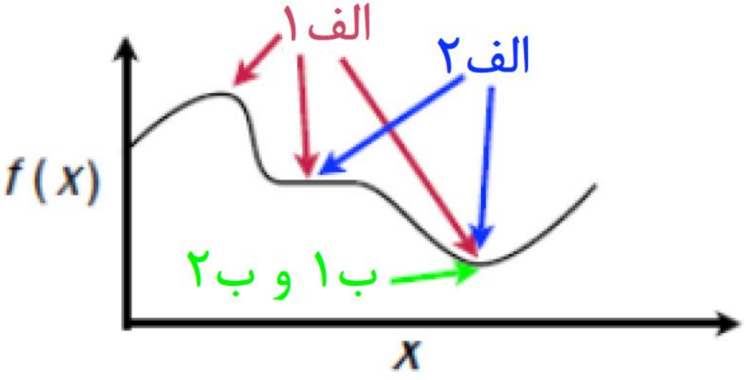
بهینه سازی
بهینه سازی مقید

محسن هوشمند
دانشکده تکنولوژی اطلاعات و علم رایانه
دانشگاه تحصیلات تکمیلی علوم پایه زنجان

شرایط بهینگی کمینه‌سازی نامقید



قضیۀ شروط لازم کمینه محلی ضعیف

- الف۱: $\nabla f(\boldsymbol{x}^*) = 0$ نقطه مانا
- الف۲: $\nabla^2 f(\boldsymbol{x}^*)$ مثبت نیمه‌معین

قضیۀ شروط کافی کمینه محلی قوی

- ب۱: $\nabla f(\boldsymbol{x}^*) = 0$
- ب۲: $\nabla^2 f(\boldsymbol{x}^*)$ مثبت معین

شرایط بهینگی بیشینه‌سازی نامقید

قضیهٔ شروط لازم بیشینه محلی ضعیف

- الف۱: $\nabla f(\boldsymbol{x}^*) = 0$ نقطه مانا
- الف۲: $\nabla^2 f(\boldsymbol{x}^*)$ منفی نیمه‌معین

قضیهٔ شروط کافی بیشینه محلی قوی

- ب۱: $\nabla f(\boldsymbol{x}^*) = 0$
- ب۲: $\nabla^2 f(\boldsymbol{x}^*)$ منفی معین

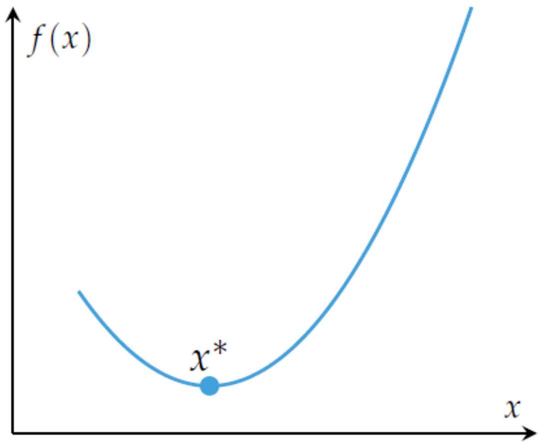
شرایط بهینگی بیشینه‌سازی نامقید

جهت بهینه‌سازی نامقید صرفا شروط اول و دوم جهت تایید نقطه بهینه

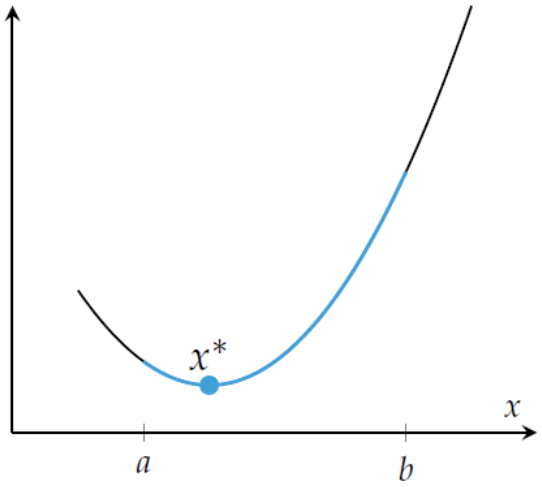
- افزوده شدن سربار منطقهٔ شدنی در بهینه‌سازی مقید
- لزوم واقع شدن نقطه بهینه در منطقهٔ شدنی

تاثیر قید و محدودیت

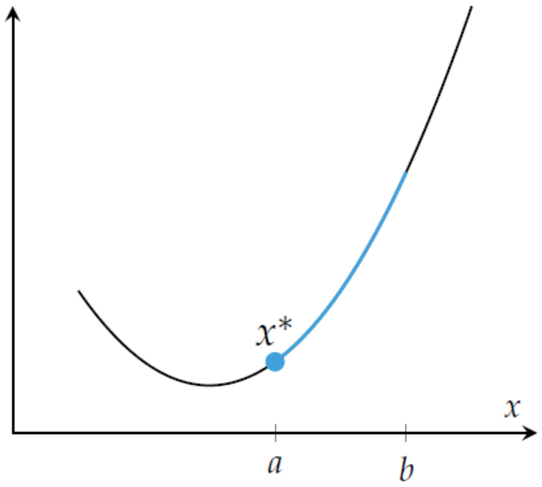
نامقید



مقید، با پاسخ یکسان



مقید، پاسخ متفاوت



Kochenderfer, Mykel J. Wheeler, Tim Allan - Algorithms for optimization-The MIT Press (2019)

تدوین ریاضی

کمینه‌سازی یا بیشینه‌سازی تابع با درنظر گرفتن قیدهایی روی متغیرهای تابع

$\boldsymbol{x}$ بردار متغیرها (یا مجهول‌ها، پارامترها)

f تابع هدف، تابعی (عددی) از x که قصد بیشینه یا کمینه کردن آن را داریم

c_i توابع نمایشگر قیدهایی شامل تساوی‌ها و نامساوی‌ها که مقادیر $\boldsymbol{x}$ پاسخ تابع f باید آن‌ها را رعایت کنند

$$\max_{\boldsymbol{x}\in\mathbb{R}^n} f(\boldsymbol{x})$$

با توجه

$$c_i(\boldsymbol{x}) = 0 \quad i \in \mathcal{E}$$

$$c_i(\boldsymbol{x}) \leq 0 \quad i \in \mathrm{I}$$

$\mathcal{E}$ و I مجموعه اندیس‌های قیدهای تساوی و نامساوی

$$\max_{\boldsymbol{x}\in\mathbb{R}^n} f(\boldsymbol{x})$$

با توجه

$$g_i(\boldsymbol{x}) = c_i,\ \ i \in \mathcal{E}$$

$$h_i(\boldsymbol{x}) \leq b_i, i \in \mathrm{I}$$

$$\min_{\boldsymbol{x}\in\mathbb{R}^n} f(\boldsymbol{x})$$

با توجه

$$g_i(\boldsymbol{x}) = c_i, i \in \mathcal{E}$$

$$h_i(\boldsymbol{x}) \geq b_i, i \in \mathrm{I}$$

مجموعه فعال

$$\mathcal{A}(\boldsymbol{x}) = \mathcal{E} \cup \{i \in I \mid c_i(\boldsymbol{x}) = 0\}.$$

تک قید تساوی

در نظر گرفتن تابعی

$f(x, y) = z$

با محدودیتی

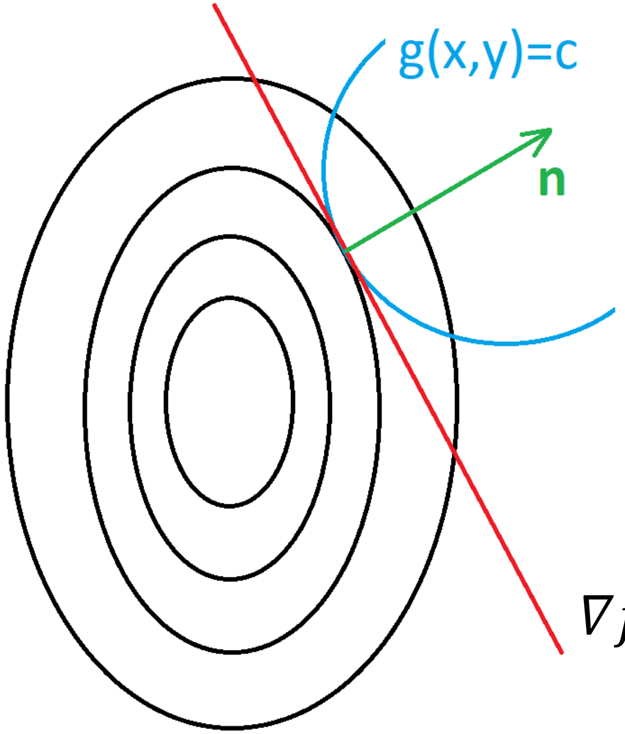
$g(x, y) = c$

f سطح و c ثابت

g سطح ترازی از سطح $g(x,y) = z$

گرادیان‌ها موازی یکدیگر

گرادیان‌های منحنی‌های تراز مضربی از یکدیگر



سطح ترازهای f

$$\nabla f(x,y) = k\nabla g(x,y)$$
$$\nabla f(x,y) = \lambda\nabla g(x,y)$$

برای سه بعد و بیشتر:

$$\nabla f(x,y,z) = \lambda\nabla g(x,y,z)$$

روش ضرائب لاگرانژ

روشی قدرتمند در حل مسائل بهینه‌سازی مقید

- بیشینه یا کمینه با استفاده از ضرائب لاگرانژ یافت می‌شود

یافتن حجم بیشینهٔ جعبه‌ای با قید اینکه هزینهٔ خرید ماده ثابت باشد.

تابع لاگرانژ

$$\mathcal{L}(\boldsymbol{x},\lambda_1) = f(\boldsymbol{x}) - \lambda_1[g(\boldsymbol{x}) - \mathrm{c}]$$

$$\nabla_{\boldsymbol{x}}\mathcal{L}(\boldsymbol{x}^*,\lambda_1^*) = 0$$

شرط لازم مرتبه اول

قضیه لاگرانژ

f و g توابع مشتق پذیری هستند. به طوری‌که f دارای نقطه اکسترمم در $(\boldsymbol{x}^*)$ بر سطح منحنی $g(\boldsymbol{x}) = c$ است. اگر $\nabla g(\boldsymbol{x}^*) \neq 0$، آن‌گاه عدد λ وجود دارد به طوری که $\nabla f(\boldsymbol{x}^*) = \lambda \nabla g(\boldsymbol{x}^*)$.

