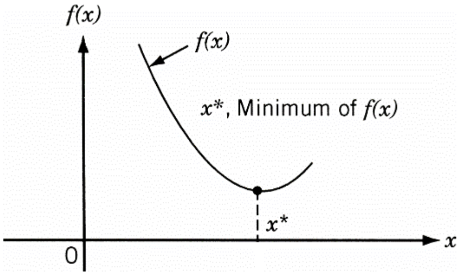
بهینه سازی (Optimization)

بهینه سازی تک پارامتری (نامقید)

$$\min f(x)$$

$$f(x^*) \le f(x) \qquad \forall x \qquad \text{(Global)}$$
$$f(x^*) \le f(x^* + \Delta x) \qquad \forall \Delta x \quad |\Delta x| < \varepsilon \qquad \text{(Local)}$$



بسط تیلور

$$f(x^* + \Delta x) = f(x^*) + \Delta x\, f'(x^*) + \frac{1}{2!}\Delta x^2 f''(x^*) + \frac{1}{3!}\Delta x^3 f'''(x^*) + \dots$$

هدف بهینه سازی (مینیمم سازی)

$$f(x^* + \Delta x) - f(x^*) \ge 0$$

شرط لازم (necessary condition)

$$\Delta x\, f'(x^*) \ge 0$$

اگر Δx می تواند مثبت یا منفی باشد

$$f'(x^*) = 0$$

اگر نمی تواند؟

شرط کافی (sufficient condition) (برای مینیمم کردن)

$$\frac{1}{2!}\Delta x^2 f''(x^*) \ge 0 \qquad \Rightarrow \qquad f''(x^*) \ge 0$$

طبیعتا برای ماکزیمم کردن

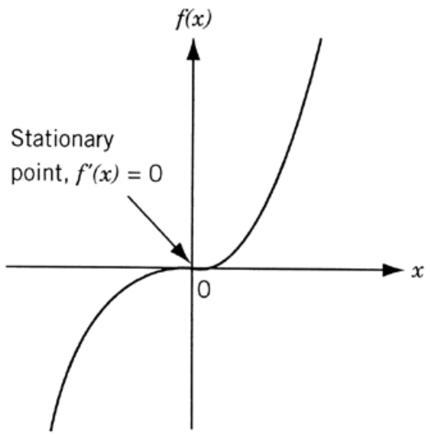
$$\Delta f = f(x^* + \Delta x) - f(x^*) \le 0 \qquad \Rightarrow \qquad f''(x^*) \le 0$$

مثال:

$$
\begin{aligned}
\min f(x) = x^2 \qquad &\Rightarrow \qquad f'(x^*) = 2x^* = 0 \qquad \Rightarrow \qquad x^* = 0 \\
&\Rightarrow \qquad f''(x^*) = 2 > 0
\end{aligned}
$$

مثال:

$$
\begin{aligned}
\min f(x) = x^3 \qquad &\Rightarrow \qquad f'(x^*) = 3x^{*2} = 0 \qquad \Rightarrow \qquad x^* = 0 \\
&\Rightarrow \qquad f''(x^*) = 6x^* = 0 \\
&\Rightarrow \qquad f'''(x^*) = 6 \\
\Delta f = f(x^* + \Delta x) - f(x^*) \quad &= \Delta x\, f'(x^*) + \frac{1}{2!}\Delta x^2 f''(x^*) + \frac{1}{3!}\Delta x^3 f'''(x^*) + \ldots \\
&= \Delta x^3
\end{aligned}
$$



بهینه سازی تک پارامتری مقید (محدوده)

$$
\begin{aligned}
&\min && f(x) \\
&\text{subject to} && a \le x \le b
\end{aligned}
$$

هدف بهینه سازی (مینیمم سازی)

$$
f(x^* + \Delta x) - f(x^*) = \Delta x\, f'(x^*) + \frac{1}{2!}\Delta x^2 f''(x^*) + \frac{1}{3!}\Delta x^3 f'''(x^*) + \ldots \ge 0
$$

شرط لازم (necessary condition)

$$
\Delta x\, f'(x^*) \ge 0
$$

اگر Δx می تواند مثبت یا منفی باشد

$$
f'(x^*) = 0
$$

اگر نمی تواند:

$$\begin{aligned} f'(x^*) > 0 \qquad &\Rightarrow \qquad x^* = a \\ f'(x^*) < 0 \qquad &\Rightarrow \qquad x^* = b \end{aligned}$$

مثال:

$$\min f(x) = x^2 \qquad \Rightarrow \qquad f'(x) = 2x$$

بهینه سازی چند پارامتری (نامقید)

$$\min f(\vec{X})$$

$$\begin{array}{lll} f(\vec{X}^*) \le f(\vec{X}) & \forall \vec{X} & \text{(Global)} \\ f(\vec{X}^*) \le f(\vec{X}^* + \Delta\vec{X}) & \forall \Delta\vec{X} \quad \left\|\Delta\vec{X}\right\| < \varepsilon & \text{(Local)} \end{array}$$

$$\vec{X} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{pmatrix}$$

بسط تیلور

$$\begin{aligned} f(\vec{X}^* + \Delta\vec{X}) &= f(\vec{X}^*) + \Delta\vec{X}^T \left.\frac{\partial f}{\partial \vec{X}}\right|_{\vec{X}^*} + \frac{1}{2!}\Delta\vec{X}^T \left.\frac{\partial^2 f}{\partial \vec{X}^2}\right|_{\vec{X}^*} \Delta\vec{X} + \ldots \\ &= f(\vec{X}^*) + \Delta\vec{X} \cdot \vec{\nabla} f(\vec{X}^*) + \frac{1}{2!}\Delta\vec{X}^T\, \mathrm{H}(\vec{X}^*)\Delta\vec{X} + \ldots \\ &= f(\vec{X}^*) + \sum_{i=1}^{N} \Delta x_i \left.\frac{\partial f}{\partial x_i}\right|_{\vec{X}^*} + \frac{1}{2!}\sum_{i=1}^{N} \Delta x_i \Delta x_j \left.\frac{\partial^2 f}{\partial x_i \partial x_j}\right|_{\vec{X}^*} + \ldots \end{aligned}$$

