

# Lagrange Multipliers - Equality Constraints

Lec 13

## First-order necessary condition

Let  $x^*$  be a local min. of  $f(x)$  s.t.  $h(x) = 0$

If  $\nabla h_1(x^*), \nabla h_2(x^*), \dots, \nabla h_m(x^*)$  are l.i.,  
then  $\exists$  a unique  $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_m)$  s.t.

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

## Second-order necessary condition

- $g^{(k)}(x) = f(x) + \frac{\kappa}{2} \|h(x)\|^2 + \frac{\alpha}{2} \|x - x^*\|^2, k=1, 2, \dots$
- $x^{(k)} = \arg \min_{x \in \mathcal{S}} g^{(k)}(x), \quad \mathcal{S} = \{x : \|x - x^*\| \leq \varepsilon\}$
- $\lim_{k \rightarrow \infty} x^{(k)} = x^*$ , 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  $\nabla^2 g^{(k)}(x^{(k)}) \succcurlyeq 0$ , for  $k$  suff. large

$$\begin{aligned} \bullet \nabla g^{(k)}(x) &= \nabla f(x) + \kappa \sum_{i=1}^m h_i(x) \nabla h_i(x) + \alpha(x - x^*) \\ \Rightarrow \nabla^2 g^{(k)}(x) &= \nabla^2 f(x) + \kappa \sum_{i=1}^m \nabla h_i(x) \nabla h_i(x)^T \\ &\quad + \kappa \sum_{i=1}^m h_i(x) \nabla^2 h_i(x) + \alpha I \end{aligned}$$

$$\nabla^2 g^{(k)}(x) = \nabla^2 f(x) + \kappa \nabla h(x) \nabla h(x)^T$$

$$+ \kappa \sum_{i=1}^m h_i(x) \nabla^2 h_i(x) + \alpha I$$

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

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

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

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

$$\begin{aligned} \text{Then } \nabla h(x^{(k)})^T \bar{z}^{(k)} &= \nabla h(x^{(k)})^T \bar{z} - \nabla h(x^{(k)})^T \bar{z} \\ &= 0. \end{aligned}$$

And (1) implies

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

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

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

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

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

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

Taking  $\alpha \rightarrow 0$ ,

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

Similarly, we can show the following second order sufficient condition:

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

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

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

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

minimize  $f(x)$

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.

$$\begin{aligned} \nabla_x L(x^*, \lambda) &= 0 \\ h(x^*) = \nabla_\lambda L(x^*, \lambda) &= 0 \end{aligned} \quad \begin{array}{l} m+n \text{ eqns} \\ \text{in } m+n \text{ unknowns} \\ x^*, \lambda \end{array}$$

$$\mathbf{z}^\top \nabla_{xx}^2 L(x^*, \lambda) \mathbf{z} \geq 0 \quad \forall \mathbf{z} \in \mathcal{V}(x^*)$$

$$\begin{aligned} (\text{i.e. } \forall \mathbf{z} \text{ s.t.} \\ \nabla h_i(x^*)^\top \mathbf{z} = 0 \forall i) \end{aligned}$$

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  $\mathbf{z}^\top \nabla_{xx}^2 L(x^*, \lambda) \mathbf{z} > 0, \forall \mathbf{z} \neq 0, \mathbf{z} \in \mathcal{V}(x^*)$

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  $x_1 + x_2 + x_3 = 3$

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

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

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

$$\nabla h(x) = [1 \ 1 \ 1]^T \leftarrow \text{rank } 1$$

$\Rightarrow$  every  $x$  is regular

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

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

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

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

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

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

$$\begin{aligned} \nabla_{xx}^2 L(x, \lambda) &= \nabla_{xx}^2 f(x) + \lambda \nabla_{xx}^2 h(x) \\ &= I + 0 = I > 0 \end{aligned}$$

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

Is  $x^*$  also the global min of  $f(x)$  s.t.  $h(x) = 0$ ?  
 f is convex and  $H = \{x : h(x) = 0\}$  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,  $S \subseteq \mathbb{R}^n$ , then  $f$  attains its min and max on  $S$ , i.e.,

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

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

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

Application to (P1)

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

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

$\Rightarrow f$  achieves its global min on  $H$

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

## Sensitivity Analysis

Let  $x^*$ , which is regular, and  $\lambda$  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

Claim  $f(x^*(u)) = f(x^*) - \lambda^T u + O(\|u\|)$

Proof Let  $\phi(u) = f(x^*(u))$

$$\phi(0) = f(x^*(0)) = f(x^*)$$

Then we need to show that

$$\phi(u) = \phi(0) - \lambda^T u + O(\|u\|)$$

i.e. we need to show that

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

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

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,

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

$$\text{But } h_i(x^*(u)) = u_i$$

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

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

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

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