λ را ضریب **لاگرانژ** خوانند.

مثال

$$Max\ f\ (x,y) = 4xy$$

$$g(x,y) = \frac{x^2}{3^2} + \frac{y^2}{4^2} = 1$$

حل

$$\nabla f(x,y) = \lambda \nabla g(x,y)$$

$$\Longrightarrow 4y\boldsymbol{i} + 4x\boldsymbol{j} = \lambda\ \frac{2x}{9}\boldsymbol{i} + \lambda \frac{y}{8}\boldsymbol{j}$$

$$\Longrightarrow \begin{cases} 4y = \lambda\ \frac{2x}{9} \\ 4x = \lambda \frac{y}{8} \\ \frac{x^2}{3^2} + \frac{y^2}{4^2} = 1 \end{cases}$$

$$\lambda = \frac{18y}{x} \xRightarrow{\text{جانشینی}} 4x = \lambda \frac{y}{8}$$

$$4x = \lambda \frac{y}{8} = \frac{18y}{x}\frac{y}{8} \Longrightarrow 4x = \frac{9y^2}{4x}$$

$$\Longrightarrow \frac{x^2}{9} = \frac{y^2}{16}$$

$$\frac{x^2}{3^2} + \frac{y^2}{4^2} = 1 \Longrightarrow \frac{y^2}{16} + \frac{y^2}{16} = 1$$

$$\Longrightarrow \mathrm{y} = 2\sqrt{2} \Longrightarrow \mathrm{x} = \frac{3\sqrt{2}}{2}$$

$$\Longrightarrow fmax = 24$$

مثال ۲

$$f(x,y) = x + 2y - 2z$$

$$x^2 + 2y^2 + 4z^2 = 1$$

$$g(x,y) = x^2 + 2y^2 + 4z^2$$

$$\nabla f(x,y) = \lambda \nabla g(x,y)$$

$$\Longrightarrow \boldsymbol{i} + 2\boldsymbol{j} - 2\boldsymbol{k} = \lambda\,(2x\boldsymbol{i} + 4y\boldsymbol{j} + 8z\boldsymbol{k})$$

$$\Longrightarrow \begin{cases} 1 = 2x\lambda \rightarrow x = \frac{1}{2\lambda} \\ 2 = 4y\lambda \rightarrow y = \frac{1}{2\lambda} \\ -2 = 8z\lambda \rightarrow z = -\frac{1}{4\lambda} \\ x^2 + 2y^2 + 4z^2 = 1 \end{cases}$$

$${\frac{1}{2\lambda}}^2 + 2{\frac{1}{2\lambda}}^2 + 4\left(-\frac{1}{4\lambda}\right)^2 = 1 \Longrightarrow \lambda = \pm 1$$

$$\Longrightarrow \begin{cases} \lambda = 1 \rightarrow x = 0.5, y = 0.5, z = -.25 \\ \lambda = -1 \rightarrow x = -0.5, y = -0.5, z = .25 \end{cases}$$

$$\Longrightarrow \begin{cases} \lambda = 1 \rightarrow f(0.5, 0.5, -.25) = 2 \text{ بیش} \\ \lambda = -1 \rightarrow f(-0.5, -0.5, .25) = -2 \text{ کم} \end{cases}$$

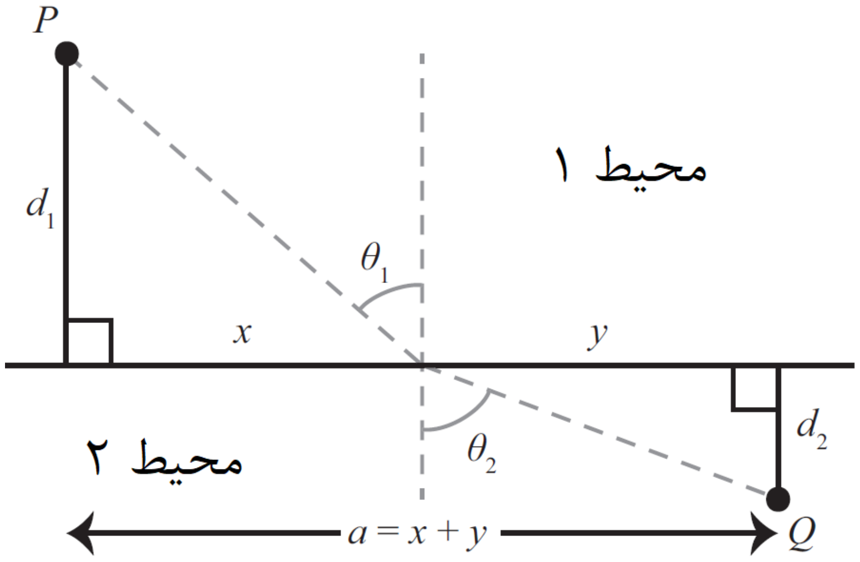
قانون اسنل (شکست نور)

نور هنگام گذر از محیط شفافی به محیط شفاف دیگر در سطح محیط دوم خم می‌شود تا مسیری با کمترین زمان را طی کند

قانون شکست اسنل

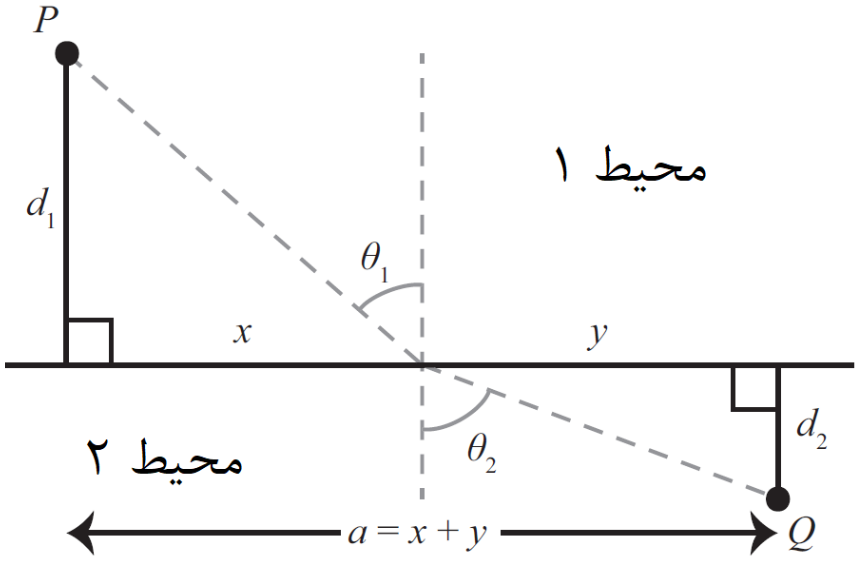
قانون شکست نور

$$\frac{\sin\theta_1}{v_1} = \frac{\sin\theta_2}{v_2}$$



قانون اسنل (شکست نور)

نور هنگام گذر از محیط شفافی به محیط شفاف دیگر در سطح محیط دوم خم می‌شود تا مسیری با کمترین زمان را طی کند



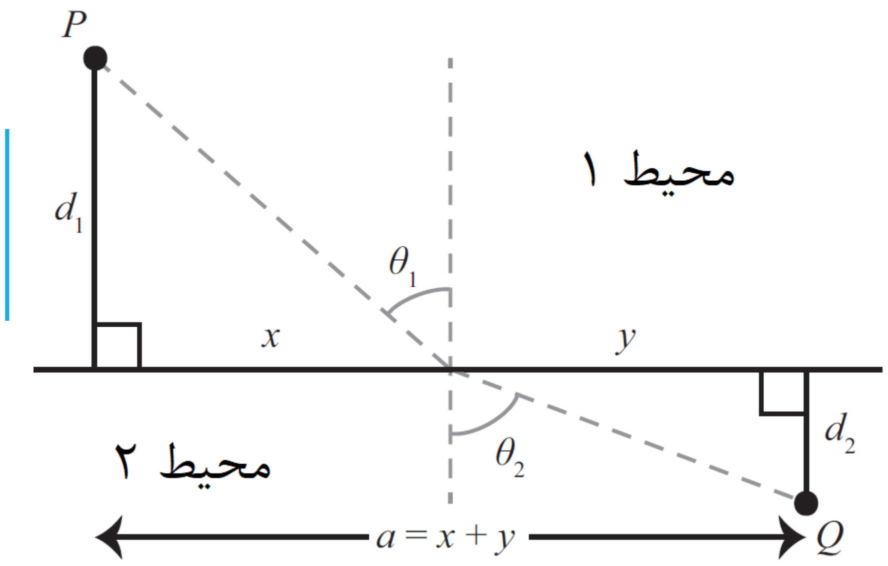
فرض: v_1و v_2 سرعت نور در دو محیط باشد

فاصله از P تا Q: $\sqrt{d_1^2 + x^2} + \sqrt{d_2^2 + y^2}$

سرعت برابر با مسافت تقسیم بر زمان، پس

- $T(x,y) = \frac{\sqrt{d_1^2+x^2}}{v_1} + \frac{\sqrt{d_2^2+y^2}}{v_2}$
- با قید $g(x,y) = x + y = a$

