LECTURE 22

LAGRANGE MULTIPLIERS AND SENSITIVITY

Intuitively Lagrange multipliers measure the sensitivity of a constrained optimization problem to perturbations in the constraints.

Consider

$$\underset{x \in \mathbb{R}^n}{\text{minimize}} \quad \underbrace{c^T x}_{f(x)}$$
$$\text{subject to}$$
$$a^T x = b$$

where $a, c \in \mathbb{R}^n$ and $b \in \mathbb{R}$.

$$x_* \in \mathbb{R}^n \text{ is a minimizer}$$
$$\Longrightarrow$$

(i) $c = \pi a$ for some $\pi \in \mathbb{R}$
and
(ii) $a^T x_* = b$

$$\Longrightarrow$$

$$f(x_*) = c^T x_* = (\pi a)^T x_*$$
$$= \pi a^T x_* = \pi b$$

Now slightly perturb the constraint and denote the minimizer for the perturbed problem by $x_*(\epsilon)$.

$$\begin{aligned} &\underset{x \in \mathbb{R}^n}{\text{minimize}} \quad c^T x \\ &\text{subject to} \\ &a^T x = b + \epsilon \end{aligned}$$

$x_*(\epsilon) \in \mathbb{R}^n$ is a minimizer

$$\Longrightarrow$$

(i) $c = \pi a$ for some $\pi \in \mathbb{R}$

and

(ii) $a^T x_*(\epsilon) = b + \epsilon$

$$\Longrightarrow$$

$$\begin{aligned} f(x_*(\epsilon)) &= c^T x_*(\epsilon) \\ &= (\pi a)^T x_*(\epsilon) \\ &= \pi a^T x_*(\epsilon) \\ &= \pi (b + \epsilon) \end{aligned}$$

Consequently

$$\frac{f(x_*(\epsilon)) - f(x_*)}{\epsilon} = \frac{\pi \cancel{\epsilon}}{\cancel{\epsilon}}$$

and taking the limit as $\epsilon \to 0$

$$\left.\frac{d f(x_*(\epsilon))}{d\epsilon}\right|_{\epsilon=0} = \pi .$$

Now consider the inequality constrained problem

$$\text{(NIP)} \quad \begin{array}{l} \underset{x \in \mathbb{R}^n}{\text{minimize}} \ f(x) \\ \text{subject to} \\ \quad c_j(x) \geq 0 \quad j = 1, \ldots, m \end{array}$$

Suppose the kth constraint is slightly perturbed by $\epsilon \|\nabla c_k(x_*)\|$ where x_* is a minimizer of problem (NIP).

$$\begin{array}{l} \underset{x \in \mathbb{R}^n}{\text{minimize}} \ f(x) \\ \text{subject to} \\ \quad c_j(x) \geq 0 \quad j = 1, \ldots, m, \ j \neq k \\ \quad c_k(x) \geq \epsilon \|\nabla c_k(x_*)\| \end{array}$$

