کنترل بهینه

جلسه سوم ۹۹/۷/۵

مسائل بهینه‌سازی مقید با قیود تساوی

$$\min \quad f(x)$$

$$\text{Subject to} \quad h_i(x) = 0 \qquad i = 1, \dots, m$$

$$f, h_i : \mathbb{R}^n \longrightarrow \mathbb{R}$$

$$h = (h_1, \dots, h_m) : \mathbb{R}^n \longrightarrow \mathbb{R}^m$$

تعریف - $S = \{x \in \mathbb{R}^n : h(x) = 0\}$ $\bar{x} \in$ منظم گوییم هرگاه $\nabla h_1(\bar{x}), \dots, \nabla h_m(\bar{x})$ مستقل خطی باشند یا

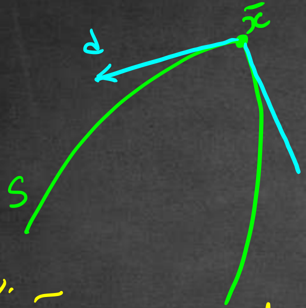
به‌عبارت دیگر ماتریس مشتق $Dh(\bar{x})$ در رابطه

$$\operatorname{rank}(Dh(\bar{x})) = m$$

صدق کند. این شرط معادل این است که S در همسایگی $\bar{x}$ یک رویه با نقص بعد m (یا بعد $n-m$)

خواهد بود.

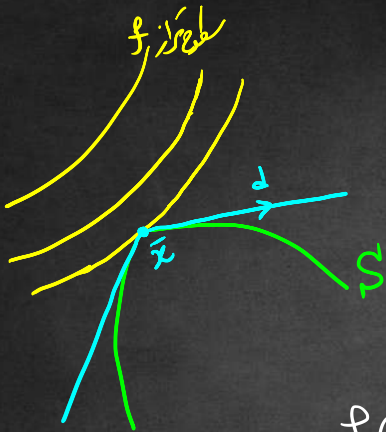
تعریف- $T_S(\bar{x}) = \{ d \in \mathbb{R}^n : \mathrm{dist}(\bar{x} + \epsilon \vec{d}, S) = o(\epsilon) \}$ مخروط مماسی S در نقطه $\bar{x}$



به طور معادل به ازای بردار $d \in \mathbb{R}^n$ خم $\gamma : [0,1] \to S$
وجود داشته باشد که $\gamma(0) = \bar{x}$ و $\gamma'(0) = d$.

به سادگی می‌توان دید که $T_S(\bar{x})$ یک مخروط است یعنی اگر $d \in T_S(\bar{x})$ و $\lambda \geq 0$ آنگاه
$\lambda d \in T_S(\bar{x})$.

اگر $\bar{x} \in S$ یک نقطه منظم باشد، $T_S(\bar{x})$ یک زیرفضای برداری است و به آن فضای مماس رویه S
در نقطه $\bar{x}$ گفته می‌شود.



شرط لازم هندسی بهینگی : اگر $\bar{x}$ نقطه بهینه باشد،

$$D_0(\bar{x}) \cap T(\bar{x}) = \phi$$

خم $\gamma: I \subseteq \mathbb{R} \to S$ را در نظر بگیرید که $\gamma(0) = \bar{x}$.

$$f(\gamma(0)) = f(\bar{x}) = \min_{t \in I} f(\gamma(t))$$

$$f(\gamma(t)) = f(\gamma(0)) + \underbrace{[\nabla f(\bar{x}) \cdot \gamma'(0)]}_{\frac{d}{dt} f(\gamma(t))\big|_{t=0}} t + o(t^2)$$

$f(\gamma(0))$ مینیمم f روی S است. $\longrightarrow \nabla f(\bar{x}) \cdot \gamma'(0) \geq 0$

$$D_0(\bar{x}) = \{ d : \nabla f(\bar{x}) \cdot d < 0 \}$$

اگر $\bar{x}$ منظم باشد، شرط $D_0(\bar{x}) \cap T(\bar{x}) = \emptyset$ معادل است با اینکه

$$(1) \qquad \nabla f(\bar{x}) \perp T(\bar{x})$$

چون وقتی $\bar{x}$ منظم است، $T(\bar{x})$ یک فضای برداری است و اگر $d \in T(\bar{x})$ آنگاه باید $-d \in T(\bar{x})$. پس $\nabla f(\bar{x}) \cdot d \geq 0$ و $\nabla f(\bar{x}) \cdot (-d) \geq 0$ هم‌زمان برقرار است. درنتیجه $\nabla f(\bar{x}) \cdot d = 0$.