قانون اسنل (شکست نور)



$$T(x,y) = \frac{\sqrt{d_1^2+x^2}}{v_1} + \frac{\sqrt{d_2^2+y^2}}{v_2}$$

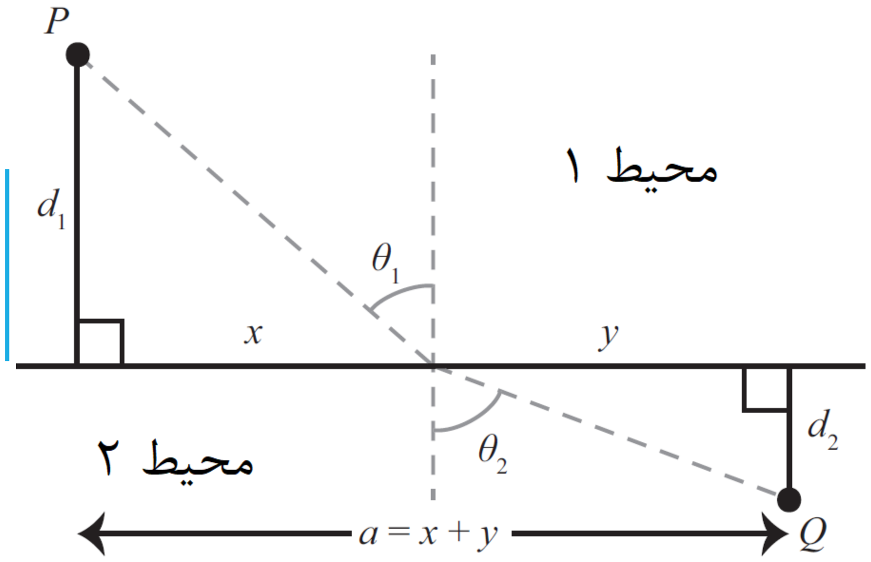
با قید $g(x,y) = x + y = a$

$$\nabla f(x,y) = \lambda \nabla g(x,y)$$

$$\Rightarrow \frac{x}{v_1\sqrt{d_1^2+x^2}}\boldsymbol{i} + \frac{y}{v_2\sqrt{d_2^2+y^2}}\mathbf{j} = \lambda\boldsymbol{i} + \lambda\boldsymbol{j}$$

$$\Rightarrow \begin{cases} \dfrac{x}{v_1\sqrt{d_1^2+x^2}} = \lambda \\ \dfrac{y}{v_2\sqrt{d_2^2+y^2}} = \lambda \\ x + y = a \end{cases} \Rightarrow \frac{x}{v_1\sqrt{d_1^2+x^2}} = \frac{y}{v_2\sqrt{d_2^2+y^2}}$$

قانون اسنل (شکست نور)



$$\frac{x}{v_1\sqrt{d_1^2+x^2}} = \frac{y}{v_2\sqrt{d_2^2+y^2}}$$

$$\begin{cases} \sin\theta_1 = \dfrac{x}{\sqrt{d_1^2+x^2}} \\ \sin\theta_2 = \dfrac{y}{\sqrt{d_2^2+y^2}} \end{cases}$$

$$\Longrightarrow \frac{\sin\theta_1}{v_1} = \frac{\sin\theta_2}{v_2}$$

مثال ۴

$f(x,y) = x^2 + xy + y^2$

$x^2 + y^2 = 8$

$g(x,y) = x^2 + y^2$

$\nabla f(x,y) = \lambda \nabla g(x,y)$

$(2\mathrm{x} + \mathrm{y})\mathbf{i} + (2\mathrm{y} + \mathrm{x})\mathbf{j} = \lambda(2\mathrm{x}\mathbf{i} + 2\mathrm{y}\mathbf{j})$

$$\Longrightarrow \begin{cases} (2\mathrm{x} + \mathrm{y}) = \lambda(2\mathrm{x}) \to y = 2\lambda x - 2x \to y = 2x(\lambda - 1) \ * \\ (2\mathrm{y} + \mathrm{x}) = \lambda(2\mathrm{y}) \to x = 2\lambda y - 2y \to x = 2y(\lambda - 1) \overset{*}{\to} x = 2\big(2x(\lambda - 1)\big)(\lambda - 1) \\ x^2 + y^2 = 8 \end{cases}$$

$$x = 4x(\lambda - 1)^2 \to x(4(\lambda - 1)^2 - 1) = 0 \begin{cases} x = 0 \ \text{نامعتبر} \\ 4(\lambda - 1)^2 - 1 = 0 \to \begin{cases} \lambda = 1.5 \to y = x \\ \lambda = 0.5 \to y = -x \end{cases} \end{cases}$$

$$\lambda = 1.5 \Longrightarrow y = x \Longrightarrow x^2 + y^2 = 8 \Longrightarrow x^2 + x^2 = 8 \Longrightarrow \begin{cases} x = 2, y = 2 \\ x = -2, y = -2 \end{cases}$$

$$\lambda = 0.5 \Longrightarrow y = -x \Longrightarrow x^2 + y^2 = 8 \Longrightarrow x^2 + x^2 = 8 \Longrightarrow \begin{matrix} x = 2, y = -2 \\ x = -2, y = 2 \end{matrix}$$

$$\begin{cases} \left.\begin{matrix} f(2,2) = 12 \\ f(2,-2) = 12 \end{matrix}\right\} \text{بیش} \\ \left.\begin{matrix} f(-2,2) = 4 \\ f(-2,-2) = 4 \end{matrix}\right\} \text{کم} \end{cases}$$

چند قید تساوی

در نظر گرفتن تابعی

$$f(\boldsymbol{x}) = z$$

با محدودیت‌های

$$g_۱(\boldsymbol{x}) = c_۱$$
$$g_۲(\boldsymbol{x}) = c_۲$$

$c_۱$ و $c_۲$ ثابت

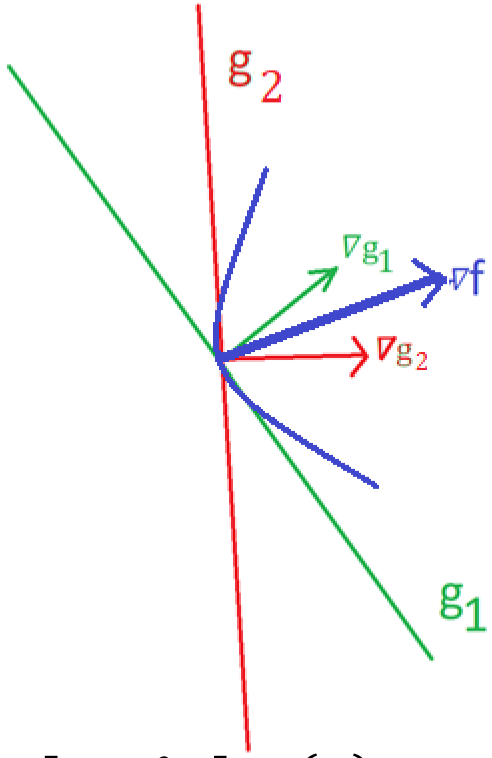
g سطح ترازی از سطح $g(\boldsymbol{x}) = z$

چند قید تساوی

$f(\boldsymbol{x}) = z$

$g_{۱}(\boldsymbol{x}) = c_{۱}$

$g_{۲}(\boldsymbol{x}) = c_{۲}$



$$\nabla f(x) = \lambda_1 \nabla g_{۱}(\boldsymbol{x}) + \lambda_2 \nabla g_2(\boldsymbol{x})$$

$$\nabla f(x) = \lambda . \nabla g(x)$$

$$\left.\begin{matrix} g(x) = c \\ \lambda > 0 \end{matrix}\right\} \Longrightarrow (g(x) - c)\lambda = 0$$

$$\mathcal{L}(\boldsymbol{x}, \boldsymbol{\lambda}) = f(\boldsymbol{x}) - \lambda_1[g_{۱}(\boldsymbol{x}) - c_{۱}] - \lambda_{۲}[g_{۲}(\boldsymbol{x}) - c_{۲}]$$

چند قید تساوی

$$\mathcal{L}(\boldsymbol{x},\boldsymbol{\lambda}) = f(\boldsymbol{x}) - \lambda_1[g_۱(\boldsymbol{x}) - c_۱] - \lambda_۲[g_۲(\boldsymbol{x}) - c_۲]$$

$$\nabla f(\boldsymbol{x}) = \lambda_۱ \nabla g_۱(\boldsymbol{x}) + \lambda_۲ \nabla g_۲(\boldsymbol{x}) = [\lambda_۱, \lambda_۲] \begin{bmatrix} \frac{\partial g_۱}{\partial x_۱} & \frac{\partial g_۱}{\partial x_۲} \\ \frac{\partial g_۲}{\partial x_۱} & \frac{\partial g_۲}{\partial x_۲} \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} \frac{\partial f}{\partial x_۱} \\ \frac{\partial f}{\partial x_۲} \end{bmatrix} = [\lambda_۱, \lambda_۲] \begin{bmatrix} \frac{\partial g_۱}{\partial x_۱} & \frac{\partial g_۱}{\partial x_۲} \\ \frac{\partial g_۲}{\partial x_۱} & \frac{\partial g_۲}{\partial x_۲} \end{bmatrix}$$

چند قید تساوی

m قید و n متغیر

$$\Rightarrow \begin{bmatrix} \dfrac{\partial f}{\partial x_{۱}} \\ \vdots \\ \dfrac{\partial f}{\partial x_n} \end{bmatrix} = [\lambda_{۱}, \dots, \lambda_m] \begin{bmatrix} \dfrac{\partial g_{۱}}{\partial x_{۱}} & \cdots & \dfrac{\partial g_{۱}}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial g_m}{\partial x_1} & \cdots & \dfrac{\partial g_m}{\partial x_n} \end{bmatrix}$$

شرط ناتباهیدگی

- رتبه m

چند قید تساوی

$$\Rightarrow \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{bmatrix} = [\lambda_1, \dots, \lambda_m] \begin{bmatrix} \frac{\partial g_1}{\partial x_1} & \cdots & \frac{\partial g_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial g_m}{\partial x_1} & \cdots & \frac{\partial g_m}{\partial x_n} \end{bmatrix}$$

$$\begin{aligned} \mathcal{L}(\boldsymbol{x}, \boldsymbol{\lambda}) &= \mathcal{L}(x_1, \dots, x_n, \lambda_1, \dots, \lambda_m) \\ &= f(x_1, \dots, x_n) - \textstyle\sum_{i=1}^{m} \lambda_i (g_i(x_1, \dots, x_n) - c_i) \end{aligned}$$

شروط لازم جهت بیشینه بودن:

$$\frac{\partial \mathcal{L}}{\partial x_1} = 0, \dots, \frac{\partial \mathcal{L}}{\partial x_n} = 0$$

$$\frac{\partial \mathcal{L}}{\partial \lambda_1} = 0, \dots, \frac{\partial \mathcal{L}}{\partial \lambda_m} = 0$$