Let $x_*(\epsilon)$ be a minimizer of the perturbed (NIP). By Taylor's thm

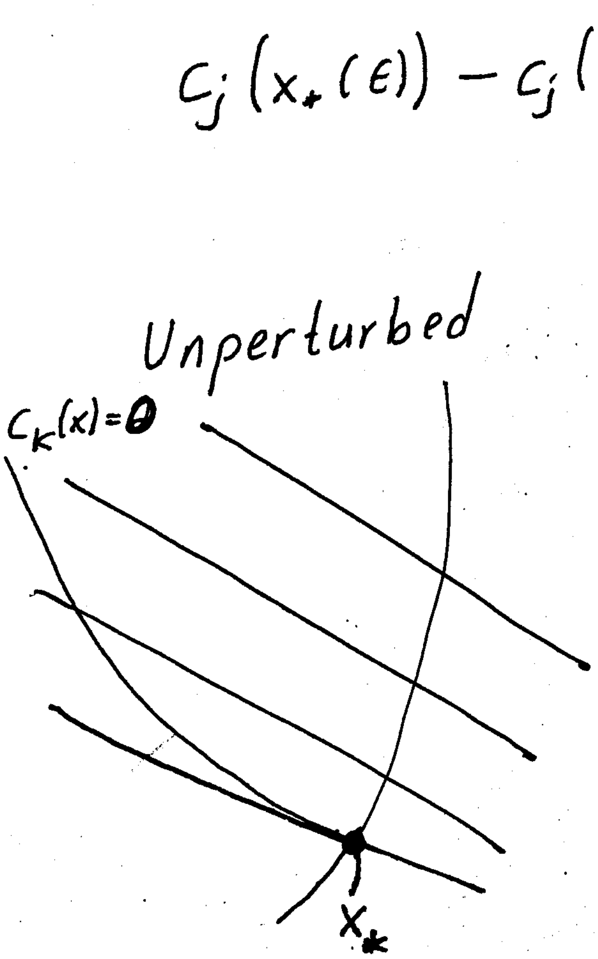
$$f(x_*(\epsilon)) - f(x_*) = \nabla f(x_*)^T (x_*(\epsilon) - x_*) + O(\|x_*(\epsilon) - x_*\|^2)$$

$$= \left(\sum_{j=1}^{m} \lambda_j \nabla c_j(x_*) \right)^T (x_*(\epsilon) - x_*) + O(\|x_*(\epsilon) - x_*\|^2)$$

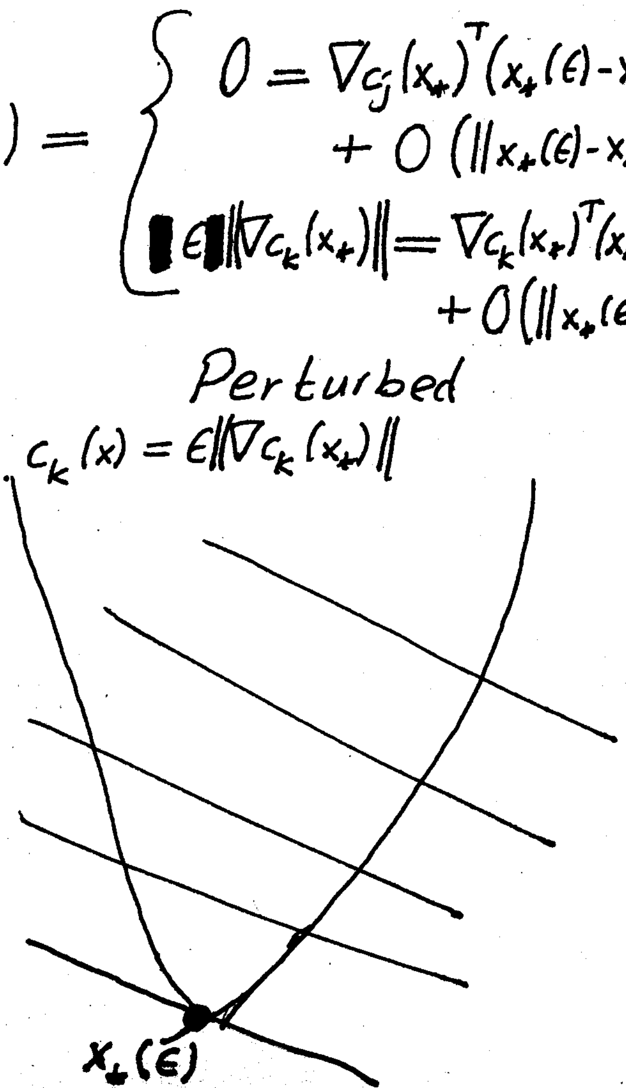
where (assuming active constraints remain same)

$$c_j(x_*(\epsilon)) - c_j(x_*) = \begin{cases} 0 = \nabla c_j(x_*)^T (x_*(\epsilon) - x_*) + O(\|x_*(\epsilon) - x_*\|^2) & j \\ \epsilon \|\nabla c_k(x_*)\| = \nabla c_k(x_*)^T (x_*(\epsilon) - x_*) + O(\|x_*(\epsilon) - x_*\|^2) & \end{cases}$$

