کنترل بهینه

جلسه دوم ۹۹/۶/۳۱

برنامه‌ریزی غیرخطی

$$\min_{x\in S} f(x) \qquad f:\mathbb{R}^n \longrightarrow \mathbb{R}$$

$$x\in S \iff \begin{cases} g_1(x)\leq 0\ ,\ \dots\ ,\ g_k(x)\leq 0 \\ \qquad\qquad g_i:\mathbb{R}^n\longrightarrow\mathbb{R} \\ h_1(x)=\dots=h_l(x)=0\ ,\quad h_i:\mathbb{R}^n\longrightarrow\mathbb{R} \end{cases}$$

$$g(x):=(g_1,\dots,g_k)\leq 0$$

$$h(x):=(h_1,\dots,h_l)=0$$

$\min f(x)$ → object

قید نامساوی ← $g(x)\leq 0$

قید تساوی ← $h(x)=0$

→ subject

اگر نقطه $x\in S$ (یا به عبارتی در قیود بالا صدق کند) آن را شدنی (feasible) گوییم.

تعریف- نقطه $\bar{x} \in S$ می‌نیمم سراسری است هرگاه

$$f(\bar{x}) \le f(x) \quad \forall x \in S$$

و آن را می‌نیمم موضعی گوییم هرگاه همسایگی $B_r(\bar{x})$ وجود داشته باشد که

$$f(\bar{x}) \le f(x) \quad \forall x \in B_r(\bar{x})$$

شرط وجود می‌نیمم: اگر S فشرده باشد، تابع f پیوسته حتماً یک نقطه می‌نیمم در S وجود دارد.

فرض کنید $x_n \in S$ و $f(x_n) \to \inf$

بنابر فشردگی یک زیردنباله همگرا در S وجود دارد $x_{n_i} \to \bar{x}$ و بنابر پیوستگی

$$f(x_{n_i}) \longrightarrow f(\bar{x}) = \inf_S f$$

$f(\bar{x}) \le \liminf f(x_{n_i})$ به جای رابطه در بالا که از پیوستگی بدست آمده کافی است داشته باشیم

که این نامساوی تعریف نیم پیوستگی پایینی است .

به جای فشردگی نیز می توان فرض کرد که تابع f روی S از پایین کران دار است به علاوه شرط تراکمی (coercivity) برقرار باشد یعنی

$$f(x) \to \infty \quad \text{if} \quad |x| \to \infty$$

قضیه- اگر تابع $f: S \to \mathbb{R}$ نیم پیوسته پایینی باشد ، از پایین کران و در شرط تراکمی صدق کند ، آنگاه $\bar{x} \in S$ وجود دارد که

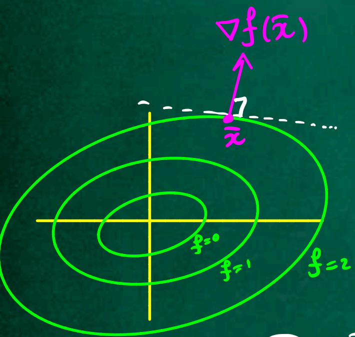
$$f(\bar{x}) = \min_{x \in S} f(x)$$

چجوری نقطه $\bar{x}$ را پیدا کنیم؟

تعریف - راستای $d \in \mathbb{R}^n$ یک جهت کاهشی در نقطه $\bar{x}$ است هرگاه تعداد $\delta > 0$ وجود داشته باشد

به طوری که

$$f(\bar{x}+td) < f(\bar{x}) \quad \forall\, t \in (0,\delta)$$



همه بردارهای کاهشی در نقطه $\bar{x}$ را با مجموعه $D(\bar{x})$ نشان می‌دهیم

اگر f در نقطه $\bar{x}$ مشتق‌پذیر باشد، آنگاه

$$D_0(\bar{x}) = \{ d \in \mathbb{R}^n : d^T \nabla f(\bar{x}) < 0 \} \subseteq D(\bar{x})$$

$$\frac{f(\bar{x}+td) - f(\bar{x})}{t} = \underbrace{\nabla f(\bar{x}) \cdot d}_{<0} + o(t) < 0$$