مثال

$f(x, y, z) = xyz$

$g_{١}(x, y, z) \equiv x^2 + y^2 = 1$

$g_2(x, y, z) \equiv x + z = 1$

حل: ماتریس ژاکوبی

$$\begin{bmatrix} 2x & 2y & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

رتبه ۱ اگر و فقط اگر x=y=0

- ممکن نیست
  - نادیده‌گیری شرط اول
- پس رتبه ۲ و ناتباهیده

$$\mathcal{L}(x, y, z, \lambda_1, \lambda_2) = xyz - \lambda_1(x^2 + y^2 - 1) - \lambda_2(x + z - 1)$$

مثال- ادامه

$$\mathcal{L}(x,y,z,\lambda_1,\lambda_2) = xyz - \lambda_1(x^2+y^2-1) - \lambda_2(x+z-1)$$

$$\frac{\partial \mathcal{L}}{\partial x} = yz - 2\lambda_1 x - \lambda_2 = 0$$

$$\frac{\partial \mathcal{L}}{\partial y} = xz - 2\lambda_1 y = 0$$

$$\frac{\partial \mathcal{L}}{\partial z} = xy - \lambda_2 = 0$$

$$\frac{\partial \mathcal{L}}{\partial \lambda_۱} = x^2 + y^2 - 1 = 0$$

$$\frac{\partial \mathcal{L}}{\partial \lambda_m} = x + z - 1 = 0$$

$$\lambda_۱ = \frac{xz}{2y}, \lambda_2 = \mathrm{xy}$$

$$yz - 2\frac{xz}{2y}x - \mathrm{xy} = 0$$

$$(1-x^2)(1-x) - x^2(1-x) - x(1-x^2) = 0$$

$$(1-x)[1-3x^2-x] = 0 \Rightarrow \begin{cases} (1-x) = 0 \\ 1-3x^2-x = 0 \end{cases}$$

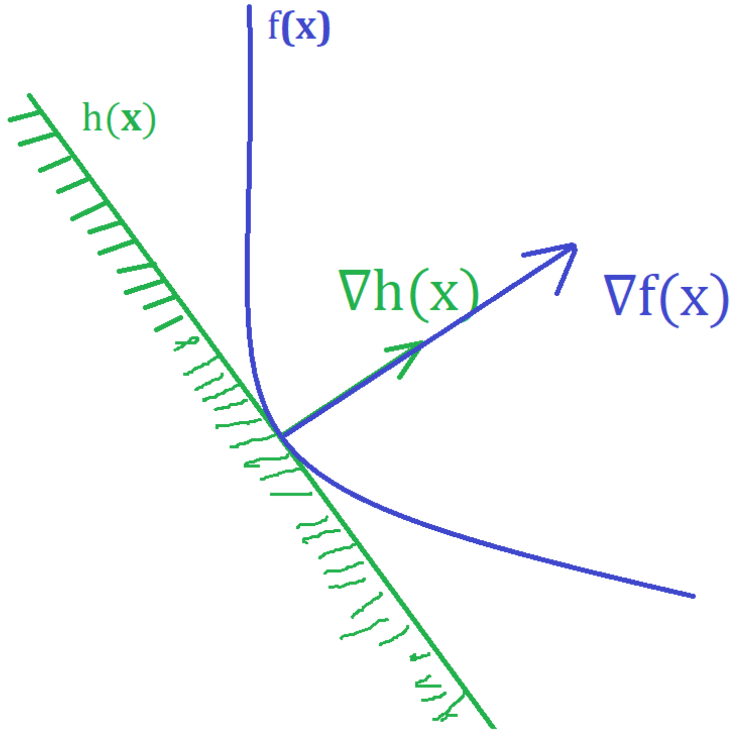
$$\begin{cases} x = 1 \\ x = \dfrac{-1+\sqrt{13}}{6} \\ \dfrac{-1-\sqrt{13}}{6} \end{cases}$$

تک قید نامساوی

$$\max_{\boldsymbol{x}\in\mathbb{R}^n} f(\boldsymbol{x})$$

با توجه

$$h(\boldsymbol{x}) \leq b$$



$$\nabla f(x) = \lambda \nabla h(x)$$

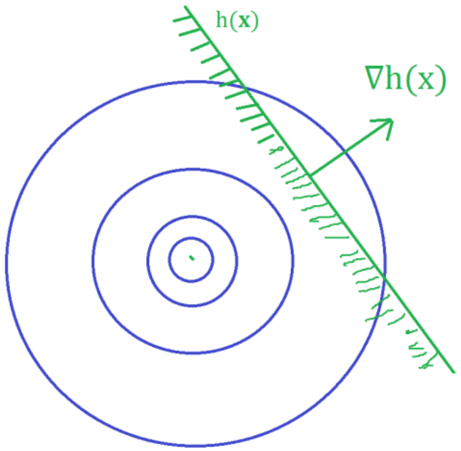
$$\left.\begin{array}{r} h(x) = b \\ \lambda > 0 \end{array}\right\} \Longrightarrow (h(x) - b)\lambda = 0$$

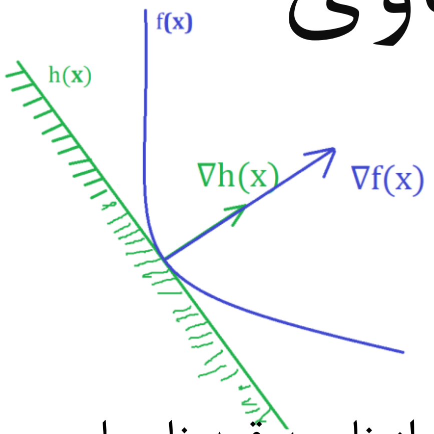
تک قید نامساوی

$$\max_{\boldsymbol{x}\in\mathbb{R}^n} f(\boldsymbol{x})$$

با توجه

$$h(\boldsymbol{x}) \leq b$$





«بیشینهٔ بدون‌محدودیت» خارج از ناحیه قید نامساوی

- قید مانع
- $h(x) - b = 0$ و $\lambda \geq 0$

«بیشینه بدون‌محدودیت» داخل ناحیه قید نامساوی

- قید غیرمانع (فارغ یا کم)
- $h(x) - b \leq 0$ و $\lambda = 0$

$\Longleftarrow$ شروط کمبود متمم

- $[h(x) - b]\lambda = 0$

$\Longleftarrow$ فارغ از مانع بودن یا نبودن

- $\lambda \geq 0$
- $h(x) \leq b$

شروط لازم مرتبه اول تک قید نامساوی

$$\max_{\boldsymbol{x}\in\mathbb{R}^n} f(\boldsymbol{x})$$

با توجه

$$h(\boldsymbol{x}) \le b$$

$$\mathcal{L}(\boldsymbol{x},\lambda) = f(\boldsymbol{x}) - \lambda[g(\boldsymbol{x}) - b]$$

$$\frac{\partial\mathcal{L}(\boldsymbol{x},\lambda)}{\partial x_i} = \frac{\partial f}{\partial x_i} - \lambda\frac{\partial g(\boldsymbol{x})}{\partial x_i} = 0, \forall i = 1,\dots,n$$

$\lambda\frac{\partial\mathcal{L}(\boldsymbol{x},\lambda)}{\partial \lambda} = \lambda[g(\boldsymbol{x}) - b] \ = 0$ کمبود مکمل

$\frac{\partial\mathcal{L}(\boldsymbol{x},\lambda)}{\partial \lambda} = [g(\boldsymbol{x}) - b] \le 0$ قید اصلی

