

# LECTURE 17

## LECTURE OUTLINE

- Sensitivity Issues
  - Exact penalty functions
  - Extended representations
- 

### Review of Lagrange Multipliers

- Problem:  $\min f(x)$  subject to  $x \in X$ , and  $g_j(x) \leq 0$ ,  $j = 1, \dots, r$ .
- Key issue is the existence of Lagrange multipliers for a given local min  $x^*$ .
- Existence is guaranteed if  $X$  is regular at  $x^*$  and we can choose  $\mu_0^* = 1$  in the FJ conditions.
- Pseudonormality of  $x^*$  guarantees that we can take  $\mu_0^* = 1$  in the FJ conditions.
- We derived several constraint qualifications on  $X$  and  $g_j$  that imply pseudonormality.

# PSEUDONORMALITY

A feasible vector  $x^*$  is *pseudonormal* if there are NO scalars  $\mu_1, \dots, \mu_r$ , and a sequence  $\{x^k\} \subset X$  such that:

(i)  $-\left(\sum_{j=1}^r \mu_j \nabla g_j(x^*)\right) \in N_X(x^*)$ .

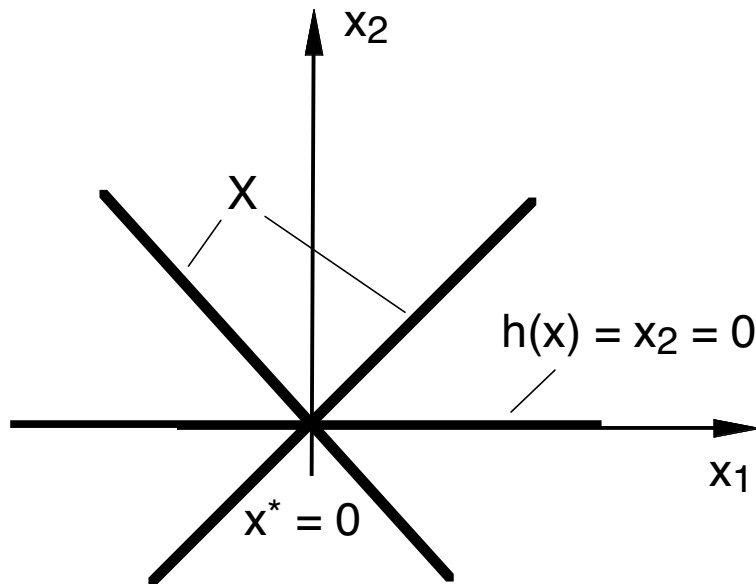
(ii)  $\mu_j \geq 0$ , for all  $j = 1, \dots, r$ , and  $\mu_j = 0$  for all  $j \notin A(x^*) = \{j \mid g_j(x^*) = 0\}$ .

(iii)  $\{x^k\}$  converges to  $x^*$  and

$$\sum_{j=1}^r \mu_j g_j(x^k) > 0, \quad \forall k.$$

- From Enhanced FJ conditions:
  - If  $x^*$  is pseudonormal, there exists an R-multiplier vector.
  - If in addition  $X$  is regular at  $x^*$ , there exists a Lagrange multiplier vector.

## EXAMPLE WHERE $X$ IS NOT REGULAR



- We have

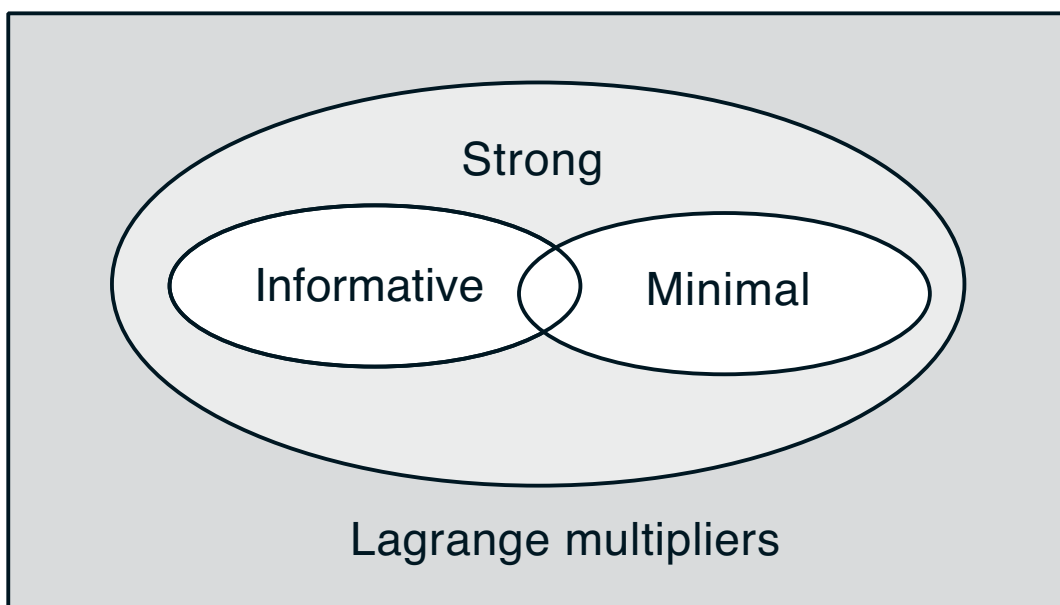
$$T_X(x^*) = X, \quad T_X(x^*)^* = \{0\}, \quad N_X(x^*) = X$$