شرط لازم مرتبه اول بهینگی : $f : X \longrightarrow \mathbb{R}$ در نقطه $\bar{x} \in X$ مشتق‌پذیر است که X یک فضای باناخ است و $\bar{x}$ نقطه می‌نیمم موضعی f روی X است. آنگاه

$$\nabla f(\bar{x}) = 0$$

نکته - این یک شرط لازم است نه کافی. ممکن است داشته باشیم $\nabla f(\bar{x}) = 0$ ولی $\bar{x}$ یک نقطه اکسترمم نباشد.

نقاطی که در شرط $\nabla f(\bar{x}) = 0$ صدق می‌کنند نقاط ایستا (Stationary) نامیده می‌شوند هیچکدام از اینها باشند.

تعریف - اگر X یک فضای باناخ باشد، در نقطه $\bar{x}$ در راستای d دارای مشتق جهتی است هرگاه

$$f : X \longrightarrow Y \quad , \quad Df(\bar{x}, d) := \lim_{t \to 0} \frac{f(\bar{x} + td) - f(\bar{x})}{t}$$

f را در نقطه $\bar{x}$ مشتق‌پذیر به معنای گاتو می‌نامیم هرگاه عملگر خطی $Df(\bar{x}) : X \to Y$ وجود داشته باشد که

$$Df(\bar{x})(d) = Df(\bar{x}, d)$$

در واقع مشتق‌های جهتی از عملگر خطی $Df(\bar{x})$ به دست می‌آیند.

و f به معنای فرشه مشتق‌پذیر است هرگاه عملگر خطی کراندار $Df(\bar{x}) : X \to Y$ وجود داشته باشد که

$$\lim_{\|d\|_X \to 0} \frac{\|f(\bar{x}+d) - f(\bar{x}) - Df(\bar{x})\cdot d\|_Y}{\|d\|_X} = 0$$

نکته ـ اگر فضای X متناهی‌بعد باشد، شرط لازم بستگی این است که مشتق گاتو f در نقطه $\bar{x}$ برابر شوند.

شرط لازم مرتبه دوم: اگر f در $\bar{x}$ دوبار مشتق‌پذیر باشد و $\bar{x}$ نقطه‌ی مینیمم تابع f روی X باشد، آنگاه $H(\bar{x}) \geq 0$ (نیمه معین مثبت) است که $H(\bar{x})$ ماتریس هسین (ماتریس مشتق دوم f) است.

$$f : \mathbb{R}^n \longrightarrow \mathbb{R} \qquad \nabla f : \mathbb{R}^n \longrightarrow \underbrace{L(\mathbb{R}^n, \mathbb{R})} = \mathbb{R}^n$$

$$H(\bar{x}) = \begin{bmatrix} \partial_{11} f & \partial_{12} f & \cdots & \partial_{1n} f \\ \vdots & & & \\ \partial_{n1} f & \partial_{n2} f & \cdots & \partial_{nn} f \end{bmatrix}$$

$$H(\bar{x}) \geq 0 \iff \forall y \in \mathbb{R}^n \quad y^T H(\bar{x}) y \geq 0$$

$$f : X \longrightarrow Y$$

$$Df : X \longrightarrow L(X, Y)$$

$$Df(x) : X \longrightarrow Y \qquad Df(x)(u) \in Y$$

$$D^2 f : X \longrightarrow L(X, L(X,Y))$$

$$D^2 f(x) : X \xrightarrow{\text{خطی}} L(X,Y)$$

$$D^2 f(x)(u) : X \xrightarrow{\text{خطی}} Y$$

$$D^2 f(x)[u,v] \in Y$$

یک عملگر دوخطی

اگر $Y=\mathbb{R}$ ، $D^2 f$ نیمه‌معین مثبت است هرگاه

$$D^2 f(x)[u,u] \geq 0$$

و معین مثبت است هرگاه

$\exists \alpha > 0$ s.th. $D^2 f(x)[u,u] \geq \alpha \|u\|_X^2$

شرط کافی ترتیب اول : اگر $f : X \longrightarrow \mathbb{R}$ محدب و در نقطه $\bar{x} \in X$ مشتق‌پذیر باشد، به‌علاوه $\nabla f(\bar{x}) = 0$. آنگاه $\bar{x}$ یک نقطه می‌نیمم سراسری است.