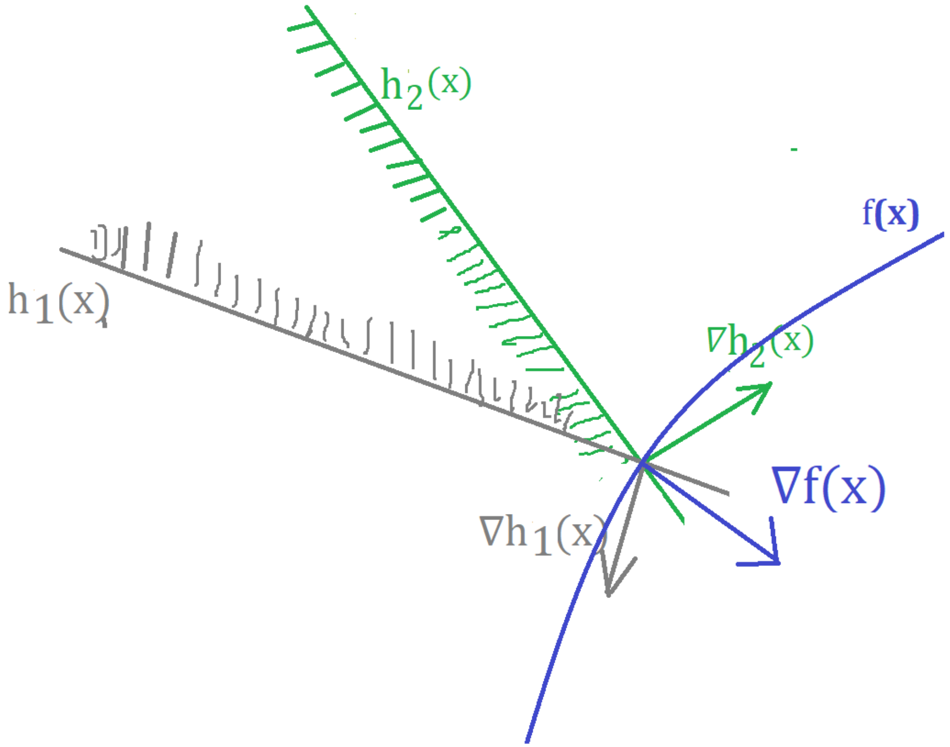
$$\lambda \ge 0$$

چند قید نامساوی

$$\max_{\boldsymbol{x}\in\mathbb{R}^n} f(\boldsymbol{x})$$

با توجه

$$h_1(\boldsymbol{x}) \le b_1$$
$$h_2(\boldsymbol{x}) \le b_2$$



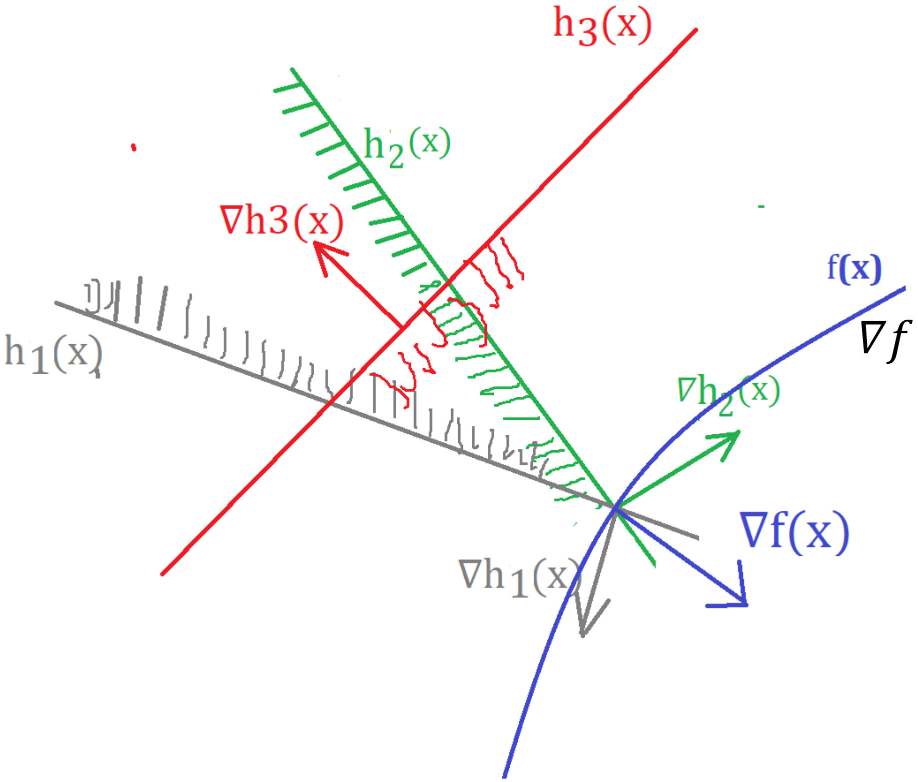
$$\nabla f(x) = \lambda_1 \nabla h_{\text{۱}}(\boldsymbol{x}) + \lambda_2 \nabla h_2(\boldsymbol{x})$$
$$\nabla f(x) = \lambda . \nabla h(x)$$
$$\left.\begin{matrix} h_1(x) = b_1 \\ \lambda_1 > 0 \end{matrix}\right\} \Longrightarrow (h_1(x) - b_1)\lambda_1 = 0$$
$$\left.\begin{matrix} h_2(x) = b_2 \\ \lambda_2 > 0 \end{matrix}\right\} \Longrightarrow (h_2(x) - b_2)\lambda_2 = 0$$

چند قید نامساوی

$$\max_{\boldsymbol{x}\in\mathbb{R}^n} f(\boldsymbol{x})$$

با توجه

$$h_1(\boldsymbol{x}) \leq b_1$$
$$h_2(\boldsymbol{x}) \leq b_2$$
$$h_3(\boldsymbol{x}) \leq b_3$$



$$\nabla f(x) = \lambda_1 \nabla h_1(\boldsymbol{x}) + \lambda_2 \nabla h_2(\boldsymbol{x}) + \lambda_3 \nabla h_3(\boldsymbol{x})$$
$$\nabla f(x) = \lambda . \nabla h(x)$$
$$\left.\begin{matrix} h_1(x) = b_1 \\ \lambda_1 > 0 \end{matrix}\right\} \Rightarrow (h_1(x) - b_1)\lambda_1 = 0$$
$$\left.\begin{matrix} h_2(x) = b_2 \\ \lambda_2 > 0 \end{matrix}\right\} \Rightarrow (h_2(x) - b_2)\lambda_2 = 0$$
$$\left.\begin{matrix} h_3(x) \leq b_3 \\ \lambda_3 = 0 \end{matrix}\right\} \Rightarrow (h_3(x) - b_3)\lambda_3 = 0$$

شروط لازم مرتبه اول چند قید نامساوی

$$\max_{\boldsymbol{x}\in\mathbb{R}^n} f(\boldsymbol{x})$$

با توجه

$$h_1(\boldsymbol{x}) \le b_1$$
$$\vdots$$
$$h_k(\boldsymbol{x}) \le b_k$$

$$\mathcal{L}(\boldsymbol{x},\lambda) = f(\boldsymbol{x}) - \lambda[h(\boldsymbol{x}) - b]$$

$$\frac{\partial\mathcal{L}(\boldsymbol{x},\lambda)}{\partial x_1} = 0, \dots, \frac{\partial\mathcal{L}(\boldsymbol{x},\lambda)}{\partial x_n} = 0$$

$$\lambda_1\,[h_1(\boldsymbol{x}) - b_1] = 0, \dots, \lambda_k\,[h_k(\boldsymbol{x}) - b_k] = 0$$

$$h_1(\boldsymbol{x}) \le b_1, \dots, h_k(\boldsymbol{x}) \le b_k$$

$$\lambda_1 \ge 0, \dots, \lambda_k \ge 0$$

ترکیب انواع قیدها

$\max f(x_{۱}, \ldots, x_n)$

$g_1(x_{۱}, \ldots, x_n) = c_1, \ldots, g_m(x_{۱}, \ldots, x_n) = c_m$

$h_1(x_{۱}, \ldots, x_n) \leq b_1, \ldots, h_k(x_{۱}, \ldots, x_n) \leq b_k$

شروط لازم مرتبه اول ترکیب انواع قیدها

x^* کمینه‌ساز محلی روی مجموعه مقید دارای m تساوی و k نامساوی

فرض تعداد k_0 ناتساوی از نوع قید مانع (binding) و $k-k_0$ ناتساوی از شمار قید غیرمانع(non binding)

فرض ماتریس ژاکوبی رتبه کامل

$$\mathcal{L}(\boldsymbol{x},\lambda_1,\dots,\lambda_k,\mu_1,\dots,\mu_m)=f(\boldsymbol{x})-\sum_{i=1}^{k}\lambda_i[h_i(\boldsymbol{x})-b_i]-\sum_{i=1}^{m}\mu_i[g_i(\boldsymbol{x})-c_i]$$

I) $\frac{\partial\mathcal{L}(\boldsymbol{x}^*,\lambda^*,\mu^*)}{\partial x_i}=0,\quad \forall i\in\{1,\dots,n\}$

II) $\lambda_i^*[h_i(\boldsymbol{x}^*)-b_1]=0,\forall i\in\{1,\dots,k\}$

III) $g_i(\boldsymbol{x}^*)=c_i,\forall i\in\{1,\dots,m\}$

IV) $h_i(\boldsymbol{x}^*)\geq b_i,\forall i\in\{1,\dots,k\}$

V) $\lambda_i\geq 0,\forall i\in\{1,\dots,k\}$

(ک‌ک‌ت)

شروط لازم مرتبه اول ترکیب انواع قیدها

$\boldsymbol{x}^*$ بیشینه‌ساز محلی روی مجموعه مقید دارای m تساوی و k نامساوی

فرض تعداد k_0 ناتساوی از نوع قید مانع (binding) و $k - k_0$ ناتساوی از شمار قید غیرمانع(non binding)

فرض ماتریس ژاکوبی رتبه کامل

$$\mathcal{L}(\boldsymbol{x}, \lambda_1, \dots, \lambda_k, \mu_1, \dots, \mu_m) = f(\boldsymbol{x}) - \sum_{i=1}^{k} \lambda_i [h_i(\boldsymbol{x}) - b_i] - \sum_{i=1}^{m} \mu_i [g_i(\boldsymbol{x}) - c_i]$$

I) $\frac{\partial \mathcal{L}(\boldsymbol{x}^*, \lambda^*, \mu^*)}{\partial x_i} = 0, \quad \forall i \in \{1, \dots, n\}$

II) $\lambda_i^* [h_i(\boldsymbol{x}^*) - b_1] = 0, \forall i \in \{1, \dots, k\}$

III) $g_i(\boldsymbol{x}^*) = c_i, \forall i \in \{1, \dots, m\}$

IV) $h_i(\boldsymbol{x}^*) \leq b_i, \forall i \in \{1, \dots, k\}$

V) $\lambda_i \geq 0, \forall i \in \{1, \dots, k\}$

KKT (ک‌ک‌ت)

شروط لازم مرتبه اول ترکیب انواع قیدها

x^* کمینه‌ساز محلی روی مجموعه مقید دارای m تساوی و k نامساوی

فرض تعداد k_0 ناتساوی از نوع قید مانع (binding) و $k-k_0$ ناتساوی از شمار قید غیرمانع(non binding)

فرض ماتریس ژاکوبی رتبه کامل

$$\mathcal{L}(\boldsymbol{x},\lambda_1,\dots,\lambda_k,\mu_1,\dots,\mu_m)=f(\boldsymbol{x})-\sum_{i=1}^{k}\lambda_i[h_i(\boldsymbol{x})-b_i]-\sum_{i=1}^{m}\mu_i[g_i(\boldsymbol{x})-c_i]$$

I) $\frac{\partial\mathcal{L}(\boldsymbol{x}^*,\lambda^*,\mu^*)}{\partial x_i}=0,\ \ \forall i\in\{1,\dots,n\}$

II) $\lambda_i^*[h_i(\boldsymbol{x}^*)-b_1]=0,\forall i\in\{1,\dots,k\}$

III) $g_i(\boldsymbol{x}^*)=c_i,\forall i\in\{1,\dots,m\}$

IV) $h_i(\boldsymbol{x}^*)\geq b_i,\forall i\in\{1,\dots,k\}$

V) $\lambda_i\geq 0,\forall i\in\{1,\dots,k\}$

نیاز به توجه به شرط ناتباهیدگی

(ک‌ک‌ت)