تعریف نرم:

$$
\begin{aligned}
&\left\|\Delta\vec{X}\right\|_1 = \left|\Delta x_1\right| + \left|\Delta x_2\right| + \cdots + \left|\Delta x_N\right| = \sum_{i=1}^{N}\left|\Delta x_i\right| \\
&\left\|\Delta\vec{X}\right\|_2 = \sqrt{\Delta x_1^2 + \Delta x_2^2 + \cdots + \Delta x_N^2} = \left(\sum_{i=1}^{N}\Delta x_i^2\right)^{\frac{1}{2}} \\
&\vdots \\
&\left\|\Delta\vec{X}\right\|_l = \left(\sum_{i=1}^{N}\left|\Delta x_i\right|^l\right)^{\frac{1}{l}} \\
&\left\|\Delta\vec{X}\right\|_\infty = \max\left|\Delta x_i\right|
\end{aligned}
$$

تعریف گرادیان (Gradient)

$$
\vec{\nabla} f(\vec{X}) = \frac{\partial f}{\partial \vec{X}} = \begin{pmatrix} \dfrac{\partial f}{\partial x_1} \\ \dfrac{\partial f}{\partial x_2} \\ \vdots \\ \dfrac{\partial f}{\partial x_N} \end{pmatrix}
$$

تعریف هشن (Hessian)

$$
H = \frac{\partial^2 f}{\partial \vec{X}^2} = \begin{pmatrix} \dfrac{\partial^2 f}{\partial x_1^2} & \dfrac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \dfrac{\partial^2 f}{\partial x_1 \partial x_N} \\ \dfrac{\partial^2 f}{\partial x_2 \partial x_1} & \dfrac{\partial^2 f}{\partial x_2^2} & & \dfrac{\partial^2 f}{\partial x_2 \partial x_N} \\ \vdots & & \ddots & \vdots \\ \dfrac{\partial^2 f}{\partial x_N \partial x_1} & \dfrac{\partial^2 f}{\partial x_N \partial x_2} & \cdots & \dfrac{\partial^2 f}{\partial x_N^2} \end{pmatrix}
$$

شرط لازم (necessary condition) برای نقطه بهینه (مینیمم سازی)

$$
\begin{aligned}
\Delta f &= f(\vec{X}^* + \Delta\vec{X}) - f(\vec{X}^*) \\
&= \Delta\vec{X}^T \left.\frac{\partial f}{\partial \vec{X}}\right|_{\vec{X}^*} + \frac{1}{2!}\Delta\vec{X}^T \left.\frac{\partial^2 f}{\partial \vec{X}^2}\right|_{\vec{X}^*} \Delta\vec{X} + \ldots \geq 0
\end{aligned}
$$

اگر $\Delta\vec{X}$ می تواند آزاد باشد

$$\vec{\nabla} f(\vec{X}^*) = \left.\frac{\partial f}{\partial \vec{X}}\right|_{\vec{X}^*} = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \vdots \\ \frac{\partial f}{\partial x_N} \end{pmatrix}_{\vec{X}^*} = \vec{0}$$

می شود N معادله برای N مجهول

شرط کافی (sufficient condition) (برای مینیمم کردن)

$$\Delta\vec{X}^T \left.\frac{\partial^2 f}{\partial \vec{X}^2}\right|_{\vec{X}^*} \Delta\vec{X} \geq 0$$

یعنی ماتریس H مثبت معین باشد (Positive definite)

طبیعتا برای ماکزیمم کردن باید ماتریس H منفی معین باشد (Negative definite)

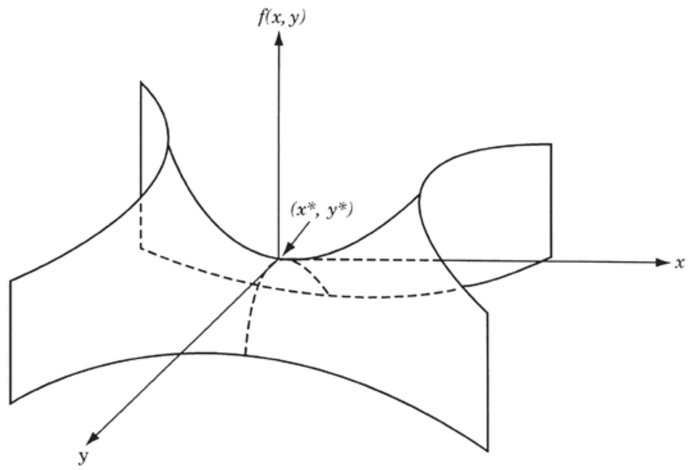
می توان با مقادیر ویژه چک کرد

اگر مقادیر ویژه مثبت و صفر باشد مثبت نیمه معین (Positive semi-definite)

باید رفت مشتقات بالاتر

اگر مقادیر ویژه مثبت و منفی باشد نامعین (Indefinite)

نقاط زینی (Saddle Point)