تعریف- تابع f محدب است هرگاه $f(\lambda x + (1-\lambda) y) \leq \lambda f(x) + (1-\lambda) f(y)$

$$\forall x, y \in X , \quad \lambda \in [0,1]$$

و اگر f در $\bar{x}$ مشتق‌پذیر باشد معادل است با اینکه نمودار تابع f بالای خط مماس قرار بگیرد

$$f(x) \geq f(\bar{x}) + \nabla f(\bar{x}) \cdot (x - \bar{x})$$

شرط کافی مرتبه دوم: اگر f در $\bar{x}$ دوبار مشتق‌پذیر باشد و $\nabla f(\bar{x})=0$ و $D^2 f(\bar{x})$ معین مثبت باشد

آنگاه $\bar{x}$ می‌نیمم موضعی است.

$$f(\bar{x}+u) = f(\bar{x}) + \underbrace{\cancel{Df(\bar{x})\cdot u}}_{0} + \underbrace{\frac{1}{2} D^2 f(\bar{x})[u,u]}_{\geq \alpha \|u\|^2} + o(\|u\|^2)$$

برای $\|u\|$ به اندازه کافی کوچک می‌توانیم داشته باشیم

$$\alpha \|u\|^2 + o(\|u\|^2) \geq \frac{\alpha}{2} \|u\|^2$$

و برای این مقادیر u خواهیم داشت

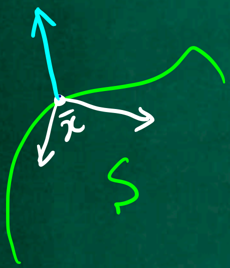
$$f(\bar{x}+u) \geq f(\bar{x})$$

مسائل بهینه‌سازی مقید:

$$\min f(x)$$
$$g(x) \le 0$$

اگر ناحیه که با قیود $g(x) \le 0$ مشخص شده باشد

تعریف- راستای d یک جهت شدنی (feasible direction) برای مجموعه S در نقطه $\bar{x}$ می‌گوییم هرگاه

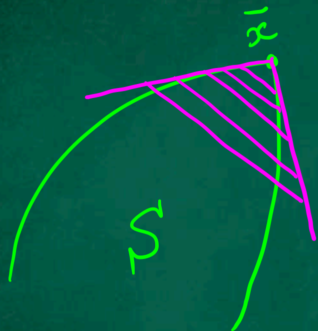


$$\exists\, \delta > 0 \ \text{s.th.} \quad \bar{x} + t d \in S \quad \forall t \in (0, \delta)$$

همه این راستاها را با نماد $F(\bar{x})$ نشان می‌دهیم.

در واقع $F(\bar{x})$ یک مخروط است به رأس مبدأ. یعنی اگر $d \in F(\bar{x})$ و

$\lambda \ge 0$ آنگاه $\lambda d \in F(\bar{x})$.



شرط لازم هندسی بهینگی: اگر $f(\bar{x}) = \min_{x \in S} f(x)$

$$F(\bar{x}) \cap D(\bar{x}) = \emptyset$$

اگر $\bar{x} \in S$ نقطه درونی S باشد، آنگاه $F(\bar{x}) = \mathbb{R}^n$.

و شرط لازم معادل این است که $D(\bar{x}) = \emptyset$ و اگر f مشتق‌پذیر باشد، $\nabla f(\bar{x}) = 0$.