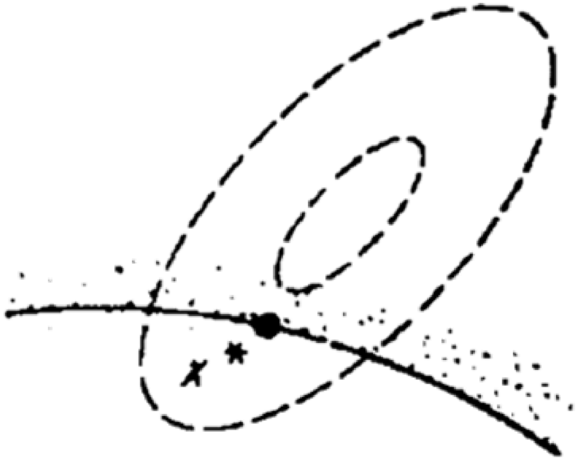
نسبت قید و بهینه

غیرفعال

$\lambda_i^* = 0, \; c_i^* > 0$

فعال ضعیف

$\lambda_i^* = c_i^* = 0$



فعال قوی

$\lambda_i^* > 0, \quad c_i^* = 0$

$f(x,y) = 3x + 4y$

$g_1(x,y) = x^2 + y^2 \leq 4$

$g_2(x,y) = -x \leq -1$

$\mathcal{L}(x,y,\lambda_1,\lambda_2) = 3x + 4y - \lambda_1(x^2 + y^2 - 4\ ) - \lambda_2(-x + 1\ ), \lambda_1, \lambda_2 > 0$

$I)\ \frac{\partial \mathcal{L}}{\partial x} = 3 - 2\lambda_1 x + \lambda_2 = 0$

$II)\ \frac{\partial \mathcal{L}}{\partial y} = 4 - 2\lambda_1 \mathrm{y} = 0$

$III)\ \lambda_1(x^2 + y^2 - 4) = 0$

$IV)\ \lambda_2(-x + 1) = 0$

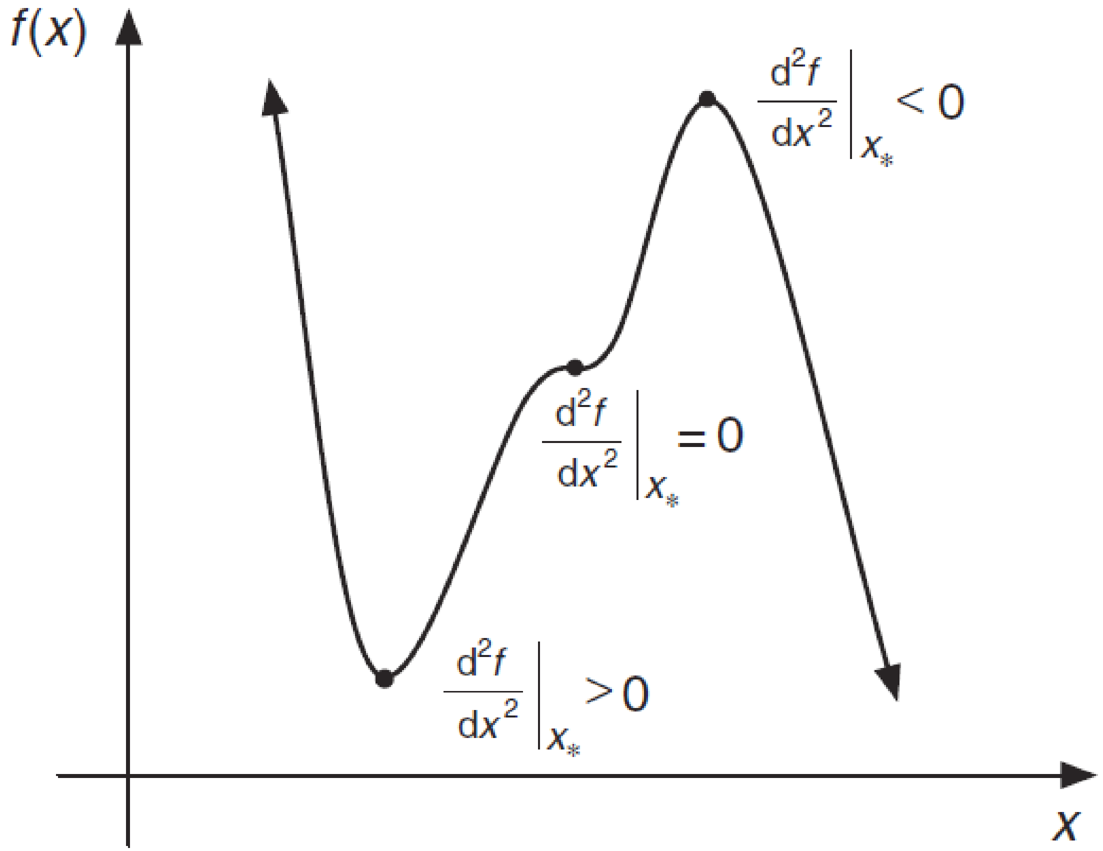
چند ویژگی

ضریب لاگرانژ فراهم کننده اطلاعات حساسیت حول نقطه بهینه

مشکلات

- امکان شکست زمان $\nabla g(x_0, y_0) = 0$
- امکان نیاز به جبر خطی پیچیده و مشکل جهت حل مسئله

شروط بهینگی مرتبه دوم



شرط لازم مرتبه دوم بیشینه

$$\boldsymbol{d}.\nabla_{xx}^2\mathcal{L}(\boldsymbol{x}^*,\lambda^*).\boldsymbol{d} \leq 0, \boldsymbol{d} \in \Lambda(\boldsymbol{x}^*)$$

$$\Lambda(\boldsymbol{x}^*) = \{\boldsymbol{d} \in R^n, \left(\nabla h_i(\boldsymbol{x}^*)\right)^T.\boldsymbol{d} = 0, i = 1, \dots, m, i \in \mathcal{J}(\boldsymbol{x}^*)\}$$

شروط بهینگی مرتبه دوم

$$\begin{aligned} f(\boldsymbol{x}^* + \delta) &= \mathcal{L}(\boldsymbol{x}^* + \delta, \lambda^*) \\ &= \mathcal{L}(\boldsymbol{x}^*, \lambda^*) + \delta^T \nabla_{\boldsymbol{x}} \mathcal{L}(\boldsymbol{x}^*, \lambda^*) + \frac{1}{2} \delta^T \nabla_{xx}^2 \mathcal{L}(\boldsymbol{x}^*, \lambda^*) \delta + O(\delta^T \delta) \\ &= f^* + \frac{1}{2} \delta^T \nabla_{xx}^2 \mathcal{L}(\boldsymbol{x}^*, \lambda^*) \delta + O(\delta^T \delta) \end{aligned}$$

$$\nabla_{xx}^2 \mathcal{L}(\boldsymbol{x}^*, \lambda^*) = \nabla_{xx}^2 f(\boldsymbol{x}^*) - \sum_{i=1}^{m} \lambda_i^* . \nabla^2 h_i(\boldsymbol{x}^*)$$

منفی معین بودن ماتریس هسی تابع لاگرانژ

نیاز به تحلیل بیشتر

شروط بهینگی مرتبه دوم

$$
\begin{aligned}
f(\boldsymbol{x}^* + \delta) &= \mathcal{L}(\boldsymbol{x}^* + \delta, \lambda^*) \\
&= \mathcal{L}(\boldsymbol{x}^*, \lambda^*) + \delta^T \nabla_{\boldsymbol{x}} \mathcal{L}(\boldsymbol{x}^*, \lambda^*) + \frac{1}{2}\delta^T \nabla_{xx}^2 \mathcal{L}(\boldsymbol{x}^*, \lambda^*)\delta + O(\delta^T \delta) \\
&= f^* + \frac{1}{2}\delta^T \nabla_{xx}^2 \mathcal{L}(\boldsymbol{x}^*, \lambda^*)\delta + O(\delta^T \delta)
\end{aligned}
$$

$$\nabla_{xx}^2 \mathcal{L}(\boldsymbol{x}^*, \lambda^*) = \nabla_{xx}^2 f(\boldsymbol{x}^*) - \sum_{i=1}^{m} \lambda_i^* . \nabla^2 h_i(\boldsymbol{x}^*)$$

$$\boldsymbol{d}. \nabla_{xx}^2 \mathcal{L}(\boldsymbol{x}^*, \lambda^*). \boldsymbol{d} < 0, \forall \boldsymbol{d}$$

$$\left(\nabla h_i(\boldsymbol{x}^*)\right)^T . \boldsymbol{d} = 0, i = 1, \dots, m$$

شرط بهینگی مرتبه دوم

$$\nabla_{xx}^2 \mathcal{L}(\boldsymbol{x}^*, \lambda^*) = \nabla^2 f(x^*) - \sum_{i=1}^{m} \lambda_i^* . \nabla^2 h_i(\boldsymbol{x}^*)$$

$$\boldsymbol{d} . \nabla_{xx}^2 \mathcal{L}(\boldsymbol{x}^*, \lambda^*) . \boldsymbol{d} < 0, \forall \boldsymbol{d} : \left(\nabla h_i(\boldsymbol{x}^*)\right)^T . \boldsymbol{d} = 0, i = 1, \dots, m$$

مثال

$$\max x_1x_2 + x_2x_3 + x_1x_3$$

$$x_1 + x_2 + x_3 = 3$$

حل

$$\mathcal{L}(\boldsymbol{x}, \lambda) = x_1x_2 + x_2x_3 + x_1x_3 - \lambda(x_1 + x_2 + x_3 - 3)$$

$$\nabla\mathcal{L}(\boldsymbol{x}, \lambda) = \begin{bmatrix} \frac{\partial\mathcal{L}(\boldsymbol{x},\lambda)}{\partial x_1} \\ \frac{\partial\mathcal{L}(\boldsymbol{x},\lambda)}{\partial x_2} \\ \frac{\partial\mathcal{L}(\boldsymbol{x},\lambda)}{\partial x_3} \\ \frac{\partial\mathcal{L}(\boldsymbol{x},\lambda)}{\partial \lambda} \end{bmatrix} = \begin{bmatrix} x_2 + x_3 - \lambda \\ x_1 + x_3 - \lambda \\ x_1 + x_2 - \lambda \\ x_1 + x_2 + x_3 - 3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \Longrightarrow \begin{bmatrix} x_1^* \\ x_2^* \\ x_3^* \\ \lambda^* \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 2 \end{bmatrix}$$

مثال- ادامه

$$\begin{bmatrix} x_1^* \\ x_2^* \\ x_3^* \\ \lambda^* \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 2 \end{bmatrix} \qquad \nabla\mathcal{L}(\boldsymbol{x},\lambda\,) = \begin{bmatrix} x_2 + x_3 - \lambda \\ x_1 + x_3 - \lambda \\ x_1 + x_2 - \lambda \\ x_1 + x_2 + x_3 - 3 \end{bmatrix}$$

$$\nabla_{xx}^2\mathcal{L}(\boldsymbol{x},\lambda\,) = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \Longrightarrow 2, -1, -1$$