Let  $h(x) = x_2 = 0$  be a single equality constraint. The only feasible point  $x^* = (0, 0)$  is pseudonormal (satisfies CQ2).

- There exists no Lagrange multiplier for some choices of  $f$ .
- For each  $f$ , there exists an R-multiplier, i.e., a  $\lambda^*$  such that  $-(\nabla f(x^*) + \lambda^* \nabla h(x^*)) \in N_X(x^*) \dots$  BUT for  $f$  such that there is no L-multiplier, the Lagrangian has negative slope along a tangent direction of  $X$  at  $x^*$ .

# TYPES OF LAGRANGE MULTIPLIERS

- **Informative:** Those that satisfy condition (iii) of the FJ Theorem
- **Strong:** Those that are informative if the constraints with  $\mu_j^* = 0$  are neglected
- **Minimal:** Those that have a minimum number of positive components
- **Proposition:** Assume that  $T_X(x^*)$  is convex. Then the inclusion properties illustrated in the following figure hold. Furthermore, if there exists at least one Lagrange multiplier, there exists one that is informative (the multiplier of min norm is informative - among possibly others).



# SENSITIVITY

- Informative multipliers provide a certain amount of sensitivity.
- They indicate the constraints that need to be violated [those with  $\mu_j^* > 0$  and  $g_j(x^k) > 0$ ] in order to be able to reduce the cost from the optimal value [ $f(x^k) < f(x^*)$ ].
- The L-multiplier  $\mu^*$  of minimum norm is informative, but it is also special; it provides quantitative sensitivity information.
- More precisely, let  $d^* \in T_X(x^*)$  be the direction of maximum cost improvement for a given value of norm of constraint violation (up to 1st order; see the text for precise definition). Then for  $\{x^k\} \subset X$  converging to  $x^*$  along  $d^*$ , we have

$$f(x^k) = f(x^*) - \sum_{j=1}^r \mu_j^* g_j(x^k) + o(\|x^k - x^*\|)$$

- In the case where there is a unique L-multiplier and  $X = \mathbb{R}^n$ , this reduces to the classical interpretation of L-multiplier.

# EXACT PENALTY FUNCTIONS

- Exact penalty function

$$F_c(x) = f(x) + c \left( \sum_{i=1}^m |h_i(x)| + \sum_{j=1}^r g_j^+(x) \right),$$

where  $c$  is a positive scalar, and

$$g_j^+(x) = \max\{0, g_j(x)\}.$$

- We say that the constraint set  $C$  admits an exact penalty at a feasible point  $x^*$  if for every smooth  $f$  for which  $x^*$  is a strict local minimum of  $f$  over  $C$ , there is a  $c > 0$  such that  $x^*$  is also a local minimum of  $F_c$  over  $X$ .
- Need for strictness in the definition.

**Main Result:** If  $x^* \in C$  is pseudonormal, the constraint set admits an exact penalty at  $x^*$ .

## INTERMEDIATE RESULT

- First use the (generalized) Mangasarian-Fromovitz CQ to obtain a necessary condition for a local minimum of the exact penalty function.