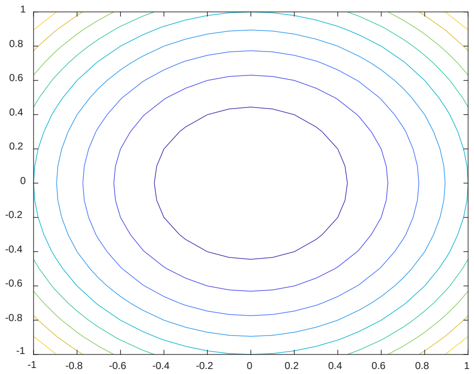


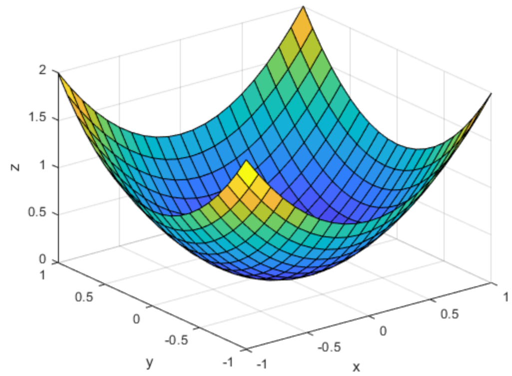
مثال:

$$\min f(\vec{X}) = x_1^2 + x_2^2 \qquad \Rightarrow \qquad \vec{\nabla} f(\vec{X}^*) = \left.\frac{\partial f}{\partial \vec{X}}\right|_{\vec{X}^*} = \begin{pmatrix} 2x_1 \\ 2x_2 \end{pmatrix}_{\vec{X}^*} = \vec{0}$$

$$\Rightarrow \begin{cases} x_1 = 0 \\ x_2 = 0 \end{cases}$$

$$H = \left.\frac{\partial^2 f}{\partial \vec{X}^2}\right|_{\vec{X}^*} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} > 0$$



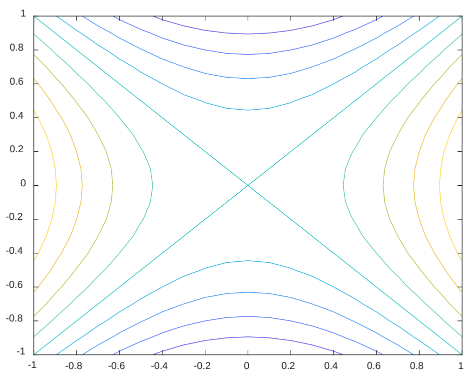


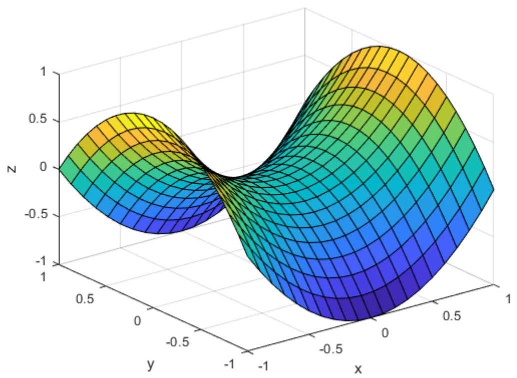
مثال:

$$\min f(\vec{X}) = x_1^2 - x_2^2 \qquad \Rightarrow \qquad \vec{\nabla} f(\vec{X}^*) = \left.\frac{\partial f}{\partial \vec{X}}\right|_{\vec{X}^*} = \begin{pmatrix} 2x_1 \\ -2x_2 \end{pmatrix}_{\vec{X}^*} = \vec{0}$$

$$\Rightarrow \begin{cases} x_1 = 0 \\ x_2 = 0 \end{cases}$$

$$H = \left.\frac{\partial^2 f}{\partial \vec{X}^2}\right|_{\vec{X}^*} = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix} \qquad \text{Indefinite}$$



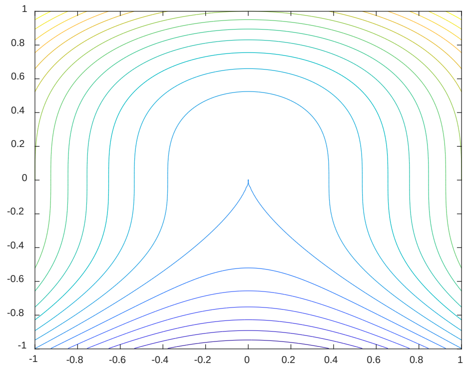


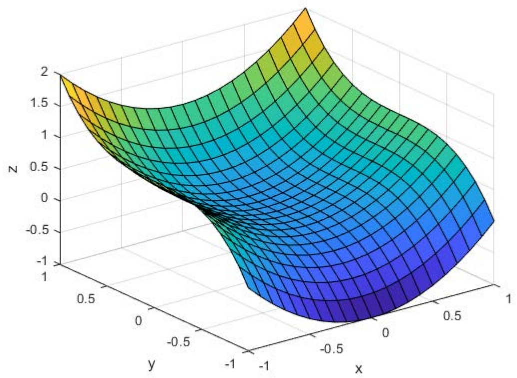
مثال:

$$\min f(\vec{X}) = x_1^2 + x_2^3 \qquad \Rightarrow \qquad \vec{\nabla} f(\vec{X}^*) = \left.\frac{\partial f}{\partial \vec{X}}\right|_{\vec{X}^*} = \begin{pmatrix} 2x_1 \\ 3x_2^2 \end{pmatrix}_{\vec{X}^*} = \vec{0}$$

$$\Rightarrow \begin{cases} x_1 = 0 \\ x_2 = 0 \end{cases}$$

$$H = \left.\frac{\partial^2 f}{\partial \vec{X}^2}\right|_{\vec{X}^*} = \begin{pmatrix} 2 & 0 \\ 0 & 6x_2 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{Positive semi-definite}$$



