A konvex programozás alapjai

Király Tamás, 2008. május 20.

Az általános konvex optimalizálási feladat a következő alakban adható meg:

$$\min f(x) \tag{1}$$
$$g_i(x) \le 0 \quad i = 1, \dots, m \tag{2}$$
$$x \in \mathcal{C} \tag{3}$$

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

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

Egy $\alpha \in \mathbb{R}$ számra a

$$\{x \in \mathcal{C} : f(x) \le \alpha\}$$

halmazt az f függvény α-hoz tartozó szinthalmazának nevezzük.

1. Állítás. *Ha f konvex függvény, akkor minden szinthalmaza konvex.*

2. Következmény. *$\mathcal{F}$ konvex halmaz.*

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

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

3. Lemma. *Ha $\mathcal{C}$ konvex nyílt halmaz, és f folytonosan differenciálható $\mathcal{C}$-n, akkor a következők ekvivalensek:*

- *f konvex a $\mathcal{C}$ halmazon.*
- *Bármely két $x, y \in \mathcal{C}$ pontra $\nabla f(x)^T(y - x) \le f(y) - f(x) \le \nabla f(y)^T(y - x)$.*
- *Ha x és $x + s$ $\mathcal{C}$-beli, akkor a $\phi(\lambda) = f(x + \lambda s)$ függvény folytonosan differenciálható $(0, 1)$-en és $\phi'(\lambda) = s^T \nabla f(x + \lambda s)$ monoton növő függvény.*

Az f függvény Hesse mátrixa a másodrendű parciális deriváltakból áll:

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

4. Lemma. *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*

A következőkben azt nézzük meg, hogy a konvex optimalizálási feladat optimális megoldásának milyen feltételeket kell teljesítenie.

5. Tétel. *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 $\nabla f(\bar{x}) = 0$.*

Bizonyítás. Ha $\nabla f(\bar{x}) = 0$, és $x \in \mathcal{F}$, akkor $f(x) - f(\bar{x}) \ge \nabla f(\bar{x})^T(x - \bar{x}) = 0$. Másik irány: ha $\nabla f(\bar{x}) \ne 0$, akkor létezik $s \in \mathbb{R}^n$ amire $\nabla f(\bar{x})^T s < 0$, azaz s irányba az iránymenti derivált negatív. De ekkor $\bar{x}$ nem lehet optimális megoldás. □

Akkor is tudunk egy szükséges és elégséges feltételt mondani, ha $\mathcal{F}$ relatív nyílt:

6. Tétel. *Ha $\mathcal{F}$ relatív nyílt, és f folytonosan differenciálható $\mathcal{C}$-n, akkor $\bar{x} \in \mathbb{R}^n$ pontosan akkor optimális megoldás, ha $\nabla f(\bar{x})^T s = 0$ minden $s \in \mathcal{L}$-re, ahol $\mathcal{L}$ az $\mathcal{F}$-et tartalmazó legszűkebb affin altér lineáris eltoltja.*

Bizonyítás. Ha $\nabla f(\bar{x})$ teljesíti a feltételt, és $x \in \mathcal{F}$, akkor $f(x) - f(\bar{x}) \geq \nabla f(\bar{x})^T(x - \bar{x}) = 0$, hiszen $x - \bar{x} \in \mathcal{L}$. Másik irány: ha van olyan $s \in \mathcal{L}$, amire $\nabla f(\bar{x})^T s < 0$, akkor s irányba az iránymenti derivált negatív. De ekkor $\bar{x}$ nem lehet optimális megoldás, hiszen $\mathcal{F}$ relatív nyíltsága miatt s irányba elég kicsit lépve $\mathcal{F}$-ben maradunk. □

Ha nem tesszük fel a relatív nyíltságot, akkor szükség van a megengedett irányok kúpjának definíciójára. Egy adott $x \in \mathcal{F}$ pontban a megengedett irányok halmaza:

$$\mathcal{D}(x) = \{s \in \mathbb{R}^n : \exists \epsilon > 0 : x + \lambda s \in \mathcal{F} \ \forall 0 < \lambda < \epsilon\}.$$