**Proposition:** Let  $x^*$  be a local minimum of  $F_c = f + c \sum_{j=1}^r g_j^+$  over  $X$ . Then there exist  $\mu_1^*, \dots, \mu_r^*$  such that

$$-\left( \nabla f(x^*) + c \sum_{j=1}^r \mu_j^* \nabla g_j(x^*) \right) \in N_X(x^*),$$

$$\begin{aligned} \mu_j^* &= 1 & \text{if } g_j(x^*) > 0, & & \mu_j^* &= 0 & \text{if } g_j(x^*) < 0, \\ \mu_j^* &\in [0, 1] & \text{if } g_j(x^*) = 0. & & & & \end{aligned}$$

**Proof:** Convert minimization of  $F_c(x)$  over  $X$  to minimizing  $f(x) + c \sum_{j=1}^r v_j$  subject to

$$x \in X, \quad g_j(x) \leq v_j, \quad 0 \leq v_j, \quad j = 1, \dots, r.$$

## PROOF THAT PN IMPLIES EXACT PENALTY

- Assume PN holds and that there exists a smooth  $f$  such that  $x^*$  is a strict local minimum of  $f$  over  $C$ , while  $x^*$  is not a local minimum over  $x \in X$  of  $F_k = f + k \sum_{j=1}^r g_j^+$  for all  $k = 1, 2, \dots$
- Let  $x^k$  minimize  $F_k$  over all  $x \in X$  satisfying  $\|x - x^*\| \leq \epsilon$  (where  $\epsilon$  is s.t.  $f(x^*) < f(x)$  for all  $x \in X$  with  $x \neq 0$  and  $\|x - x^*\| < \epsilon$ ). Then  $x^k \neq x^*$ ,  $x^k$  is infeasible, and

$$F_k(x^k) = f(x^k) + k \sum_{j=1}^r g_j^+(x^k) \leq f(x^*)$$

so  $g_j^+(x^k) \rightarrow 0$  and limit points of  $x^k$  are feasible.

- Can assume  $x^k \rightarrow x^*$ , so  $\|x^k - x^*\| < \epsilon$  for large  $k$ , and we have the necessary conditions

$$-\left( \frac{1}{k} \nabla f(x^k) + \sum_{j=1}^r \mu_j^k \nabla g_j(x^k) \right) \in N_X(x^k)$$

where  $\mu_j^k = 1$  if  $g_j(x^k) > 0$ ,  $\mu_j^k = 0$  if  $g_j(x^k) < 0$ , and  $\mu_j^k \in [0, 1]$  if  $g_j(x^k) = 0$ .

## PROOF CONTINUED

- We can find a subsequence  $\{\mu^k\}_{k \in \mathcal{K}}$  such that for some  $j$  we have  $\mu_j^k = 1$  and  $g_j(x^k) > 0$  for all  $k \in \mathcal{K}$ . Let  $\mu$  be a limit point of this subsequence. Then  $\mu \neq 0$ ,  $\mu \geq 0$ , and

$$-\sum_{j=1}^r \mu_j \nabla g_j(x^*) \in N_X(x^*)$$

[using the closure of the mapping  $x \mapsto N_X(x)$ ].

- Finally, for all  $k \in \mathcal{K}$ , we have  $\mu_j^k g_j(x^k) \geq 0$  for all  $j$ , so that, for all  $k \in \mathcal{K}$ ,  $\mu_j g_j(x^k) \geq 0$  for all  $j$ . Since by construction of the subsequence  $\{\mu^k\}_{k \in \mathcal{K}}$ , we have for some  $j$  and all  $k \in \mathcal{K}$ ,  $\mu_j^k = 1$  and  $g_j(x^k) > 0$ , it follows that for all  $k \in \mathcal{K}$ ,

$$\sum_{j=1}^r \mu_j g_j(x^k) > 0.$$

This contradicts the pseudonormality of  $x^*$ . **Q.E.D.**

## EXTENDED REPRESENTATION

- $X$  can often be described as

$$X = \{x \mid g_j(x) \leq 0, j = r + 1, \dots, \bar{r}\}.$$

- Then  $C$  can alternatively be described without an abstract set constraint,

$$C = \{x \mid g_j(x) \leq 0, j = 1, \dots, \bar{r}\}.$$

We call this the *extended representation* of  $C$ .

### Proposition:

- (a) If the constraint set admits Lagrange multipliers in the extended representation, it admits Lagrange multipliers in the original representation.
- (b) If the constraint set admits an exact penalty in the extended representation, it admits an exact penalty in the original representation.