خود مثبت معین یا منفی معین نیست.

$$\boldsymbol{d}\colon d_1 + d_2 + d_3 = 0$$

$$\boldsymbol{d}^T.\nabla_{xx}^2\mathcal{L}(\boldsymbol{x}^*,\lambda^*).\boldsymbol{d} = -(d_1^2 + d_2^2 + d_3^2)<0$$

مثال -۲- ادامه

بیشینه xyz با قید $(xy + yz + xz) = \frac{c}{2}$

$$\mathcal{L}(\boldsymbol{x}, \lambda) = xyz - \lambda\left(xy + yz + xz - \frac{c}{2}\right)$$

$$\nabla\mathcal{L}(\boldsymbol{x}, \lambda) = \begin{bmatrix} \frac{\partial\mathcal{L}(\boldsymbol{x},\lambda)}{\partial x} \\ \frac{\partial\mathcal{L}(\boldsymbol{x},\lambda)}{\partial y} \\ \frac{\partial\mathcal{L}(\boldsymbol{x},\lambda)}{\partial z} \\ \frac{\partial\mathcal{L}(\boldsymbol{x},\lambda)}{\partial \lambda} \end{bmatrix} = \begin{bmatrix} yz - \lambda y - \lambda z \\ xz - \lambda x - \lambda z \\ xy - \lambda x - \lambda y \\ xy + yz + xz - \frac{c}{2} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

جمع سه معادله اول $\Longrightarrow xy + yz + xz - 2\lambda(x + y + z) = 0 \Longrightarrow \frac{c}{2} - 2\lambda(x + y + z) = 0 \Longrightarrow \lambda \neq 0$

اگر $x=0$ آن‌گاه x و y و z هر سه برابر صفر که ممکن نیست

هر سه متغیر برابر $\sqrt{\frac{c}{6}}$

هسی حاشیه‌دار [مرزی یا کرانی؟]

$$\begin{bmatrix} 0_{m\times m} & \nabla h(x^*)_{m\times n} \\ \nabla h(x^*)^T_{n\times m} & H(\mathcal{L}(x^*))_{n\times n} \end{bmatrix}_{(m+n)\times(m+n)}$$

هر لاگرانژی محدود با تساوی و نامساوی موثر (فعال)

کهاد n-m

- تعداد محدودیت‌ها زوج
  - تابع دارای کمینه اگر همه کهادها مثبت $H_{2m+1} > 0, H_{2m+2} > 0, \ldots, H_{n+m} > 0$
  - تابع دارای بیشینه، یک درمیان منفی و سپس مثبت $H_{2m+1} < 0, H_{2m+2} > 0, \ldots$
- تعداد محدودیت‌ها فرد
  - تابع دارای کمینه، اگر همه کهادها منفی $H_{2m+1} < 0, H_{2m+2} < 0, \ldots, H_{n+m} < 0$
  - تابع دارای بیشینه، یک درمیان مثبت و سپس منفی $H_{2m+1} > 0, H_{2m+2} < 0, \ldots$

کوچکترین باید هم ضریب $(-1)^{m+1}$ و سپس تغییر متناوب علامت

- بیشینه

مثال

$$\max x_1x_2 + x_2x_3 + x_1x_3$$

$$x_1 + x_2 + x_3 = 3$$

حل

- n-m

$$\begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix}$$

$$\begin{bmatrix} \frac{\partial \mathcal{L}(\boldsymbol{x},\lambda)}{\partial x_1} \\ \frac{\partial \mathcal{L}(\boldsymbol{x},\lambda)}{\partial x_2} \\ \frac{\partial \mathcal{L}(\boldsymbol{x},\lambda)}{\partial x_3} \\ \frac{\partial \mathcal{L}(\boldsymbol{x},\lambda)}{\partial \lambda} \end{bmatrix} = \begin{bmatrix} x_2 + x_3 - \lambda \\ x_1 + x_3 - \lambda \\ x_1 + x_2 - \lambda \\ x_1 + x_2 + x_3 - 3 \end{bmatrix} \Rightarrow \begin{bmatrix} x_1^* \\ x_2^* \\ x_3^* \\ \lambda^* \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 2 \end{bmatrix}$$

$$\begin{vmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{vmatrix} = 2$$

$$\begin{vmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{vmatrix} = -3$$

مثال ۲ - ادامه

$$\max f(w;\ x;\ y;\ z) = -w^2 - x^2 - y^2 - z^2$$
$$g(w;\ x;\ y;\ z) = 4w - 3y + z + 15 = 0$$
$$h(w;\ x;\ y;\ z) = -2x - y + z + 5 = 0$$

$$\mathcal{L}(w, x, y, z, \lambda, \mu) = f - \lambda g - \mu h$$

$$\nabla \mathcal{L} = \begin{bmatrix} \mathcal{L}_\lambda \\ \mathcal{L}_\mu \\ \mathcal{L}_w \\ \mathcal{L}_x \\ \mathcal{L}_y \\ \mathcal{L}_z \end{bmatrix} = \begin{bmatrix} -g \\ -h \\ f_w - \lambda g_w - \mu h_w \\ f_x - \lambda g_x - \mu h_x \\ f_y - \lambda g_y - \mu h_y \\ f_z - \lambda g_z - \mu h_z \end{bmatrix} = \begin{bmatrix} -4w + 3y - z - 15 \\ 2x + y - z - 5 \\ -2w - 4\lambda \\ -2x + 2\mu \\ -2y + 3\lambda + \mu \\ -2z - \lambda - \mu \end{bmatrix} = \vec{0}$$

$$\lambda = -1$$
$$\mu = -1$$
$$w = 2$$
$$x = -1$$
$$y = -2$$
$$z = 1$$

مثال ۲ - ادامه

$$-1 = (-1)^3 = (-1)^{m+1}$$

$$H = \begin{bmatrix} \mathcal{L}_{\lambda\lambda} & \mathcal{L}_{\lambda\mu} & \mathcal{L}_{\lambda w} & \mathcal{L}_{\lambda x} & \mathcal{L}_{\lambda y} & \mathcal{L}_{\lambda z} \\ \mathcal{L}_{\mu\lambda} & \mathcal{L}_{\mu\mu} & \mathcal{L}_{\mu w} & \mathcal{L}_{\mu x} & \mathcal{L}_{\mu y} & \mathcal{L}_{\mu z} \\ \mathcal{L}_{w\lambda} & \mathcal{L}_{w\mu} & \mathcal{L}_{ww} & \mathcal{L}_{wx} & \mathcal{L}_{wy} & \mathcal{L}_{wz} \\ \mathcal{L}_{x\lambda} & \mathcal{L}_{x\mu} & \mathcal{L}_{xw} & \mathcal{L}_{xx} & \mathcal{L}_{xy} & \mathcal{L}_{xz} \\ \mathcal{L}_{y\lambda} & \mathcal{L}_{y\mu} & \mathcal{L}_{yw} & \mathcal{L}_{yx} & \mathcal{L}_{yy} & \mathcal{L}_{yz} \\ \mathcal{L}_{z\lambda} & \mathcal{L}_{z\mu} & \mathcal{L}_{zw} & \mathcal{L}_{zx} & \mathcal{L}_{zy} & \mathcal{L}_{zz} \end{bmatrix} = \begin{bmatrix} 0 & 0 & -4 & 0 & 3 & -1 \\ 0 & 0 & 0 & 2 & 1 & -1 \\ -4 & 0 & -2 & 0 & 0 & 0 \\ 0 & 2 & 0 & -2 & 0 & 0 \\ 3 & 1 & 0 & 0 & -2 & 0 \\ -1 & -1 & 0 & 0 & 0 & -2 \end{bmatrix}$$

$$H_5 = \begin{vmatrix} 0 & 0 & -4 & 0 & 3 \\ 0 & 0 & 0 & 2 & 1 \\ -4 & 0 & -2 & 0 & 0 \\ 0 & 2 & 0 & -2 & 0 \\ 3 & 1 & 0 & 0 & -2 \end{vmatrix} = -232 < 0$$

$$H_6 = \begin{vmatrix} 0 & 0 & -4 & 0 & 3 & -1 \\ 0 & 0 & 0 & 2 & 1 & -1 \\ -4 & 0 & -2 & 0 & 0 & 0 \\ 0 & 2 & 0 & -2 & 0 & 0 \\ 3 & 1 & 0 & 0 & -2 & 0 \\ -1 & -1 & 0 & 0 & 0 & -2 \end{vmatrix} = 560 > 0$$

دوگانی

کران بالای مقدار بیشینهٔ تابع مقید
- بیشینهٔ تابع هدف ≤ بیشینهٔ تابع مقید

کران پایین مقدار کمینه تابع مقید
- کمینهٔ تابع هدف ≥ کمینهٔ تابع مقید

دوگانی

$\max -x_1^2 - x_2^2 - 2x_1$

باتوجه

$x_1 + x_2 = 0$

$\Longrightarrow x_2 = -x_1 \Longrightarrow -x_1^2 - x_2^2 - 2x_1 = -2x_1^2 - 2x_1 \Longrightarrow x_1 = -\frac{1}{2}$

جواب بهینه$\Longrightarrow \left(-\frac{1}{2}, \frac{1}{2}\right), f^* = \frac{1}{2}$

دوگانی

کران بالای مقدار بیشینهٔ تابع

- بیشینهٔ تابع هدف $\leq$ بیشینهٔ تابع مقید

کران پایین مقدار کمینه تابع

- کمینهٔ تابع هدف $\geq$ کمینهٔ تابع مقید

$$P\begin{cases}\max -x_1^2 - x_2^2 - 2x_1 \\ \text{باتوجه} \\ x_1 + x_2 = 0\end{cases}$$

$$P_0\left\{\max -x_1^2 - x_2^2 - 2x_1\right.$$

$$P_0 \geq P?\ (-1,0),\ \mathrm{f} = 1$$

دوگانی

کران بالای مقدار بیشینهٔ تابع

- بیشینهٔ تابع هدف $\leq$ بیشینهٔ تابع مقید

$$P\begin{cases}\max -x_1^2 - x_2^2 - 2x_1 \\ \text{باتوجه} \\ x_1 + x_2 = 0\end{cases}$$

$$P_\mu \left\{\max_{x_1,x_2} -x_1^2 - x_2^2 - 2x_1 - \mu(x_1 + x_2)\right.$$

$$P_\mu \geq P? \left(-1 - \frac{\mu}{2}, -\frac{\mu}{2}\right)$$

$$\mathrm{q}(\mu) = P_\mu{}^* = -\left(-1 - \frac{\mu}{2}\right)^2 - \left(-\frac{\mu}{2}\right)^2 - 2\left(-1 - \frac{\mu}{2}\right) - \mu(-1 - \mu)$$

$$= \frac{\mu^2}{2} + \mu + 1$$

دوگانی

$$\max_{\boldsymbol{x}\in\mathbb{R}^n} f(\boldsymbol{x}) \qquad\qquad \mathbb{R}^n \to \mathbb{R}$$