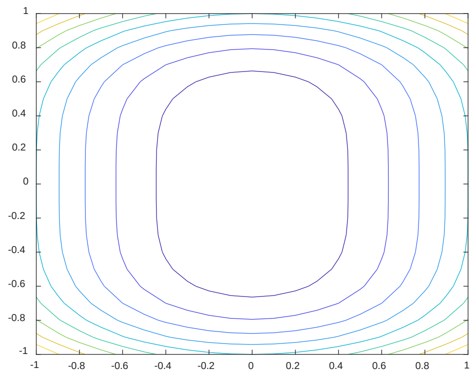


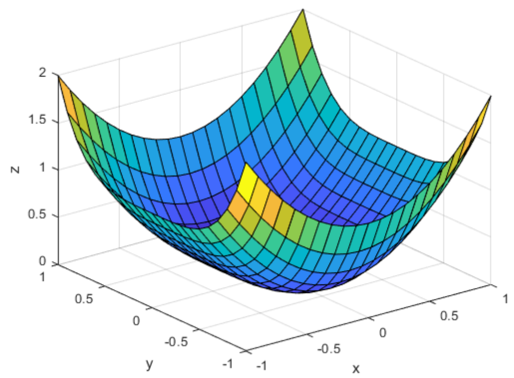
مثال:

$$\min f(\vec{X}) = x_1^2 + x_2^4 \qquad \Rightarrow \qquad \vec{\nabla} f(\vec{X}^*) = \left.\frac{\partial f}{\partial \vec{X}}\right|_{\vec{X}^*} = \begin{pmatrix} 2x_1 \\ 4x_2^3 \end{pmatrix}_{\vec{X}^*} = \vec{0}$$

$$\Rightarrow \begin{cases} x_1 = 0 \\ x_2 = 0 \end{cases}$$

$$H = \left.\frac{\partial^2 f}{\partial \vec{X}^2}\right|_{\vec{X}^*} = \begin{pmatrix} 2 & 0 \\ 0 & 12x_2^2 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{Positive semi-definite}$$





بهینه سازی چند پارامتری مقید

مقید به قید مساوی (Equality Constraint)

$$\begin{aligned} \min \quad & f(\vec{X}) \\ \text{subject to} \quad & g_i(\vec{X}) = 0 \qquad i = 1,\ldots,m < n \end{aligned}$$

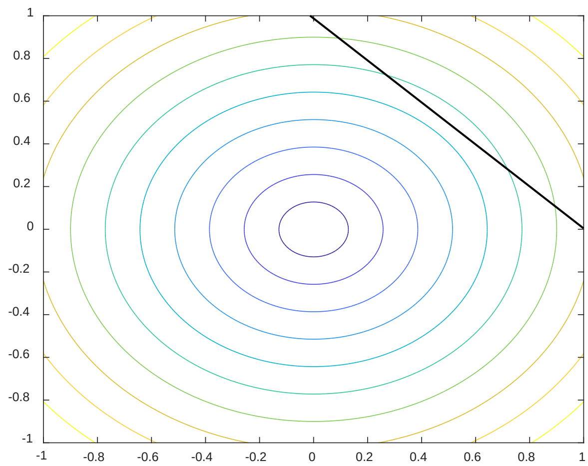
یا برداری

$$\begin{aligned} \min \quad & f(\vec{X}) \\ \text{subject to} \quad & \vec{g}(\vec{X}) = \vec{0} \end{aligned}$$

که $\vec{g}(\vec{X}) = \begin{pmatrix} g_1 & g_2 & \cdots & g_m \end{pmatrix}$

مثلا

$$\begin{aligned} \min \quad & f(\vec{X}) = x_1^2 + x_2^2 \\ \text{subject to} \quad & g(\vec{X}) = x_1 + x_2 - 1 = 0 \end{aligned}$$



جایگذاری مستقیم (Direct Substitiution)

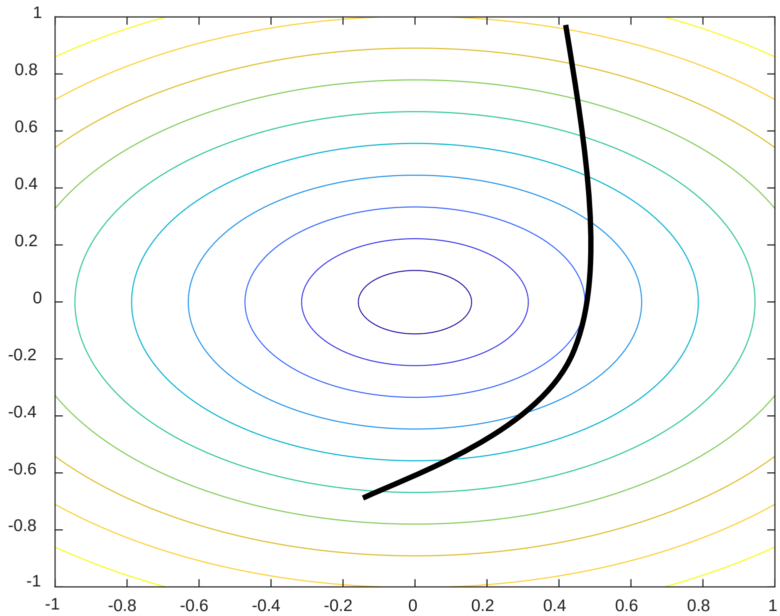
حل m معادله m مجهول $\vec{g}(\vec{X}) = \vec{0}$ و بدست آوردن m متغیر $\tilde{X}_m$ بر حسب بقیه n-m متغیر $\tilde{X}_{n-m}$

$$\min f(\tilde{X}_{n-m})$$

برای مثال

$$\begin{aligned} x_1 = 1 - x_2 \quad & \xrightarrow{f(\vec{X}) = x_1^2 + x_2^2} \quad f(\vec{X}) = f(x_2) = \left(1 - x_2\right)^2 + x_2^2 \\ \Rightarrow \quad & f' = -2(1 - x_2) + 2x_2 = 4x_2 - 2 = 0 \\ \Rightarrow \quad & x_1 = x_2 = \frac{1}{2} \end{aligned}$$