نکته- وقتی $\bar{x}$ منظم است، $T(\bar{x}) = \{ d \in \mathbb{R}^n : \nabla h_i(\bar{x}) \cdot d = 0 \ , \ i=1,\dots,m \}$

$$= \{ d \in \mathbb{R}^n : Dh(\bar{x})(d) = 0 \}$$

$$Dh(\bar{x}) : \mathbb{R}^n \longrightarrow \mathbb{R}^m$$

شرط (1) معادل این است که $\nabla f(\bar{x}) \perp \mathrm{Nul}\left(Dh(\bar{x})\right)$

(یادآوری از جبرخطی : $(\mathrm{Nul}\, A)^{\perp} = \mathrm{Im}\, A^{T}$)

درنتیجه $\nabla f(\bar{x}) \in \mathrm{Im}\left(\left(Dh(\bar{x})\right)^{T}\right)$ یا به طور معادل بردار $\lambda \in \mathbb{R}^{m}$ وجود دارد که

$$Dh(\bar{x})^{T} \cdot \lambda = -\nabla f(\bar{x})$$

$$\Rightarrow \quad \nabla f(\bar{x}) + Dh(\bar{x})^{T} \cdot \lambda = 0 \qquad \lambda = (\lambda_1, \dots, \lambda_m)^{T}$$

$$\nabla f(\bar{x}) + \lambda_1 \nabla h_1(\bar{x}) + \cdots + \lambda_m \nabla h_m(\bar{x}) = 0$$

تعریف- وقتی مسأله بهینه‌سازی

$$\begin{cases} \min f(x) \\ \text{subject to } \; h(x)=0 \end{cases}$$

را داریم، تابع

$$L(x,\lambda) = f(x) + \lambda \cdot h(x) \; : \; \mathbb{R}^n \times \mathbb{R}^m \longrightarrow \mathbb{R}$$

را لاگرانژین می‌نامیم. شرط لازم (۱) معادل این است که

$$D_x \mathcal{L}(\bar{x},\bar{\lambda}) = 0 \quad , \quad D_\lambda \mathcal{L}(\bar{x},\bar{\lambda}) = 0$$

مقدار $\bar{\lambda}$ که تساوی بالا برای آن برقرار باشد را ضریب لاگرانژین می‌نامیم.

تساوی $D_\lambda \mathcal{L} = 0$ قید $h(x)=0$ را بیان می‌کند و تساوی $D_x \mathcal{L} = 0$ معادل شرط لازم (۱) است.

شرط لازم مرتبه دوم: اگر $\bar{x}$ جواب مینیمم و منظم باشد، آنگاه بردار $\bar{\lambda} \in \mathbb{R}^m$ وجود دارد که

$$\nabla f(\bar{x}) + Dh(\bar{x})^T \bar{\lambda} = 0$$

$$(2) \qquad d^T D^2 f(\bar{x}) d + \bar{\lambda} \underset{\text{ضرب داخلی در } \mathbb{R}^m}{\cdot} D^2 h(\bar{x})[d,d] \geq 0 \qquad \forall \underbrace{d \in T(\bar{x})}_{Dh(\bar{x}) \cdot d = 0}$$

تذکر: $h : \mathbb{R}^n \to \mathbb{R}^m$ بنابراین $D^2 h : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^m$ یک فرم دوخطی است

$$D^2 h\, [d,d] \in \mathbb{R}^m$$

در واقع داریم:

$$\bar{\lambda} \cdot D^2 h(\bar{x})[d,d] = d^T \left(\sum_{i=1}^{m} \lambda_i D^2 h_i(\bar{x}) \right) d$$