Unperturbed



Perturbed



Consequently

$$f(x_*(\epsilon)) - f(x_*) = \lambda_k \nabla c_k(x_*)^T (x_*(\epsilon) - x_*) + O(\|x_*(\epsilon) - x_*\|$$

$$= \lambda_k \epsilon \|\nabla c_k(x_*)\| + \underbrace{O(\|x_*(\epsilon) - x_*\|}_{O(\epsilon^2)}$$

implying

$$\boxed{\frac{df(x_*(\epsilon))}{d\epsilon} = \lambda_k \|\nabla c_k(x_*)\|}$$

INTERIOR POINT METHODS FOR LINEAR PROGRAMS

$$\text{(LP)} \quad \begin{array}{ll} \underset{x \in \mathbb{R}^n}{\text{minimize}} & \underbrace{c^T x}_{f(x)} \\ \text{subject to} & \\ & Ax = b \\ & x \geq 0 \end{array}$$

where
$A \in \mathbb{R}^{m \times n}$
$c \in \mathbb{R}^n$
$b \in \mathbb{R}^m$

Associated KKT conditions

(P1) $Ax = b$

(P2) $x \geq 0$

(P3) $c = A^T \pi + s$ for some $\pi \in \mathbb{R}^m$, $s \in \mathbb{R}$

Lagrange multipliers for equality const.

Lagrange multipliers for inequality const.

(P4) $s \geq 0$

(P5) $s_i x_i = 0 \quad i = 1, \dots, n$ (equivalently $s^T x = 0$)

THM (Global Min for LP)

A point $x_* \in \mathbb{R}^n$ is a global minimizer of (LP) if and only if it satisfies KKT conditions.

PROOF

If a point does not satisfy KKT conditions, it cannot be a local minimizer (by first order necessary conditions) and consequently cannot be a global minimizer of (LP).

Now suppose $x_* \in \mathbb{R}^n$ satisfies KKT conditions. Then x_* is feasible and

$$
\begin{aligned}
f(x_*) &= c^T x_* \\
&= (A^T \pi + s)^T x_* \\
&= \pi^T A x_* + \underbrace{s^T x_*}_{0 \text{ (by complementarity)}} \\
&= \pi^T b
\end{aligned}
$$

Let $\hat{x}$ be any feasible point

$$
\begin{aligned}
f(\hat{x}) &= c^T \hat{x} \\
&= (A^T \pi + s)^T \hat{x}
\end{aligned}
$$

$$= \pi^T \underbrace{A\hat{x}}_{b \,(\text{since } \hat{x}\in\mathcal{F})} + \underbrace{\hat{x}^T}_{\geq 0} \underbrace{s}_{\geq 0}$$

$$= \pi^T b + \hat{x}^T s \geq \pi^T b = f(x_*)$$

meaning x_* is a global minimizer.