ضرایب لاگرانژ (Lagrange Multipliers)



جهت گرادیان بر روی کانتورهای f ثابت

بسط تیلور مرتبه اول

$$\Delta f = f(\vec{X}^* + \Delta\vec{X}) - f(\vec{X}^*) \simeq \Delta\vec{X}^T \left.\frac{\partial f}{\partial \vec{X}}\right|_{\vec{X}^*}$$

شرط بهینگی این است که کانتورهای f ثابت و g مماس باشند. یا $\vec{\nabla} f$ و $\vec{\nabla} g$ ها در یک راستا باشند

$$\vec{\nabla} f = -\lambda \vec{\nabla} g$$

وقتی چند تا قید g_i باشند

$$\vec{\nabla} f = -\sum_{i=1}^{m} \lambda_i \vec{\nabla} g_i = -\frac{\partial \vec{g}}{\partial \vec{X}}^T \vec{\lambda}$$

که

$$\frac{\partial \vec{g}}{\partial \vec{X}} = \begin{pmatrix} \frac{\partial g_1}{\partial x_1} & \frac{\partial g_1}{\partial x_2} & \cdots & \frac{\partial g_1}{\partial x_n} \\ \frac{\partial g_2}{\partial x_1} & \frac{\partial g_2}{\partial x_2} & \cdots & \frac{\partial g_2}{\partial x_n} \\ \vdots & & \ddots & \vdots \\ \frac{\partial g_m}{\partial x_1} & \frac{\partial g_m}{\partial x_2} & \cdots & \frac{\partial g_m}{\partial x_n} \end{pmatrix} \qquad \frac{\partial \vec{g}}{\partial \vec{X}}^T = \begin{pmatrix} \vec{\nabla} g_1 & \vec{\nabla} g_2 & \cdots & \vec{\nabla} g_m \end{pmatrix}$$

نهایتا حل بهینه شامل n+m معادله و n+m مجهول $\vec{X}$ و $\vec{\lambda}$

$$\vec{\nabla} f+\sum_{i=1}^{m}\lambda_i \vec{\nabla} g_i=\vec{0}$$
$$\vec{g}(\vec{X})=\vec{0}$$

معادل بهینه سازی لاگرانژین (Lagrangian) زیر

$$\min \mathcal{L}(\vec{X},\vec{\lambda})=f(\vec{X})+\vec{\lambda}^T\vec{g}=f(\vec{X})+\sum_{i=1}^{m}\lambda_i g_i$$

که شرط لازم بهینگی می شود

$$\vec{\nabla}\mathcal{L}=\vec{0} \qquad \Rightarrow \qquad \begin{cases}\vec{\nabla}_{\vec{X}}\mathcal{L}=\dfrac{\partial \mathcal{L}}{\partial \vec{X}}=\vec{\nabla} f+\sum_{i=1}^{m}\lambda_i\vec{\nabla} g_i=\vec{0}\\ \vec{\nabla}_{\vec{\lambda}}\mathcal{L}=\dfrac{\partial \mathcal{L}}{\partial \vec{\lambda}}=\vec{g}(\vec{X})=\vec{0}\end{cases}$$

برای مثال

$$\begin{array}{rl}\min & f(\vec{X})=x_1^2+x_2^2\\ \text{subject to} & g(\vec{X})=x_1+x_2-1=0\end{array}$$

$$\min \mathcal{L}(\vec{X},\vec{\lambda})=x_1^2+x_2^2+\lambda\left(x_1+x_2-1\right)$$

$$\begin{cases}\dfrac{\partial \mathcal{L}}{\partial x_1}=2x_1+\lambda=0\\ \dfrac{\partial \mathcal{L}}{\partial x_2}=2x_2+\lambda=0\\ \dfrac{\partial \mathcal{L}}{\partial \lambda}=x_1+x_2-1=0\end{cases} \qquad \Rightarrow \qquad \begin{cases}x_1=x_2=\dfrac{1}{2}\\ \lambda=-1\end{cases}$$

روش تغییرات مقید (Constrained variation)

مثلا برای دو متغیره

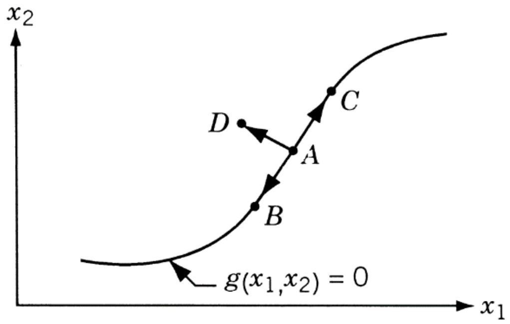
$$\begin{array}{rl}\min & f(x_1,x_2)\\ \text{subject to} & g(x_1,x_2)=0\end{array}$$

شرط لازم بهینگی

$$df=\frac{\partial f}{\partial x_1}dx_1+\frac{\partial f}{\partial x_2}dx_2=0$$

ولی به شرط اینکه روی قید باشد

$$g\left(x_1^*+dx_1,x_2^*+dx_2\right)=0$$



با بسط تیلور مرتبه اول

$$g\left(x_1^* + dx_1, x_2^* + dx_2\right) \simeq g\left(x_1^*, x_2^*\right) + \frac{\partial g}{\partial x_1}\left(x_1^*, x_2^*\right) dx_1 + \frac{\partial g}{\partial x_2}\left(x_1^*, x_2^*\right) dx_2 = 0$$

