Tutorial 8

05 March 2022 11:01

$\min f(x)$

s.t. $h(x) = \underline{0}$

$f: \mathbb{R}^n \to \mathbb{R}$

$h(x) = \begin{bmatrix} h_1(x) \\ \vdots \\ h_m(x) \end{bmatrix}$

$h: \mathbb{R}^n \to \mathbb{R}^m$

Necessary conditions

Lagrangian function

$L: \mathbb{R}^{n+m} \to \mathbb{R}$

$L(x, \lambda) = f(x) + \sum_{i=1}^{m} \lambda_i h_i(x)$

If x^* is a local minimum that is regular, then $\exists\, \lambda^*$ s.t.

($\nabla h_1(x^*), \ldots, \nabla h_m(x^*)$ are linearly independent.

$\nabla_x L(x^*, \lambda^*) = \underline{0}$ —— (1)

$\left(\nabla f(x^*) + \sum_{i=1}^{m} \lambda_i^* h_i(x^*) = 0 \right)$

$\nabla_\lambda L(x^*, \lambda^*) = \underline{0}$ —— (2)

$\left(h_i(x^*) = 0 \quad \forall\, i = 1, \ldots, m \right)$

$y^T \nabla_{xx}^T L(x^*, \lambda^*) y \geq 0 \quad \forall\, y: \underbrace{y^T \nabla h_i(x^*) = 0}\ \forall\, i = 1, \ldots, m$ —— (3)

$\left(y^T \left(\nabla^2 f(x^*) + \sum_{i=1}^{m} \nabla^2 h_i(x^*) \right) y \geq 0 \right)$

(1) & (2) are called first order necessary conditions
($~~n+~~m$ equations in $n+m$ variables)

(3) is called second order necessary condition

Sufficient conditions

If $\exists\, x^*$ and λ^* satisfying

$\nabla_x L(x^*, \lambda^*) = 0$

$\nabla_\lambda L(x^*, \lambda^*) = 0$

and $y^T \nabla_{xx}^2 L(x^*, \lambda^*) y > 0 \quad \forall \begin{cases} y: y^T \nabla h_i(x) = 0 \ \forall\, i = 1, \ldots, m \\ y \neq 0 \end{cases}$

$\Rightarrow x^*$ is a strict local minimum

Sensitivity theorem

Sensitivity theorem

$$\min f(x)$$
$$\text{s.t. } h(x) = c \qquad c \in \mathbb{R}^m$$

- $(x^*(c), \lambda^*(c))$ is a local minimum, lagrange multiplier pair that satisfies sufficiency condition
- $x^*(0)$ is regular
- $p(c)$ is the optimal value

$$\boxed{\nabla p(c) = -\lambda^*(c)}$$

$$\begin{bmatrix} \frac{\partial p(c)}{\partial c_1} \\ \vdots \\ \frac{\partial p(c)}{\partial c_m} \end{bmatrix} \in \mathbb{R}^m$$

Example $\min \frac{1}{2}(x_1^2 - x_2^2) - x_2$

s.t. $x_2 = 0$

$$\nabla f(x^*) + \lambda^* \nabla h(x^*) = 0$$

$$\begin{bmatrix} x_1^* \\ -x_2^* - 1 \end{bmatrix} + \lambda^* \begin{bmatrix} 0 \\ 1 \end{bmatrix} = 0$$

$$\left.\begin{aligned} x_1^* &= 0 \\ -x_2^* - 1 + \lambda^* &= 0 \\ x_2^* &= 0 \end{aligned}\right\}$$

$$\Rightarrow x^* = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad \lambda^* = 1$$

Also, x^* is regular

$$\nabla^2_{xx} L(x^*, \lambda^*) = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

$$y^T \nabla^2_{xx} L(x^*, \lambda^*) y > 0 \qquad \forall y : \left.\begin{bmatrix} y_1 & y_2 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = 0 \atop y \neq 0 \right\} \Rightarrow \begin{aligned} y_2 &= 0 \\ y &\neq 0 \end{aligned}$$

i.e., $y^T \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} y > 0$

i.e., $y' \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} y > 0$

i.e. $[y_1 \quad y_2] \begin{bmatrix} y_1 \\ -y_2 \end{bmatrix} > 0$

i.e. $y_1^2 - y_2^2 > 0$

– this condition is satisfied.