سؤال: آیا توصیف خوبی از $F(\bar{x})$ می‌توانیم به دست آوریم؟

فرض کنید $S = \{x : g_i(x) \le 0 , \ i = 1, \dots, m\}$

تعریف- اگر $\bar{x} \in S$ قیدِ g_i را در نقطه $\bar{x}$ فعال (active) می‌نامیم هرگاه

$g_i(\bar{x}) = 0$ وگرنه غیرفعال (inactive) نامیده می‌شود، $g_i(\bar{x}) < 0$.

$\mathcal{A}(\bar{x}) = \{ i : \ g_i(\bar{x}) = 0 \}$ مجموعه قیودِ فعال

گزاره: $\mathcal{F}(\bar{x}) \supseteq \mathcal{F}_0(\bar{x}) = \{ d : \ \nabla g_i(\bar{x}) \cdot d < 0 \quad \forall i \in \mathcal{A}(\bar{x}) \}$

نتیجه: $\operatorname{argmin}\{ f(x) : x \in S \} \subseteq \{ \bar{x} : \ \mathcal{F}_0(\bar{x}) \cap D_0(\bar{x}) = \phi \}$

شرایط لازم KKT (Karush - Kuhn - Tucker)

تعریف - نقطه $\bar{x} \in S$ منظم است اگر $\nabla g_i(\bar{x})$ برای همه $i \in A(\bar{x})$ مستقل خطی باشند.

این تعریف تضمین می‌کند که مجموعه $\{x : g_i(x) = 0 \quad \forall i \in A(x)\}$ یک رویه با بعد $n - |A(x)|$ خواهد شد.

قضیه - اگر f و g مشتق‌پذیر باشند و $\bar{x}$ نقطه می‌نیمم نسبی در مسأله

$$\min \ f(x)$$
$$x \in S = \{ g(x) \le 0 \}$$

و $\bar{x}$ یک نقطه منظم S باشد. آنگاه بردار $\bar{\lambda}$ وجود دارد به طوری که

$$\nabla f(\bar{x}) + \bar{\nu}^T \nabla g(\bar{x}) = 0$$

$$\bar{\nu} \geq 0$$

$$g(\bar{x}) \leq 0$$

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

قضیه (Gordan) A یک ماتریس $m \times n$ است. دقیقاً یکی از دو حالت زیر اتفاق می‌افتد.

$$\exists x \in \mathbb{R}^n \quad \text{s.th.} \quad Ax < 0 \qquad (1)$$

$$\exists y \in \mathbb{R}^m, \quad y \neq 0, \quad A^T y = 0, \; y \geq 0 \qquad (2)$$

اثبات KKT:

$A = \left[\nabla f(\bar{x}) \mid \nabla g_i(\bar{x}) \quad i \in \mathcal{A}(\bar{x})\right]$ ← ماتریس $n \times (|\mathcal{A}(\bar{x})|+1)$

بنابر شرط هندسی $F_0(\bar{x}) \cap D_0(\bar{x}) = \phi$. درنتیجه حالت (۱) قضیه Gordan نمی‌تواند اتفاق بیفتد.

بنابر حالت (۲) بردار $p \geq 0$, $p \neq 0$ وجود دارد که

$$A^T p = 0 \Rightarrow p_0 \nabla f(\bar{x}) + \sum_{i \in \mathcal{A}(\bar{x})} p_i \nabla g_i(\bar{x}) = 0$$

می‌دانیم که $p_0 \neq 0$ زیرا درغیراین صورت $\{\nabla g_i(\bar{x})\}$ وابسته خطی هستند.

$$\Rightarrow \nabla f(\bar{x}) + \sum_{i \in \mathcal{A}(\bar{x})} \frac{p_i}{p_0} \nabla g_i(\bar{x}) = 0$$

$v_i = \frac{p_i}{p_0}$ if $i \in \mathcal{A}(\bar{x})$ وگرنه $v_i = 0$ و دراین صورت

$$\nabla f(\bar{x}) + v^T \nabla \bar{g}(\bar{x}) = 0$$

بنابر تعریف v به وضوح شرط $v^T g(\bar{x}) = 0$ برقرار است.