فرض $\gamma : I \longrightarrow S$ در نظر بگیرید که $\gamma(0) = \bar{x}$, $\gamma'(0) = d$ و قرار دهید $\varphi(t) = f(\gamma(t))$

$$\varphi(0) = \min_{t \in I} \varphi(t) \Longrightarrow \begin{cases} \varphi'(0) = 0 \\ \varphi''(0) \geq 0 \end{cases}$$

$$\varphi'(t) = Df(\gamma(t)) . \gamma'(t)$$

$$\varphi''(t) = D^2 f(\gamma(t)) \, [\gamma'(t), \gamma'(t)] + Df(\gamma(t)) . \gamma''(t)$$

$$0 \leq \varphi''(0) = \underbrace{D^2 f(\bar{x}) \, [d, d]}_{d^T \, D^2 f(\bar{x}) \, d} + \nabla f(\bar{x}) \cdot \gamma''(0)$$

از شرط لازم مرتبه اول $(\varphi'(0) = 0)$ داریم:

$$\nabla f(\bar{x}) + Dh(\bar{x})^T . \lambda = 0$$

از طرفی $h(\gamma(t)) = 0$ برای هر $t \in I$. (چون $\gamma(t) \in S$)

$$Dh(\gamma(t))\,\gamma'(t) = 0$$

$$D^2h(\gamma(t))[\gamma'(t), \gamma'(t)] + Dh(\gamma(t))\,\gamma''(t) = 0$$

$$D^2h(\bar{x})[d, d] + Dh(\bar{x})\cdot\gamma''(0) = 0$$

برگردیم به محاسبه $\varphi''(0)$:

$$\nabla f(\bar{x}) \cdot \gamma''(0) = -\left(Dh(\bar{x})^T \lambda\right) \cdot \gamma''(0)$$

$$= -\lambda^T Dh(\bar{x})\,\gamma''(0)$$

$$= \lambda^T D^2h(\bar{x})[d, d]$$

تذکر: رابطه (2) معادل این است که

$$d^T D_x^2 \mathcal{L}(\bar{x},\bar{\lambda})\, d \geq 0 \qquad d \in T(\bar{x})$$

از طرفی داریم $\mathcal{L} : \mathbb{R}^n \times \mathbb{R}^m \longrightarrow \mathbb{R}$ برای $\bar{\lambda}$ ثابت لاگرانژین یک تابع از $\mathbb{R}^n$ به $\mathbb{R}^n$

است و مشتق دوم آن باید نیمه‌معین مثبت باشد روی زیرفضای $T(\bar{x})$.

این مطلب معادل این است که به جای ماتریس $D_x^2\mathcal{L}$ عملگر $P_{T(\bar{x})}[D_x^2\mathcal{L}]$ را درنظر بگیریم

$$P_{T(\bar{x})}[D_x^2\mathcal{L}] : T(\bar{x}) \longrightarrow T(\bar{x})$$

شرط لازم (2) معادل این است که عملگر بالا نیمه‌معین مثبت باشد یا به عبارت هم مقادیر ویژه آن نامنفی باشند. برای بررسی این مطلب کافی است یک پایه برای $T(\bar{x})$ درنظر بگیرید.

فرض کنید $T(\bar{x}) = \langle e_1(\bar{x}), \dots, e_{n-m}(\bar{x}) \rangle$.

آنگاه ماتریس نمایش $P_{T(\bar{x})}[D_x^2 \mathcal{L}]$ در این پایه برابر

$$E^T D_x^2 \mathcal{L}(\bar{x},\bar{\lambda})\, E$$

است که $E = [e_1 | \cdots | e_{n-m}]$

مثال-

$$f(x) = x_1 + x_2$$

$$h(x) = x_1^2 + x_2^2 - 2$$

$$\longleftarrow \quad \mathcal{L}(x,\lambda) = x_1 + x_2 + \lambda(x_1^2 + x_2^2 - 2)$$

شرط لازم مرتبه اول:

$$\begin{cases} \partial_{x_1}\mathcal{L} = 2\bar{\lambda}\bar{x}_1 + 1 = 0 \\ \partial_{x_2}\mathcal{L} = 2\bar{\lambda}\bar{x}_2 + 1 = 0 \\ \partial_{\lambda}\mathcal{L} = \bar{x}_1^2 + \bar{x}_2^2 - 2 = 0 \end{cases}$$

$$\rightsquigarrow \begin{cases} \bar{x}_1 = \bar{x}_2 = -1, & \bar{\lambda} = 1/2 \\ \bar{x}_1 = \bar{x}_2 = 1, & \bar{\lambda} = -1/2 \end{cases}$$

