).$$

5. Definíció. *Egy (1) alakú feladatot* konvex optimalizálásnak (CO) *nevezünk, ha $f, g_1, \ldots, g_m$ konvex függvények, $h_1, \ldots, h_p$ affin (lineáris) függvények és $\mathcal{C}$ konvex halmaz.*

1. Lemma (Szükséges és elégséges feltétel a konvexitásra). *Ha $\mathcal{C}$ konvex, nyílt halmaz, és f kétszer folytonosan differenciálható $\mathcal{C}$-n, akkor a következők ekvivalensek:*

- *f konvex $\mathcal{C}$-n*
- *f Hesse-mátrixa pozitív szemidefinit*

1. Tétel (Megoldás létezésének szükséges és elégséges feltétele). *Ha $\mathcal{F}$ (a megengedett megoldások halmaza) nyílt, és f folytonosan differenciálható $\mathcal{C}$-n, akkor $\bar{x} \in \mathbb{R}^n$ pontosan akkor optimális megoldás, ha $\nabla f(\bar{x}) = 0$.*

3. Feltételes optimalizálás

6. Definíció. *Egy (1) alakú feladatot* feltétel nélküli optimalizálásnak *nevezünk, ha az I és J indexhalmazok üresek, valamint $\mathcal{C} = \mathbb{R}^n$. Ellenkező eseben a feladatot* feltételes optimalizálásnak nevezzük.

3.1. Egyenlőség-feltételes optimalizálás: a Lagrange-féle multiplikátoros módszer

Tekintsünk egy (1) alakú komplex optimalizálási (CO) feladatot, amelyben J üres halmaz, I nem:

$$\begin{aligned} \min f(x) & \\ g_i(x) = 0 & \quad i \in I = \{1, \dots, m\} \\ x \in \mathcal{C} & \end{aligned} \tag{2}$$

ahol $\mathcal{C} \subseteq \mathbb{R}^n$ egy konvex, nyílt halmaz és az $f, g_1, \dots, g_m$ konvex függvények $\mathcal{C}$-n. A megengedett megoldások halmaza

$$\mathcal{F} = \{x \in \mathcal{C} : g_i(x) = 0; i = 1, \dots, m\}. \tag{3}$$

7. Definíció. *Vezessünk be új változókat (λ_i); ezek segítségével definiálhatjuk az $L : \mathbb{R}^{m+n} \to \mathbb{R}$* Lagrange-függvényt *a következőképpen:*

$$L(x, \lambda) = f(x) + \sum_{i=1}^{m} \lambda_i g_i(x) \quad (i = 1, \dots, m).$$

Először *szükséges feltételeket* adunk meg arra, hogy egy vektor a fenti feladatnak lokális szélsőérték helye legyen:

2. Tétel (Lagrange-tétel). *Tegyük fel, hogy az $f, g_1, \dots, g_m$ függvények folytonosan differenciálhatóak egy, az $\mathcal{F}$ halmazt tartalmazó nyílt halmazon. Ha $\bar{x} \in \mathbb{R}^n$ a (2) feladat lokális szélsőértékhelye és a $\nabla g_i(\bar{x})$ gradiensvektorok lineárisan függetlenek, akkor van olyan $\bar{\lambda} = (\bar{\lambda}_1, \dots, \bar{\lambda}_n) \in \mathbb{R}^n$ multiplikátorvektor, hogy $(\bar{x}, \bar{\lambda})$ stacionárius pontja a Lagrange-függvénynek, vagyis*

$$\nabla L(\bar{x}, \bar{\lambda}) = \nabla f(\bar{x}) + \sum_{i=1}^{m} \bar{\lambda}_i \nabla g_i(\bar{x}) = 0.$$

A Lagrange-függvény segítségével *elégséges feltételt* is tudunk adni arra, hogy egy $\bar{x} \in \mathcal{F}$ pont globális optimális megoldása legyen az (2) feladatnak.

3. Tétel. *Legyen $(\bar{x}, \bar{\lambda})$ a Lagrange-függvény stacionárius pontja. Ha $\bar{x}$ rögzített $\bar{\lambda}$ mellett feltétel nélküli globális minimumpontja a Lagrange-függvénynek, akkor $\bar{x}$ globális minimumpontja az (2) feladatnak.*

3.2. Egyenlőtlenség-feltételes optimalizálás: a Karush-Kuhn-Tucker optimalitási feltételek

Tekintsünk az (1) alakú komplex optimalizálási (CO) feladatot, amelyben I üres halmaz, J nem:

$$\begin{aligned} \min f(x) & & (4) \\ h_j(x) \le 0 & \quad j \in J = \{1, \dots, p\} \\ x \in \mathcal{C} & \qquad\qquad . \end{aligned}$$

ahol $\mathcal{C} \subseteq \mathbb{R}^n$ konvex halmaz és $f, g_1, \dots, g_m$ konvex függvények $\mathcal{C}$ felett (vagy egy $\mathcal{C}$ halmazt tartalmazó nyílt halmaz felett). Ekkor a megengedett megoldások halmaza

$$\mathcal{F} = \{x \in \mathcal{C} : h_j(x) \le 0, \forall j \in J\}.$$

4. Tétel (Kuhn-Tucker optimalitási feltételek). *Ha f és $h_1, \dots, h_p$ folytonosan differenciálhatók $\mathcal{C}$-n, és egy $\bar{x} \in \mathcal{F}$ pontra léteznek $\lambda_j \ge 0$ számok ($j = 1, \dots, p$), hogy*

$$\nabla f(\bar{x}) = -\sum_{j=1}^{m} \lambda_j \nabla h_j(\bar{x}), \tag{5}$$

$$\lambda_j h_j(\bar{x}) = 0 \quad (j = 1, \dots, p), \tag{6}$$

akkor $\bar{x}$ optimális megoldása a konvex optimalizálási feladatnak.

Kérdés, hogy a Kuhn-Tucker feltételek milyen feladat esetén szükségesek? Erre mutatunk egy feltételt.

8. Definíció. *Az $x^0 \in \mathcal{C}^0$ pontot a konvex optimalizálási feladat* Slater-pontjának *hívjuk, ha*

- *$h_j(x^0) < 0$, minden j-re, ha h_j nemlineáris,*
- *$h_j(x^0) \le 0$, minden j-re, ha h_j lineáris.*

Ha létezik ilyen pont, akkor azt mondjuk, hogy (CO) Slater-reguláris *vagy* (CO) *kielégíti a* Slater-feltételt.

5. Tétel. *Ha a konvex optimalizálási feladat Slater-reguláris, akkor minden optimális megoldás teljesíti a Kuhn-Tucker feltételeket.*

4. Felhasznált irodalom

1. http://www.cs.elte.hu/~tkiraly/nemlin.pdf
2. http://erettsegi.eu/index.php?dir=I.%20%E9vfolyam/Line%E1ris%20algebra%20%E9s%20programoz%E1s/&file=Lagrange.doc
3. http://web.uni-corvinus.hu/~opkut/elibd/nemlinearis_optimalizalas.pdf