$\Rightarrow x^* = (0,0)$ is unique global minimum.

Now consider

$$\min \frac{1}{2}(x_1^2 - x_2^2) - x_2$$

s.t. $x_2 = u$

$$p(u) = \min \frac{1}{2}(x_1^2 - u^2) - u$$
$$= -\frac{1}{2}u^2 - u$$

$$p'(u) = -u - 1$$

From sensitivity theorem we expect that

$$p'(0) = -\lambda^*$$

this is true as both equal -1.

Inequality Constraints

$f: \mathbb{R}^n \to \mathbb{R}$

$$\min f(x)$$
$$\text{s.t. } h_i(x) = 0, \quad i = 1, \ldots, m$$
$$g_j(x) \le 0, \quad j = 1, \ldots, r$$

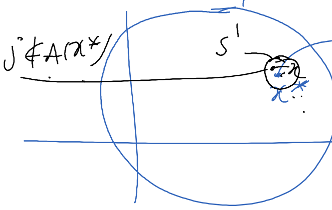
– Let x^* be a local minimum.
– We want to come up with Necessary conditions.

Suppose $h_i(x^*) = 0 \quad i = 1, \ldots, m$
$g_j(x^*) = 0 \quad j \in A(x^*) \subseteq \{1, 2, \ldots, r\}$

$g_j(x^*) < 0 \quad \forall j \notin A(x^*)$

($A(x^*)$ is the set of active constraints at x^*).

$S = \{ y : h(x) = 0,\ g(x) \le 0 \}$



$$\begin{aligned} \min\ & f(x) \\ \text{s.t. } & h_i(x) = 0 \quad \forall i = 1, \dots m \\ & g_j(x) \le 0 \quad \forall j \in A(x^*) \end{aligned}$$

$\Rightarrow x^*$ is a local min. of this problem.

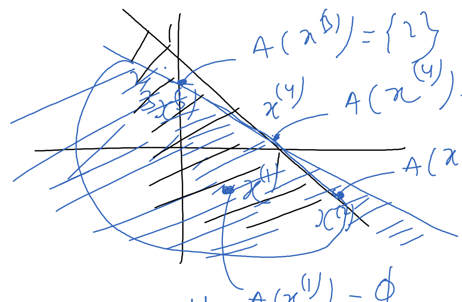
Example

$g_1(x) = x_1 + x_2 \le 1$

$g_2(x) \quad 2x_1 + 3x_2 \le 2$

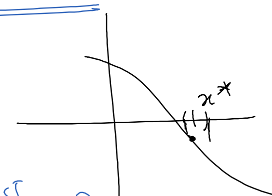
Consider the following problem

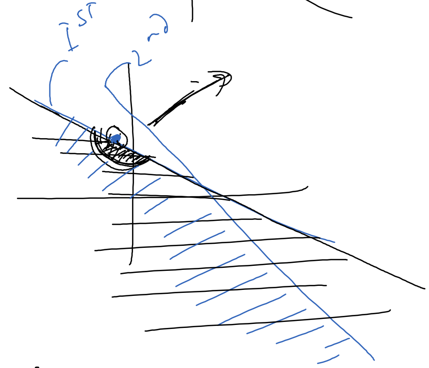
$$\begin{aligned} \min\ & f(x) \\ \text{s.t. } & h_i(x) = 0 \quad i = 1, \dots m \\ & g_j(x) = 0 \quad \forall j \in A(x^*) \end{aligned}$$



if x^* is a local minimum of the original problem

$\Rightarrow x^*$ is a local minimum of this problem as well.





there exist $\lambda^* = \begin{bmatrix} \lambda_1^* \\ \vdots \\ \lambda_m^* \end{bmatrix}$ and $\mu^* = \begin{bmatrix} \mu_1^* \\ \vdots \\ \mu_r^* \end{bmatrix}$ s.t.