$\square$

DUAL PROBLEM

Constrained optimization problems have dual problems

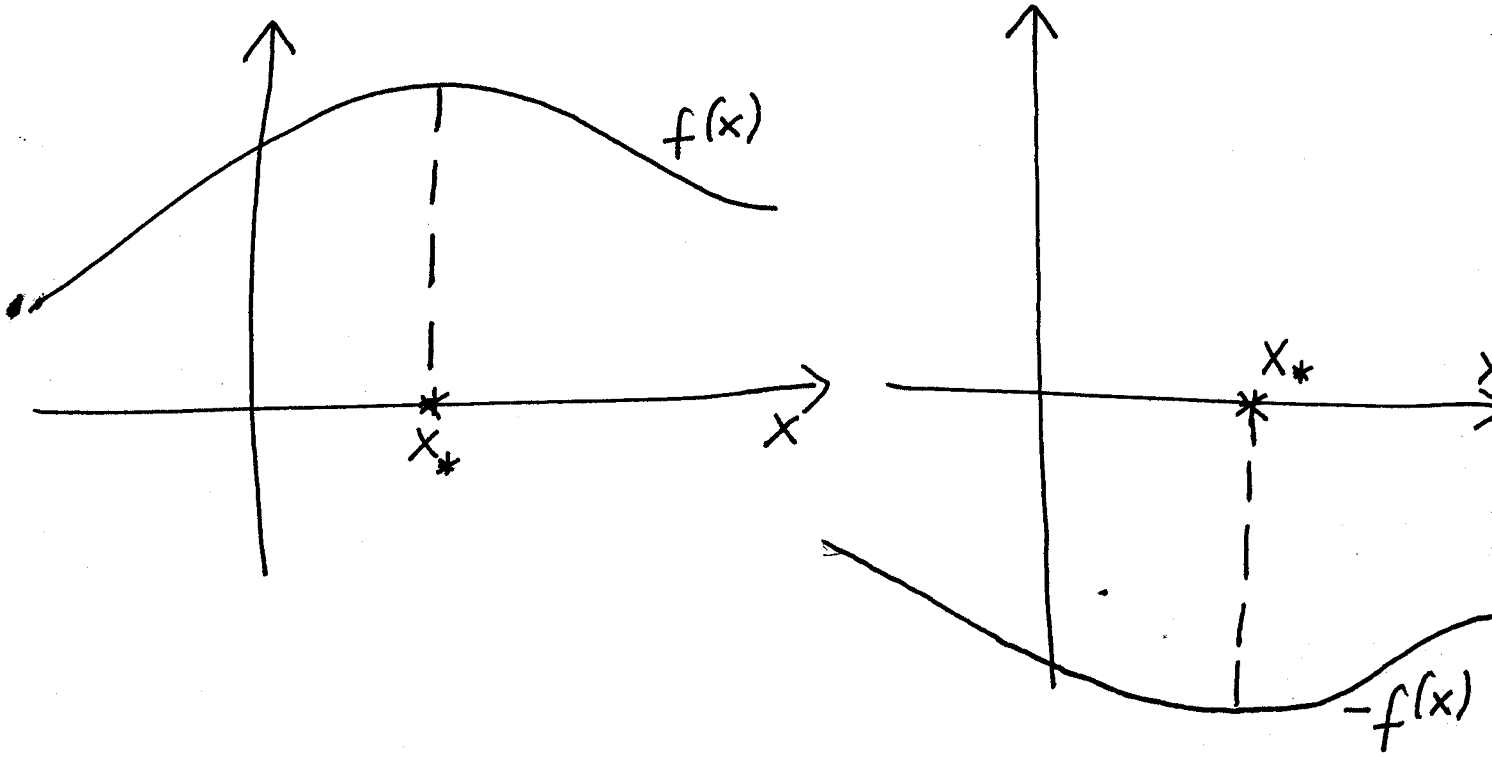
* that yield lower bounds for actual problems,
* For special problems such as (LP) the dual problem yields the same value (objective function) as the actual problem.

Dual linear program (for (LP))

$$\text{(DLP)} \quad \begin{aligned} &\underset{\pi \in \mathbb{R}^m}{\text{maximize}} \; \pi^T b \\ &\text{subject to} \\ &\quad c - A^T \pi \geq 0 \end{aligned}$$

EQUIVALENT FORMS

① Turn maximization into minimization



* $\underset{x \in \mathbb{R}^n}{\text{maximize}}\ f(x) \equiv -\underset{x \in \mathbb{R}^n}{\text{minimize}}\ -f(x)$

* local maximizers of $f(x)$ are same as the local minimizers of $-f(x)$.

Dual linear program can be posed as

$$\text{(DLP2)} \quad \underset{\pi \in \mathbb{R}^m}{\text{minimize}}\ -\pi^T b \quad \text{subject to} \quad c - A^T\pi \geq 0$$

Maximizing π values for (DLP) are same as the minimizing π values for (DLP2).

KKT conditions for (DLP2)

(1) $c - A^T\pi \geqslant 0$

(2) $-b = -Ax$ Lagrange multipliers for some $x \in \mathbb{R}^n$

(3) $x \geqslant 0$

(4) $(c - A^T\pi)^T x = 0$ $\boxed{\text{COMPLEMENTARITY}}$

equivalently letting $s = c - A^T\pi$

(D1) $s \geqslant 0$

(D2) $Ax = b$

(D3) $x \geqslant 0$

(D4) $c = A^T\pi + s$

(D5) $s^T x = 0$

REMARK

Conditions (D1-D5) are same as the KKT conditions for (LP).