و نتیجتا تغییرات مقید باید در رابطه زیر صدق کند

$$dg = \left(\frac{\partial g}{\partial x_1} dx_1 + \frac{\partial g}{\partial x_2} dx_2\right)\Bigg|_{\left(x_1^*, x_2^*\right)} = 0 \quad \Rightarrow \qquad dx_2 = -\frac{\partial g / \partial x_1}{\partial g / \partial x_2}\Bigg|_{\left(x_1^*, x_2^*\right)} dx_1$$

و با جایگذاری در شرط لازم

$$df = \left(\frac{\partial f}{\partial x_1} - \frac{\partial g / \partial x_1}{\partial g / \partial x_2}\frac{\partial f}{\partial x_2}\right)\Bigg|_{\left(x_1^*, x_2^*\right)} dx_1 = 0$$

دیگر dx_1 آزاد است

$$\left(\frac{\partial f}{\partial x_1} - \frac{\partial g / \partial x_1}{\partial g / \partial x_2}\frac{\partial f}{\partial x_2}\right)\Bigg|_{\left(x_1^*, x_2^*\right)} = 0$$

به عبارتی دو معادله و دو مجهول زیر:

$$g = 0$$

$$\begin{vmatrix} \dfrac{\partial f}{\partial x_1} & \dfrac{\partial f}{\partial x_2} \\ \dfrac{\partial g}{\partial x_1} & \dfrac{\partial g}{\partial x_2} \end{vmatrix} = 0$$

مشابهت با ضرایب لاگرانژ

$$\left.\left(\frac{\partial f}{\partial x_1}-\frac{\partial f/\partial x_2}{\partial g/\partial x_2}\frac{\partial g}{\partial x_1}\right)\right|_{\left(x_1^*,x_2^*\right)}=0 \qquad \longrightarrow \qquad \lambda=-\frac{\partial f/\partial x_2}{\partial g/\partial x_2}$$

$$\left.\left(\frac{\partial f}{\partial x_2}-\frac{\partial f/\partial x_1}{\partial g/\partial x_1}\frac{\partial g}{\partial x_2}\right)\right|_{\left(x_1^*,x_2^*\right)}=0 \qquad \longrightarrow \qquad \lambda=-\frac{\partial f/\partial x_1}{\partial g/\partial x_1}$$

در حالت عمومی

مشابه حل m معادله m مجهول $\vec{g}(\vec{X})=\vec{0}$ و بدست آوردن m متغیر $\tilde{X}_m$ بر حسب بقیه n-m متغیر $\tilde{X}_{n-m}$ و نهایتا بهینه سازی برای $\min f(\tilde{X}_{n-m})$

حالا در حقیقت تغییرات $dx_1,dx_2,\ldots,dx_m$ بر حسب $dx_{m+1},dx_{m+2},\ldots,dx_n$ نوشته می شود

$$J_k\left(\frac{f,g_1,g_2,\ldots,g_m}{x_k,x_1,x_2,x_3,\ldots,x_m}\right)=\begin{vmatrix}\frac{\partial f}{\partial x_k} & \frac{\partial f}{\partial x_1} & \frac{\partial f}{\partial x_2} & \cdots & \frac{\partial f}{\partial x_m}\\ \frac{\partial g_1}{\partial x_k} & \frac{\partial g_1}{\partial x_1} & \frac{\partial g_1}{\partial x_2} & \cdots & \frac{\partial g_1}{\partial x_m}\\ \frac{\partial g_2}{\partial x_k} & \frac{\partial g_2}{\partial x_1} & \frac{\partial g_2}{\partial x_2} & \cdots & \frac{\partial g_2}{\partial x_m}\\ \vdots & & & & \\ \frac{\partial g_m}{\partial x_k} & \frac{\partial g_m}{\partial x_1} & \frac{\partial g_m}{\partial x_2} & \cdots & \frac{\partial g_m}{\partial x_m}\end{vmatrix}=0 \qquad k=m+1,m+2,\ldots,n$$

مجموعه معادلات

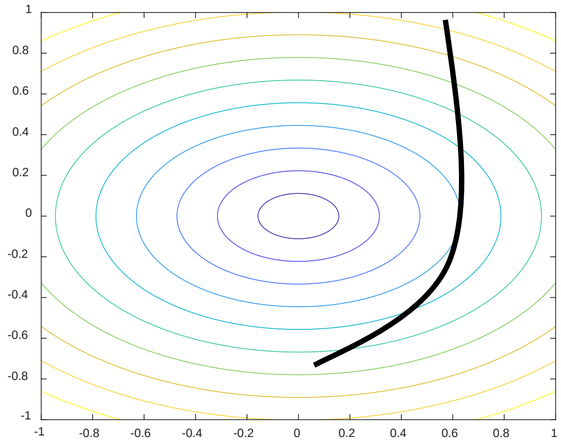
$$g_i=0 \qquad i=1,\cdots,m$$

$$J_k\left(\frac{f,g_1,g_2,\ldots,g_m}{x_k,x_1,x_2,x_3,\ldots,x_m}\right)=0 \qquad k=m+1,m+2,\ldots,n$$

باید ترکیبی از متغیرهای اولیه (m تا x اول) انتخاب شود که

$$J\left(\frac{g_1,g_2,\ldots,g_m}{x_1,x_2,\ldots,x_m}\right)\neq 0$$

تعبیر ضریب لاگرانژ



برای بهینه سازی مقید زیر

$$\begin{aligned} \min \quad & f(\vec{X}) \\ \text{subject to} \quad & g(\vec{X}) = 0 \end{aligned}$$