با توجه

$$g_i(\boldsymbol{x}) = c_i, \quad i \in \mathcal{E} \qquad \mathbb{R}^n \to \mathbb{R}^m$$

$$h_i(\boldsymbol{x}) \leq b_i, \quad i \in \mathrm{I} \qquad \mathbb{R}^n \to \mathbb{R}^k$$

لاگرانژ

$$\mathcal{L}(\boldsymbol{x},\boldsymbol{\lambda},\boldsymbol{\mu}) = f(\boldsymbol{x}) - \boldsymbol{\lambda}^T(\boldsymbol{h}(\boldsymbol{x}) - \boldsymbol{b}) - \boldsymbol{\mu}^T(\boldsymbol{g}(\boldsymbol{x}) - \boldsymbol{c})$$

دوگان

$$q: \mathbb{R}^{m+k} \to \mathbb{R} :: q(\boldsymbol{\lambda},\boldsymbol{\mu}) = \max_{\boldsymbol{x}\in\mathbb{R}} \mathcal{L}(\boldsymbol{x},\boldsymbol{\lambda},\boldsymbol{\mu})$$

بدون محدودیت

قضیه

فرض $\boldsymbol{x}^*$ جواب بهینهٔ مسئله اصلی (یگان) باشد. اگر $\boldsymbol{\lambda} \in \mathbb{R}^k$ و $\boldsymbol{\mu} \in \mathbb{R}^m$ و $\boldsymbol{\lambda} \geq 0$

آن‌گاه $q(\boldsymbol{\lambda},\boldsymbol{\mu}) \geq \boldsymbol{f}(\boldsymbol{x}^*)$

اثبات

$$q(\boldsymbol{\lambda},\boldsymbol{\mu}) = \max_{\boldsymbol{x}\in\mathbb{R}} \mathcal{L}(\boldsymbol{x},\boldsymbol{\lambda},\boldsymbol{\mu}) = \mathcal{L}(\boldsymbol{x}^*,\boldsymbol{\lambda},\boldsymbol{\mu}) = \boldsymbol{f}(\boldsymbol{x}^*) - \boldsymbol{\lambda}^T[\boldsymbol{h}(\boldsymbol{x}^*) - \boldsymbol{b}] - \boldsymbol{\mu}^T[\boldsymbol{g}(\boldsymbol{x}^*) - \boldsymbol{c}]$$

$$= \boldsymbol{f}(\boldsymbol{x}^*) - \underbrace{\underbrace{\boldsymbol{\lambda}^T[\boldsymbol{h}(\boldsymbol{x}^*) - \boldsymbol{b}]}_{\boldsymbol{\lambda}\geq 0,[\boldsymbol{h}(\boldsymbol{x}^*)-\boldsymbol{b}]\leq 0}}_{\geq 0} - \underbrace{\boldsymbol{\mu}^T[\boldsymbol{g}(\boldsymbol{x}^*) - \boldsymbol{c}]}_{=0} \geq \boldsymbol{f}(\boldsymbol{x}^*)$$

نتیجه

فرض $\boldsymbol{x}^*$ جواب بهینهٔ مسئله اصلی (یگان) باشد. همچنین $\boldsymbol{x}$ جواب شدنی مسئلهٔ اصلی باشد. اگر $\boldsymbol{\lambda} \in \mathbb{R}^k$ و $\boldsymbol{\mu} \in \mathbb{R}^m$ و $\boldsymbol{\lambda} \geq 0$

آن‌گاه $q(\boldsymbol{\lambda},\boldsymbol{\mu}) \geq \boldsymbol{f}(\boldsymbol{x}^*) \geq f(\boldsymbol{x})$

مسئلهٔ دوگان

یافتن بهترین کران بالا

$$\min_{\boldsymbol{\lambda},\boldsymbol{\mu}} q(\boldsymbol{\lambda},\boldsymbol{\mu})$$

با توجه

$$\boldsymbol{\lambda} \geq \mathbf{0}$$

همچنین $(\boldsymbol{\lambda},\boldsymbol{\mu}) \in \{\boldsymbol{\lambda},\boldsymbol{\mu}: \boldsymbol{q}(\boldsymbol{\lambda},\boldsymbol{\mu}) < +\infty\}$

برنامه‌ریزی خطی

$\max \boldsymbol{c}^T\boldsymbol{x}$
با توجه
$A\boldsymbol{x} \leq \boldsymbol{b}$

$$q(\boldsymbol{\lambda}) = \min_{\boldsymbol{\lambda}} \left[\max_{\boldsymbol{x}}\{\boldsymbol{c}^T\boldsymbol{x} - \boldsymbol{\lambda}^T\ (A\boldsymbol{x} - \boldsymbol{b})\}\right] = \min_{\boldsymbol{\lambda}} \left[\max_{\boldsymbol{x}}\{(\boldsymbol{c} - A^T\boldsymbol{\lambda})^T\boldsymbol{x} + \boldsymbol{b}^T\boldsymbol{\lambda}\}\right]$$

اگر $\boldsymbol{c} - A^T\boldsymbol{\lambda} \neq \boldsymbol{0}$ آن‌گاه مقدار اکسترمم برابر $+\infty$

$$q(\boldsymbol{\lambda}) = \begin{cases} +\infty, & \boldsymbol{c} - A^T\boldsymbol{\lambda} \neq \boldsymbol{0} \\ \boldsymbol{b}^T\boldsymbol{\lambda}, & \boldsymbol{c} - A^T\boldsymbol{\lambda} = \boldsymbol{0} \end{cases}$$

$\min \boldsymbol{b}^T\boldsymbol{y}$
با توجه
$A^T\boldsymbol{y} = \boldsymbol{c}$
$\boldsymbol{y} \geq 0$

برنامه‌ریزی خطی

$\max \boldsymbol{c}^T \boldsymbol{x}$
با توجه
$A\boldsymbol{x} = \boldsymbol{b}$
$\boldsymbol{x} \geq 0$

$$q(\boldsymbol{\lambda},\boldsymbol{\mu}) = \min_{\boldsymbol{\lambda}} \left[\max_{\boldsymbol{x}}\{\boldsymbol{c}^T\boldsymbol{x} - \boldsymbol{\mu}^T\ (A^T\boldsymbol{x} - \boldsymbol{b}) - \boldsymbol{\lambda}^T\boldsymbol{x}\}\right] = \min_{\boldsymbol{\lambda}} \left[\max_{\boldsymbol{x}}\{(\boldsymbol{c} - A^T\boldsymbol{\mu} - \boldsymbol{\lambda})^T\boldsymbol{x} + \boldsymbol{b}^T\boldsymbol{\mu}\}\right]$$

اگر $\boldsymbol{c} - A^T\boldsymbol{\mu} - \boldsymbol{\lambda} \neq \boldsymbol{0}$ آن‌گاه مقدار اکسترمم برابر $+\infty$

$$q(\boldsymbol{\lambda},\boldsymbol{\mu}) = \begin{cases} +\infty, & \boldsymbol{c} - A^T\boldsymbol{\mu} - \boldsymbol{\lambda} \neq \boldsymbol{0} \\ \boldsymbol{b}^T\boldsymbol{\mu}, & \boldsymbol{c} - A^T\boldsymbol{\mu} - \boldsymbol{\lambda} = \boldsymbol{0} \end{cases}$$

$\min \boldsymbol{b}^T\boldsymbol{y}$
با توجه
$A^T\boldsymbol{y} \geq \boldsymbol{c}$
$\boldsymbol{y} \geq 0$

قضیهٔ دوگان ضعیف

اگر x^* جواب بهینهٔ مسئله اصلی (یگان) باشد.

اگر $(\boldsymbol{\lambda}^*, \boldsymbol{\mu}^*)$ جواب بهینهٔ مسئلهٔ دوگان باشد.

آن‌گاه $q(\boldsymbol{\lambda}^*, \boldsymbol{\mu}^*) \geq \boldsymbol{f}(x^*)$

نتیجه

اگر یکی از مسئله‌ها (اعم از یگان یا دوگان) بی‌کران باشد، دیگری ناشدنی است.

قضیهٔ دوگان و شدنی (شدن‌پذیری!)

مسئله دوگان				
ناشدنی	**بی‌کران**	**بهینه**		
خیر	خیر	بله	**بهینه**	**مسئله یگان**
بله	خیر	خیر	**بی‌کران**	
بله	بله	خیر	**ناشدنی**	

دوگان ولف

بهینه‌سازی - انواع

نام	f(x)	c(x)
بهینه‌سازی نامقید	غیرخطی	-
برنامه‌ریزی خطی	خطی	خطی
برنامه‌ریزی درجه دو	درجه دو	خطی
بهینه‌سازی مقید خطی	غیرخطی	خطی
بهینه‌سازی مقید یا بهینه‌سازی غیرخطی	غیرخطی	غیرخطی

الگوریتم‌های بهینه‌سازی مقید

برنامه‌ریزی خطی

- روش سیمپلکس
- روش نقطه درونی

برنامه‌ریزی غیرخطی

- روش‌های جریمه‌ای
- روش‌های برنامه‌ریزی دنباله‌ای درجه‌دو
- روش‌های نقطه درونی

تقریبا همگی مبتنی بر جستجو خط و منطقه اعتماد

منابع

B. H. Edwards, “Understanding Multivariable Calculus: Problems, Solutions, and Tips-Course Workbook,” The Great Courses, 2014.

J. Wilde, et al., “Constrained optimization,” 2013.

[فلچر]

B. W. Bader, “Constrained and unconstrained optimization,” 2009