② Standard LP form for dual problem

Introducing slack variables $s = c - A^T\pi$ as in for conditions (D1-D5)

$$\text{(DLP3)} \quad \begin{array}{c} \underset{\pi \in \mathbb{R}^m,\, s \in \mathbb{R}^n}{\text{minimize}} \; -\pi^T b \\ \text{subject to} \\ A^T\pi + s = c \\ s \geqslant 0 \end{array}$$

· Still not in standard LP form without any nonnegativity condition on π

For any $\pi \in \mathbb{R}^m$ let

$$\pi = \pi^+ - \pi^-$$

where $\pi^+, \pi^- \geqslant 0$.

$$\left(\text{e.g. } \underbrace{\begin{bmatrix} 3 \\ -2 \\ -1 \end{bmatrix}}_{\pi} = \underbrace{\begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix}}_{\pi^+} - \underbrace{\begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix}}_{\pi^-} \right)$$

Then (DLP3) can be expressed as

$$\text{minimize } \underbrace{\begin{bmatrix} -b & b & 0 \end{bmatrix}}_{c_d^T} \underbrace{\begin{bmatrix} \pi^+ \\ \pi^- \\ s \end{bmatrix}}_{x_d}$$

$$x_d = \begin{bmatrix} \pi^+ \\ \pi^- \\ s \end{bmatrix} \in \mathbb{R}^{2m+n}$$

(DLP4) subject to

$$\underbrace{\begin{bmatrix} A^T & -A^T & I \end{bmatrix}}_{A_d} \underbrace{\begin{bmatrix} \pi^+ \\ \pi^- \\ s \end{bmatrix}}_{x_d} = \underset{b_d}{c}$$

$$\underbrace{\begin{bmatrix} \pi^+ \\ \pi^- \\ s \end{bmatrix}}_{x_d} \geqslant 0$$

REMARK

(DLP4) is equivalent to (DLP 2) (whose minimizers are same as the maximizers of (DLP))

THM

The problem (LP) has a global minimizer x_* if and only if (DLP) has a global maximizer π_*. Furthermore the optimal objective values satisfy

$$\begin{pmatrix}\text{minimal}\\ \text{value of (LP)}\end{pmatrix} \; c^T x_* = b^T \pi_* \; \begin{pmatrix}\text{maximal}\\ \text{value of (DLP)}\end{pmatrix}$$

PROOF

Suppose (LP) has a global [redacted] minimizer, x_*
Then it must satisfy KKT conditions (P1 - P5) for some (π_*, s_*).

But then conditions (D1 - D5) are satisfied by (x_*, π_*, s_*), or equivalently KKT conditions are satisfied by $(x_*, \pi_*^+, \pi_*^-, s_*)$ for (DLP4) where $\pi_* = \pi_*^+ - \pi_*^-$ s.t. $\pi_*^+, \pi_*^- \geq$

By thm (global min for LP) $(\pi_*^+, \pi_*^-, s_*$ is a global minimizer of (DLP4) which implies π_* is a global maximizer of (DLP)

Now

$$c^T x_* = \begin{array}{l} \underset{x \in \mathbb{R}^n}{\text{minimize}} \quad c^T x \\ \text{subject to} \\ \quad Ax = b, \; x \geqslant 0 \end{array}$$

$$= (A^T \pi_* + s_*)^T x_* \quad \left(\begin{array}{l}\text{since} \\ c = A^T\pi_* + s_*\end{array}\right.$$

$$= \pi_*^T \underbrace{A x_*}_{b} + \underbrace{s_*^T x_*}_{0 \text{ (COMPLEMENTARITY)}}$$

$$= \pi_*^T b$$

$$= \begin{array}{l} \underset{\pi \in \mathbb{R}^m}{\text{maximize}} \quad b^T \pi \\ \text{subject to} \\ \quad c - A^T \pi \geqslant 0 \end{array}$$

The converse can be proven similarly $\square$