تبدیل به لاگرانژین

$$\min \mathcal{L}(\vec{Y}) = \mathcal{L}(\vec{X},\vec{\lambda}) = f(\vec{X}) + \lambda\, g(\vec{X})$$

که شرط لازم بهینگی می شود

$$\begin{cases} \dfrac{\partial \mathcal{L}}{\partial \vec{X}} = \left(\vec{\nabla} f + \lambda \vec{\nabla} g\right)_{\left(\vec{X}^*,\vec{\lambda}^*\right)} = \vec{0} \\ \dfrac{\partial \mathcal{L}}{\partial \lambda} = g(\vec{X}^*) = 0 \end{cases}$$

به قید اجازه کمی تغییر دهیم $g(\vec{X}) = b$ تا ببینیم جواب بهینه چقدر جابجا می شود

$$\begin{aligned} \min \quad & f(\vec{X}) \\ \text{subject to} \quad & \tilde{g}(\vec{X}) = b - g(\vec{X}) = 0 \end{aligned}$$

تبدیل به لاگرانژین

$$\min \mathcal{L}(\vec{X},\vec{\lambda}) = f(\vec{X}) + \lambda\, \tilde{g} = f(\vec{X}) + \lambda \Big(b - g(\vec{X})\Big)$$

که شرط لازم بهینگی می شود

$$\begin{cases} \dfrac{\partial \mathcal{L}}{\partial \vec{X}} = \left(\vec{\nabla} f - \lambda \vec{\nabla} g\right)_{\left(\vec{X}^* + \Delta\vec{X}^*,\vec{\lambda}^*\right)} = \vec{0} \\ \dfrac{\partial \mathcal{L}}{\partial \vec{\lambda}} = b - g(\vec{X}^* + \Delta\vec{X}^*) = 0 \end{cases}$$

با توجه به روابط زیر:

$$\begin{aligned} df^* &= \vec{\nabla} f \cdot \Delta \vec{X}^* \\ dg = db &= \vec{\nabla} g \cdot \Delta \vec{X}^* \end{aligned}$$

رابطه اول بهینگی می شود:

$$d\mathcal{L} = df - \lambda dg = 0 \quad \Rightarrow \quad \lambda = \frac{df}{db}$$

بهینه سازی چند پارامتری با قید نامساوی (Inequality Constraint)

$$\begin{aligned} &\min && f(\vec{X}) \\ &\text{subject to} && g_i(\vec{X}) \le 0 \qquad i = 1,\dots,m \end{aligned}$$

تبدیل به قیود مساوی با slack variables

$$\begin{aligned} &\min && f(\vec{X}) \\ &\text{subject to} && g_i(\vec{X}) + y_i^2 = 0 \qquad i = 1,\dots,m \end{aligned}$$

تشکیل لاگرانژین (تبدیل به مساله نامقید)

$$\min \mathcal{L}(\vec{X},\vec{\lambda},\vec{Y}) = f(\vec{X}) + \sum_{i=1}^{m} \lambda_i \left(g_i + y_i^2 \right)$$

شروط لازم

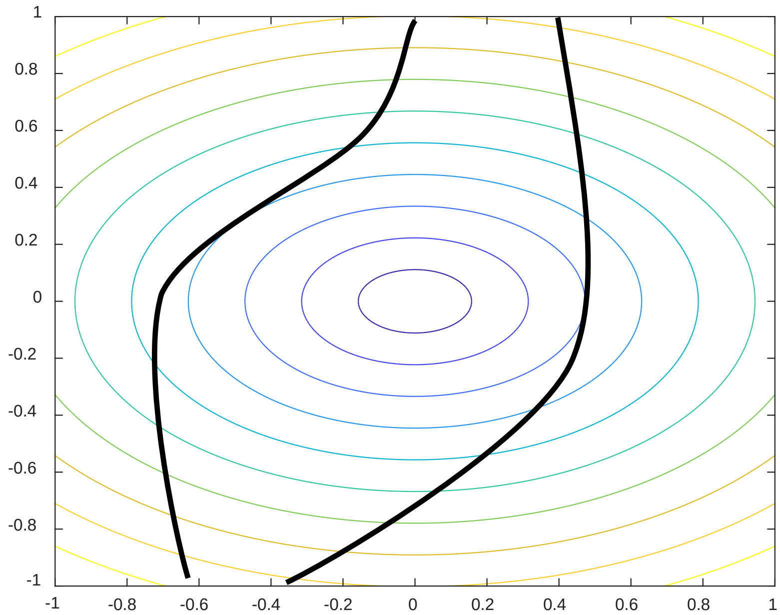
$$\begin{cases} \dfrac{\partial \mathcal{L}}{\partial \vec{X}} = \vec{\nabla} f + \sum_{i=1}^{m} \lambda_i \vec{\nabla} g_i = \vec{0} \\ \dfrac{\partial \mathcal{L}}{\partial \lambda_i} = g_i + y_i^2 = 0 \qquad i = 1,\dots,m \\ \dfrac{\partial \mathcal{L}}{\partial y_i} = 2\lambda_i y_i = 0 \qquad i = 1,\dots,m \end{cases}$$

جمعا $2m+n$ معادله برای $2m+n$ مجهول $\vec{X},\vec{\lambda},\vec{Y}$

دو حالت برای معادله آخر

۱- $y_i = 0$ قید نامساوی فعال (Active constraint) است و مانند قید مساوی شده است

۲- $\lambda_i = 0$ قید غیرفعال (Inactive constraint) است



تعیین علامت ضرایب لاگرانژ قیود نامساوی فعال

برای قیود غیرفعال $\lambda_i = 0$

معادله اول شروط لازم برای قیود فعال

$$\frac{\partial \mathcal{L}}{\partial \vec{X}} = \vec{\nabla} f + \sum_{i=1}^{m} \lambda_i \vec{\nabla} g_i = \vec{\nabla} f + \sum_{i \in s_{active}} \lambda_i \vec{\nabla} g_i = \vec{0}$$

$$\Rightarrow \qquad \vec{\nabla} f = -\sum_{i \in s_{active}} \lambda_i \vec{\nabla} g_i$$