$d^T \cdot D_x^2 \mathcal{L} \cdot d \geq 0$ $\quad d \cdot \nabla h(\bar{x}) = 0$ شرط لازم مرتبه دوم:

$$D_x^2 \mathcal{L} = \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} \rightsquigarrow d^T D_x^2 \mathcal{L} \cdot d = \lambda |d|^2 \geq 0$$

$T(\bar{x}) = \{ d : d \cdot \nabla h(\bar{x}) = 0 \}$ از فرض

$$= \{ d \in \mathbb{R}^2 : d \cdot (x_1, x_2) = 0 \}$$

$T(\bar{x})$ پایه برای فضای مماس $e(\bar{x}) = \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix}$

$$E = \begin{bmatrix} -x_2 \\ x_1 \end{bmatrix}$$

$$E^T D_x^2 \mathcal{L} E = [-\bar{x}_2, \bar{x}_1] \begin{bmatrix} \bar{\lambda} & 0 \\ 0 & \bar{\lambda} \end{bmatrix} \begin{bmatrix} -\bar{x}_2 \\ \bar{x}_1 \end{bmatrix} = \bar{\lambda}(\bar{x}_1^2 + \bar{x}_2^2)$$
$$= 2\bar{\lambda}$$

شرط کافی مرتبه دوم : اگر $\bar{x}$ و $\bar{\lambda}$ در شرایط

$$D_x \mathcal{L}(\bar{x}, \bar{\lambda}) = 0, \quad D_\lambda \mathcal{L}(\bar{x}, \bar{\lambda}) = 0$$

صدق کند به علاوه داشته باشیم :

$$\exists \alpha > 0 \text{ s.t.h. } \quad d^T D_x^2 \mathcal{L}(\bar{x}, \bar{\lambda})\, d \geq \alpha \|d\|^2 \quad \forall\, 0 \neq d \in T(\bar{x})$$

آنگاه $\bar{x}$ یک نقطه ی مینیمم موضعی مسئله

$$\begin{cases} \min f(x) \\ \text{subject to } \; h(x) = 0 \end{cases}$$

است.

تذکر: این شرط معادل این است که مقادیر ویژه ی $E^T D^2 \mathcal{L}\, E$ همگی اکیداً مثبت باشند.

مسأله غیرخطی در حالت کلی :

$$\min \ f(x)$$
$$\text{subject to} \quad g(x) \leq 0$$
$$h(x) = 0$$

$$f : \mathbb{R}^n \longrightarrow \mathbb{R}$$
$$g : \mathbb{R}^n \longrightarrow \mathbb{R}^k$$
$$h : \mathbb{R}^n \longrightarrow \mathbb{R}^m$$

به طور طبیعی شرط لازم هندسی این است که $D(\bar{x}) \cap F(\bar{x}) \cap T(\bar{x}) = \phi$

در این حالت لاگرانژین به صورت زیر تعریف می‌شود:

$$\mathcal{L}(x, \nu, \lambda) = f(x) + \nu \cdot g(x) + \lambda \cdot h(x)$$

شرط لازم مرتبه اول و دوم: اگر $\bar{x}$ نقطه منظم برای قیود $h=0$ و $g\geq 0$ باشد و جواب مسئله

بهینه سازی. آنگاه

$$D_x L(\bar{x},\bar{\nu},\bar{\lambda}) = \nabla f(\bar{x}) + \bar{\nu}\cdot\nabla g(\bar{x}) + \bar{\lambda}\cdot\nabla h(\bar{x}) = 0$$

$$\bar{\nu}\geq 0 \quad , \quad g(\bar{x})\leq 0$$

$$\bar{\nu}\cdot g(\bar{x}) = 0$$

$$D_\lambda L(\bar{x},\bar{\nu},\bar{\lambda}) = h(\bar{x}) = 0$$

$$d^T D_x^2 L(\bar{x},\bar{\nu},\bar{\lambda})\, d \geq 0 \qquad \forall d \in T(\bar{x}) \cap F_0(\bar{x})$$

تذکر- کلاً شرط کافی مرتبه دوم باید نامساوی مرتبه دوم بالا را به نامساوی اکید (در بعد متناهی) یا در بعد نامتناهی

درست است به جای صفر $\alpha\|d\|^2$ قرار دهیم.