7. Állítás. *A $\mathcal{D}(x)$ halmaz egy konvex kúp minden $x \in \mathcal{F}$-re.*

8. Tétel. *Ha f folytonosan differenciálható $\mathcal{C}$-n, akkor egy $\bar{x} \in \mathbb{R}^n$ pont pontosan akkor optimális megoldása a konvex optimalizálási feladatnak, ha $\bar{x} \in \mathcal{F}$ és $\nabla f(\bar{x})^T s \geq 0$ minden $s \in \mathcal{D}(\bar{x})$-re.*

Bizonyítás. Ha $\nabla f(\bar{x})$ teljesíti a feltételt, és $x \in \mathcal{F}$, akkor $f(x) - f(\bar{x}) \geq \nabla f(\bar{x})^T(x - \bar{x}) \geq 0$, hiszen $x - \bar{x} \in \mathcal{D}(\bar{x})$. Másik irány: ha van olyan $s \in \mathcal{D}(\bar{x})$, amire $\nabla f(\bar{x})^T s < 0$, akkor s irányba az iránymenti derivált negatív. De ekkor $\bar{x}$ nem lehet optimális megoldás, hiszen s irányba elég kicsit lépve $\mathcal{F}$-ben maradunk. □

A fenti tétel hátránya, hogy alkalmazásához ismerni kell a megengedett irányok kúpjait, amit nehéz lehet minden pontra kiszámolni. Ehelyett olyan feltételeket fogunk mutatni, amik elégségesek ugyan, de nem feltétlenül szükségesek (azaz nem minden optimális megoldás teljesíti őket, de ha valami teljesíti, az optimális). Ezek a Kuhn-Tucker optimalitási feltételek.

9. Tétel (Kuhn-Tucker optimalitási feltételek). *Ha f és g_j $(j = 1, \ldots, m)$ folytonosan differenciálhatók $\mathcal{C}$-n, és egy $\bar{x} \in \mathcal{F}$ pontra léteznek $\lambda_j \geq 0$ számok $(j = 1, \ldots, m)$, hogy*

$$\nabla f(\bar{x}) = -\sum_{j=1}^{m} \lambda_j \nabla g_j(\bar{x}), \tag{4}$$

$$\lambda_j g_j(\bar{x}) = 0 \quad (j = 1, \ldots, m), \tag{5}$$

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

Bizonyítás. Megmutatjuk, hogy $\bar{x}$-ban teljesül a 8. Tétel feltétele. Legyen $s \in \mathcal{D}(\bar{x})$. Ekkor $g_j(\bar{x}) = 0$ esetén $\nabla g_j(\bar{x})^T s \leq 0$, hiszen s irányba nem nőhet g_j. Így

$$\nabla f(\bar{x})^T s = -\sum_{j=1}^{m} \lambda_j \nabla g_j(\bar{x})^T s = -\sum_{j : g_j(\bar{x}) = 0} \lambda_j \nabla g_j(\bar{x})^T s \geq 0.$$

□

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

10. Definíció. *Egy $x^0 \in \mathcal{F}$ pont a konvex optimalizálási feladat Slater pontja, ha $g_j(x^0) < 0$ minden nemlineáris g_j-re. A konvex optimalizálási feladat Slater-reguláris, ha van Slater pontja.*

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

Végezetül kimondunk egy Farkas lemma jellegű tételt is a konvex optimalizálási feladatra.

12. Tétel. *Ha a konvex optimalizálási feladat Slater-reguláris, és $\alpha \in \mathbb{R}$, akkor az $\{x \in \mathcal{F} : f(x) < \alpha\}$ halmaz pontosan akkor üres, ha léteznek $\lambda_j \geq 0$ számok $(j = 1, \ldots, m)$, hogy*

$$f(x) + \sum_{j=1}^{m} \lambda_j g_j(x) \geq \alpha \quad \forall \ x \in \mathcal{C}.$$