Lagrange Multipliers - Equality Constraints

First-order necessary condition

Let x^* be a local min. of $f(x)$ s.t. $h(x) = 0$
If $\nabla h_1(x^*), \nabla h_2(x^*), \ldots, \nabla h_m(x^*)$ are l.i.,
then $\exists$ a unique $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_m)$ s.t.

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

Second-order necessary condition

- $g^{(k)}(x) = f(x) + \frac{k}{2} \|h(x)\|^2 + \frac{\alpha}{2} \|x - x^*\|^2, \quad k = 1, 2, \ldots$
- $x^{(k)} = \arg\min_{x \in \mathcal{S}} g^{(k)}(x), \quad \mathcal{S} = \{x : \|x - x^*\| \le \varepsilon\}$
- $\lim_{k \to \infty} x^{(k)} = x^*$, and $x^{(k)}$ is an interior point of $\mathcal{S}$ for suff. large k.
- Thus using second-order necessary condition for unconstrained minimization of $g^{(k)}(x)$,

  we have $\nabla^2 g^{(k)}(x^{(k)}) \succeq 0$, for k suff. large

- $\nabla g^{(k)}(x) = \nabla f(x) + k \sum_{i=1}^{m} h_i(x) \nabla h_i(x) + \alpha (x - x^*)$

$$\Rightarrow \nabla^2 g^{(k)}(x) = \nabla^2 f(x) + k \sum_{i=1}^{m} \nabla h_i(x) \nabla h_i(x)^T$$
$$+ k \sum_{i=1}^{m} h_i(x) \nabla^2 h_i(x) + \alpha I$$

$$\nabla^2 g^{(k)}(x) = \nabla^2 f(x) + k \, \nabla h(x) \nabla h(x)^T + k \sum_{i=1}^{m} h_i(x) \nabla^2 h_i(x) + \alpha I$$

$$\nabla^2 g^{(k)}(x^{(k)}) \succcurlyeq 0$$

$$\Rightarrow \nabla^2 f(x^{(k)}) + k \, \nabla h(x^{(k)}) \nabla h(x^{(k)})^T + k \sum_{i=1}^{m} h_i(x^{(k)}) \nabla^2 h_i(x^{(k)}) + \alpha I \succcurlyeq 0 \qquad (1)$$

Consider $z \in \mathcal{V}^g(x^*)$, i.e., $\nabla h_i(x^*)^T z = 0, \; i = 1, \dots, m.$

Let $z^{(k)} = z - \nabla h(x^{(k)}) \underbrace{\left(\nabla h(x^{(k)})^T \nabla h(x^{(k)}) \right)^{-1} \nabla h(x^{(k)})^T}_{\text{pseudo-inverse of } \nabla h(x^{(k)})} z$

Then $\nabla h(x^{(k)})^T z^{(k)} = \nabla h(x^{(k)})^T z - \nabla h(x^{(k)})^T z$

$= 0.$

and (1) implies

$$(z^{(k)})^T \left(\nabla^2 f(x^{(k)}) + k \sum_{i=1}^{m} h_i(x^{(k)}) \nabla^2 h_i(x^{(k)}) + \alpha I \right) z^{(k)} \geq 0$$

As $k \to \infty$, $x^{(k)} \to x^*$, $k h_i(x^{(k)}) \to \lambda_i$, and $z^{(k)} \to z$

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

$$\forall z \in \mathcal{V}^g(x^*)$$

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

$$\forall\, z \in \mathcal{V}(x^*)$$

Taking $\alpha \to 0$,

$$\boxed{z^T\left(\nabla^2 f(x^*) + \sum_{i=1}^{m} \lambda_i \nabla^2 h_i(x^*)\right) z \geq 0 \quad \forall z \in \mathcal{V}(x^*)}$$

Similarly, we can show the following second order <u>sufficient condition</u>:

For x^* that is feasible and regular if $\exists\, \lambda$ s.t.

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

and

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

$$\text{for all } z \in \mathcal{V}(x^*),\ z \neq 0$$

Then x^* is a (strict) local min for

$$\text{minimize } f(x)$$

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

Lagrangian Function

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

Necessary conditions in terms of $L(x,\lambda)$

If x^* is a local min. and regular,

then $\exists\ \lambda$ s.t.

$$\left.\begin{aligned} \nabla_x L(x^*,\lambda) &= 0 \\ h(x^*) = \nabla_\lambda L(x^*,\lambda) &= 0 \end{aligned}\right\}$$

$m+n$ eqns in $m+n$ unknowns x^*, λ

$$z^T \nabla^2_{xx} L(x^*,\lambda)\, z \geq 0 \quad \forall\ z \in \mathcal{V}(x^*)$$

(i.e. $\forall\ z$ s.t. $\nabla h_i(x^*)^T z = 0\ \forall i$)

Sufficient Conditions in terms of $L(x,\lambda)$

For some x^* that is feasible and regular,

if $\exists\ \lambda$ s.t.

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

and $z^T \nabla^2_{xx} L(x^*,\lambda)\, z > 0$, $\forall z \neq 0, z \in \mathcal{V}(x^*)$

then x^* is local min. for

minimize $f(x)$ s.t. $h(x) = 0$.