اگر $\vec{S}$ یک بردار دلخواه باشد

$$\vec{S} \cdot \vec{\nabla} f = -\sum_{i \in s_{active}} \lambda_i \vec{S} \cdot \vec{\nabla} g_i$$

اگر $\vec{S}$ در جهت مجاز (feasible direction) باشد:

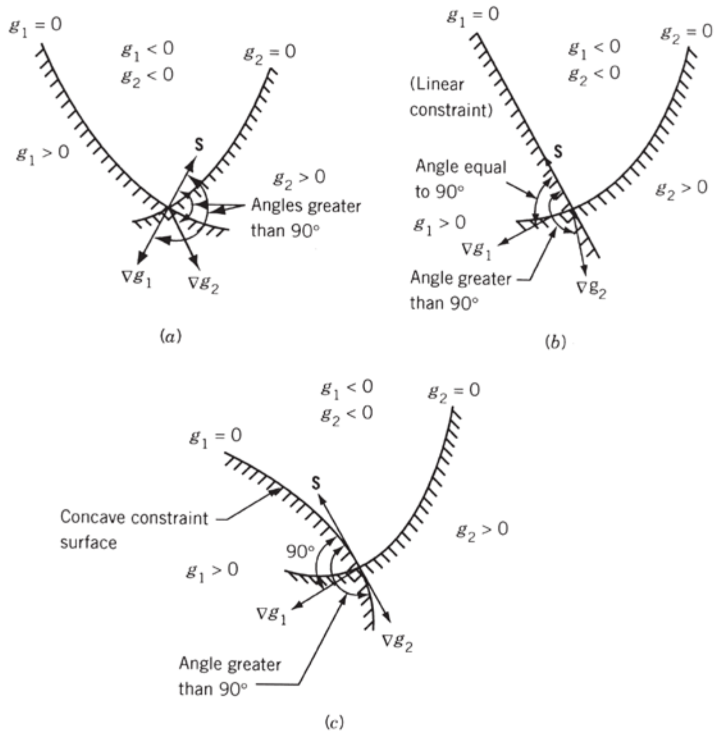
$$\vec{S} \cdot \vec{\nabla} g_i < 0$$

اگر نقطه بهینه باشد (حتما در جهت مجاز f باید زیاد شود)

$$\vec{S} \cdot \vec{\nabla} f > 0$$

نتیجتا همه λ_i ها باید مثبت باشند

$$\lambda_i > 0$$



بهینه سازی چند پارامتری با قیود مساوی و نامساوی

$$
\begin{array}{lll}
\min & f(\vec{X}) & \\
\text{subject to} & g_i(\vec{X}) \le 0 & i = 1,\dots,m \\
 & h_j(\vec{X}) = 0 & j = 1,\dots,p
\end{array}
$$

تبدیل به قیود مساوی با slack variables

$$
\begin{array}{lll}
\min & f(\vec{X}) & \\
\text{subject to} & g_i(\vec{X}) + y_i^2 = 0 & i = 1,\dots,m \\
 & h_j(\vec{X}) = 0 & j = 1,\dots,p
\end{array}
$$

تشکیل لاگرانژین (تبدیل به مساله نامقید)

$$
\min \mathcal{L}(\vec{X},\vec{\lambda},\vec{Y},\vec{\mu}) = f(\vec{X}) + \sum_{i=1}^{m} \lambda_i \left(g_i + y_i^2\right) + \sum_{j=1}^{p} \mu_j h_j
$$

شروط لازم بهینگی

$$
\begin{cases}
\dfrac{\partial \mathcal{L}}{\partial \vec{X}} = \vec{\nabla} f + \sum_{i=1}^{m} \lambda_i \vec{\nabla} g_i + \sum_{j=1}^{p} \mu_j \vec{\nabla} h_j = \vec{0} \\
\dfrac{\partial \mathcal{L}}{\partial \lambda_i} = g_i + y_i^2 = 0 & i = 1,\dots,m \\
\dfrac{\partial \mathcal{L}}{\partial y_i} = 2\lambda_i y_i = 0 & i = 1,\dots,m \\
\dfrac{\partial \mathcal{L}}{\partial \mu_j} = h_j = 0 & j = 1,\dots,p
\end{cases}
$$

جمعا $2m+n+p$ معادله برای $2m+n+p$ مجهول $\vec{X}, \vec{\lambda}, \vec{Y}, \vec{\mu}$

شروط Kuhn–Tucker

یا Karush–Kuhn–Tucker conditions

برای مساله

$$
\begin{array}{lll}
\min & f(\vec{X}) & \\
\text{subject to} & g_i(\vec{X}) \le 0 & i = 1,\dots,m \\
 & h_j(\vec{X}) = 0 & j = 1,\dots,p
\end{array}
$$

شروط لازم بهینگی

$$
\left.
\begin{cases}
\vec{\nabla} f + \sum_{i=1}^{m} \lambda_i \vec{\nabla} g_i + \sum_{j=1}^{p} \mu_j \vec{\nabla} h_j = \vec{0} \\
g_i \le 0 & i = 1,\dots,m \\
\lambda_i \ge 0 & i = 1,\dots,m \\
h_j = 0 & j = 1,\dots,p
\end{cases}
\right\} \qquad \lambda_i g_i = 0 \qquad i = 1,\dots,m
$$