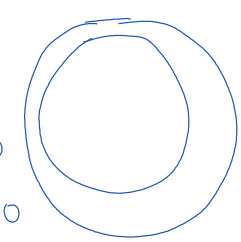
$$f(x^*) + \sum_{i=1}^{m} \lambda_i^* h_i(x^*) + \sum_{j \in A(x^*)} \mu_j^* g_j(x^*) = 0$$

$$\begin{aligned} \min_{x \in \mathbb{R}^n}\ & f(x) \\ \text{s.t. } & h_i(x) = 0 \quad \forall i = 1, \dots m \\ & \dots \quad i, \dots r \end{aligned}$$

$f(x^*)$

$x^* \quad h_i(x^*) = 0$

$g_i(x^*) \le 0$



$x \in \mathbb{R}^n$

s.t. $h_i(x) = 0 \quad \forall i = 1, \dots m$

$g_j(x) \le 0 \quad \forall j = 1, \dots r$

$h_i(x^*) = 0$

$g_j(x^*) \le 0$

define $s_j^* = \sqrt{-g_j(x^*)} \quad \forall j = 1, \dots r$

An equivalent problem is

$$\begin{aligned} &\min_{x \in \mathbb{R}^n, s \in \mathbb{R}^r} f(x) \\ &\text{s.t. } h_i(x) = 0 \quad \forall i = 1, \dots m \\ &\quad g_j(x) + s_j^2 = 0 \quad \forall j = 1, \dots r. \end{aligned}$$

(x^*, s^*) will satisfy these constraints

and a relaxed problem is

$$\begin{aligned} &\min_{x, s} f(x) \\ &h_i(x) = 0 \quad \forall i = 1, \dots m \\ &g_j(x) + s_j^2 = c_j \quad \text{for some } j \\ &g_j(x) + s_j^2 = 0 \quad \text{for other } j \end{aligned}$$

$(c_j > 0)$

— suppose $x^*(c)$ is a soln to this

let (λ^*, μ^*) be the corresponding Lagrange multiplier

$g_j(x) \le c_j$, $c_j > 0$

x^* is soln to the original problem

$g_j(x^*) < 0$

$$f(x^*(c)) \le f(x^*)$$

$$\Leftrightarrow \cancel{f(x^*)} + \underbrace{\frac{\partial f(x^*(0))}{\partial c_j}(c_j - 0)}_{\text{1st order approximation}} \le \cancel{f(x^*)}$$

$$\Leftrightarrow -\mu_j^* c_j \le 0$$

$$\Leftrightarrow \mu_j^* \ge 0$$

(Lagrange multiplier's corresponding to non-negativity constraints are non-negative)

* $\mu_j^* = 0 \quad \forall j \notin A(x^*)$

So, the necessary condition.

$$f(x^*) + \sum^m \lambda_i^* h_i(x^*) + \sum \mu_j^* g_j(x^*) = 0$$

$$f(x^*) + \sum_{i=1}^{m} \lambda_i^* h_i(x^*) + \sum_{j \in A(x^*)} \mu_j^* g_j(x^*) = 0$$

is equivalent to

$$f(x^*) + \sum_{i=1}^{m} \lambda_i^* h_i(x^*) + \sum_{j=1}^{r} \mu_j^* g_j(x^*) = 0$$

(First order necessary conditions for inequality constrained problems)

$$\begin{cases} \min f(x) \\ \text{s.t. } h_i(x) = c_i \quad i=1,\dots,m \end{cases}$$

$\lambda^*(c) \in \mathbb{R}^m$ is the corresponding L.M.

$$\boxed{\frac{\partial f(x^*(c))}{\partial c_i} = -\lambda_i^*(c)} \quad \forall i=1,\dots,m$$