Example Minimize $\frac{1}{2}(x_1^2 + x_2^2 + x_3^2)$ — (P1)

subject to $x_1 + x_2 + x_3 = 3$

$h(x) = x_1 + x_2 + x_3 - 3$

$L(x, \lambda) = \frac{1}{2}(x_1^2 + x_2^2 + x_3^2) + \lambda(x_1 + x_2 + x_3 - 3)$

$\nabla f(x) = [x_1, x_2, x_3]^T$

$\nabla h(x) = [1 \quad 1 \quad 1]^T \leftarrow$ rank 1

$\Rightarrow$ every x is regular

First-order necc. condition : $\nabla_x L(x^*, \lambda) = 0$

$\Rightarrow \quad x_1^* + \lambda = 0, \quad x_2^* + \lambda = 0, \quad x_3^* + \lambda = 0$

$\Rightarrow \quad x_1^* = x_2^* = x_3^* = -\lambda$

$h(x^*) = 0 \Rightarrow -3\lambda - 3 = 0 \Rightarrow \lambda = -1$

$\Rightarrow x^* = [1 \quad 1 \quad 1]$

$\Rightarrow x^*$ is the unique candidate for local min.

$\nabla_{xx}^2 L(x, \lambda) = \nabla^2 f(x) + \lambda \nabla^2 h(x)$

$= I + 0 = I \succ 0$

$\Rightarrow x^*$ is (strict) local min for (P1)

Is x^* also the global min of $f(x)$ s.t. $h(x)=0$?
f is convex and $\mathcal{H} = \{x : h(x)=0\}$ is convex and closed. Therefore x^* is also global min for (P1).

Alternative proof Recall from lec 1,

Weierstrass' Theorem (Extreme value Theorem)

If f is a continuous function on a compact set, $\mathcal{S} \subseteq \mathbb{R}^n$, then f attains its min and max on $\mathcal{S}$, i.e.,

$$\exists\, x_1 \in \mathcal{S} \text{ s.t. } f(x_1) = \inf_{x \in \mathcal{S}} f(x)$$

$$\exists\, x_2 \in \mathcal{S} \text{ s.t. } f(x_2) = \sup_{x \in \mathcal{S}} f(x)$$

Corollary Let f be continuous on closed set $\mathcal{S}$, that is not necessarily bounded. If f $(-f)$ is coercive on $\mathcal{S}$, it attains its min (max) on $\mathcal{S}$.

Application to (P1)

$f(x) = \frac{1}{2}(x_1^2 + x_2^2 + x_3^2)$ is coercive on $\mathcal{H}$

constraint set $\mathcal{H} = \{x : x_1 + x_2 + x_3 = 3\}$ is closed

$\Rightarrow$ f achieves its global min on $\mathcal{H}$

Since there is only one local min of f on $\mathcal{H}$, it must be the global min.

Sensitivity Analysis

Let x^*, which is regular, and λ be a local min. and Lagrange multiplier vector satisfying the second order sufficiency conditions for the problem:

$$\text{minimize } f(x) \quad \text{s.t. } h(x) = 0 .$$

Now suppose we perturb the constraint from 0 to $u \in \mathbb{R}^m$, i.e., we wish to

$$\text{minimize } f(x) \quad \text{s.t. } h(x) = u$$

Prop 4.2.2 of text shows that for "small" u, i.e., $u \in S$ (open sphere around 0), there is a $x^*(u)$ and $\lambda(u)$ that are the local min. and Lagrange multiplier for perturbed problem. Also $x^*(u)$ and $\lambda(u)$ are cont. diff. functions

<u>Claim</u> $\quad f(x^*(u)) = f(x^*) - \lambda^T u + o(\|u\|)$

<u>Proof</u> $\quad$ Let $p(u) = f(x^*(u))$

Then $p(0) = f(x^*(0)) = f(x^*)$

Then we need to show that

$$p(u) = p(0) - \lambda^T u + o(\|u\|)$$

i.e. we need to show that $\boxed{\nabla p(0) = -\lambda}$

$$\frac{\partial p(u)}{\partial u_j} = \frac{\partial f(x^*(u))}{\partial u_j}$$

$$= \sum_{k=1}^{n} \frac{\partial f}{\partial x_k}(x^*(u)) \cdot \frac{\partial x_k^*(u)}{\partial u_j}$$

First-order NC for perturbed problem

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

$$\Rightarrow \frac{\partial f}{\partial x_k}(x^*(u)) = -\sum_{i=1}^{m} \lambda_i(u) \frac{\partial h_i}{\partial x_k}(x^*(u))$$

Thus,

$$\frac{\partial p(u)}{\partial u_j} = -\sum_{k=1}^{n} \sum_{i=1}^{m} \lambda_i(u) \frac{\partial h_i}{\partial x_k}(x^*(u)) \frac{\partial x_k^*(u)}{\partial u_j}$$

$$= -\sum_{i=1}^{m} \lambda_i(u) \frac{\partial h_i(x^*(u))}{\partial u_j}$$

But $h_i(x^*(u)) = u_i$

$$\Rightarrow \frac{\partial h_i(x^*(u))}{\partial u_j} = \begin{cases} 1 & \text{if } i=j \\ 0 & \text{if } i \neq j \end{cases}$$

$$\Rightarrow \frac{\partial p(u)}{\partial u_j} = -\lambda_j(u)$$

$$\Rightarrow \nabla p(u) = -\lambda(u)$$

$$\nabla p(0) = -\lambda(0) = -\lambda$$