## PROOF OF (A)

- By conditions for case  $X = \Re^n$  there exist  $\mu_1^*, \dots, \mu_{\bar{r}}^*$  satisfying

$$\nabla f(x^*) + \sum_{j=1}^{\bar{r}} \mu_j^* \nabla g_j(x^*) = 0,$$

$$\mu_j^* \geq 0, \quad \forall j = 0, 1, \dots, \bar{r}, \quad \mu_j^* = 0, \quad \forall j \notin \bar{A}(x^*),$$

where

$$\bar{A}(x^*) = \{j \mid g_j(x^*) = 0, j = 1, \dots, \bar{r}\}.$$

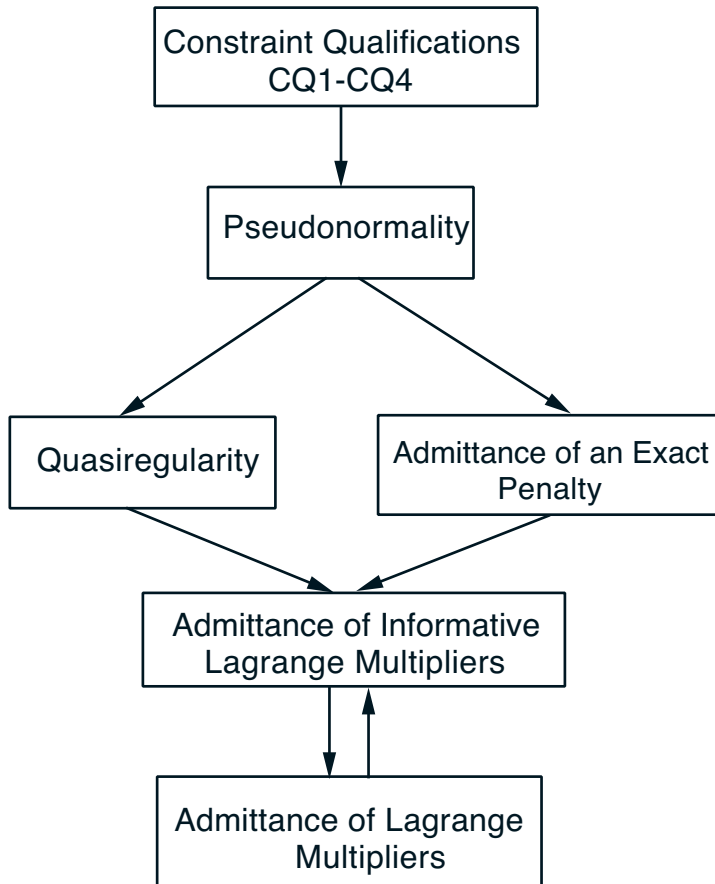
For  $y \in T_X(x^*)$ , we have  $\nabla g_j(x^*)'y \leq 0$  for all  $j = r + 1, \dots, \bar{r}$  with  $j \in \bar{A}(x^*)$ . Hence

$$\left( \nabla f(x^*) + \sum_{j=1}^r \mu_j^* \nabla g_j(x^*) \right)' y \geq 0, \quad \forall y \in T_X(x^*),$$

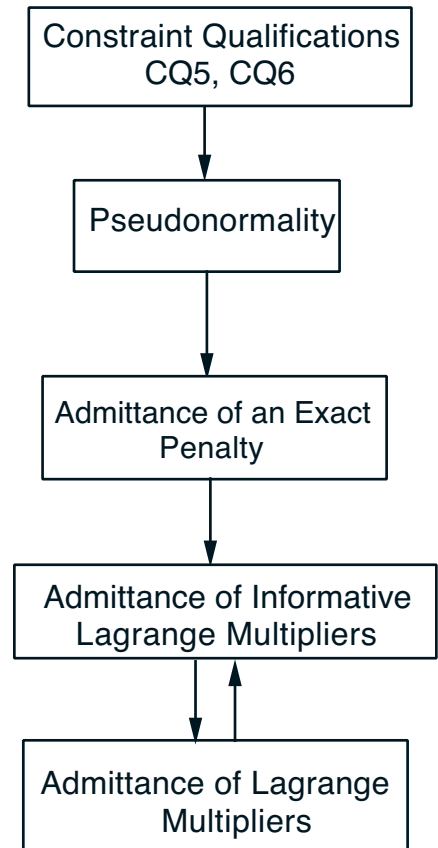
and the  $\mu_j^*$ ,  $j = 1, \dots, r$ , are Lagrange multipliers for the original representation.

# THE BIG PICTURE

$X = \mathbb{R}^n$



$X \neq \mathbb{R}^n$  and Regular



$X \neq \mathbb{R}^n$

