LECTURE 16

MATH 409/509

SPRING 2011

NONLINEAR OPTIMIZATION WITH EQUALITY CONSTRAINTS (PART II)

OPTIMALITY CONDITIONS FOR NEP (Continued)

So far we deduced

x_* is a local minimizer

$$\Longrightarrow$$

$\nabla f(x_*)^T p \geqslant 0$ for all $p \in T^\circ$(at x_*)

Algebraic Characterization for T°

Suppose $\overset{(\text{nonzero})}{p} \in T^\circ$ so that $p = x'(0)$ for some feasible path $x(\alpha)$. Then

for $j=1,\dots,m$

$$c_j(x(\alpha)) = 0 \text{ for small } \alpha > 0$$

$$\Longrightarrow$$

$$\left.\frac{d\,(c_j(x(\alpha)))}{d\alpha}\right|_{\alpha=0} = 0$$

$$\Longrightarrow$$

$$c_j'(x(0))\,x'(0) = \nabla c_j(x_*)^T x'(0) = 0$$

Consequently

$$\left.\begin{aligned} \nabla c_1(x_*)^T p &= 0 \\ \nabla c_2(x_*)^T p &= 0 \\ &\vdots \\ \nabla c_m(x_*)^T p &= 0 \end{aligned}\right\} \Longrightarrow J(x_*)p = 0$$

where

$$J(x) = \underbrace{c'(x)}_{\substack{m \times n \\ \text{Jacobian}}} \quad \text{with} \quad c(x) = \begin{bmatrix} c_1(x) \\ c_2(x) \\ \vdots \\ c_m(x) \end{bmatrix}$$

<u>THM</u> (Algebraic Charac for T°)

For all $x_* \in \mathbb{F}$

$$T^\circ \text{ at } x_* \subseteq \text{Null}(J(x_*))$$

<u>EXAMPLE</u>

Let $c(x) = 2x_1^2 - x_2$ and $x_* = (0,0)$

$$J(x) = \nabla c(x)^T = [4x_1 \quad -1]$$

Tangent Cone (see Lecture 15, page 8)

$$T^\circ = \left\{ \alpha \begin{bmatrix} 1 \\ 0 \end{bmatrix} : \alpha \in \mathbb{R} \right\} \text{ at } x_*$$

Null Space of the Jacobian

$$\text{Null}(J(x_*)) = \{ p \in \mathbb{R}^2 : J(x_*)p = 0 \}$$
$$= \{ p \in \mathbb{R}^2 : [0 \;\; -1]\, p = 0 \}$$
$$= \left\{ \alpha \begin{bmatrix} 1 \\ 0 \end{bmatrix} : \alpha \in \mathbb{R} \right\}$$

that is

$$(*) \quad T^0 \text{ at } x_* = \text{Null}(J(x_*))$$

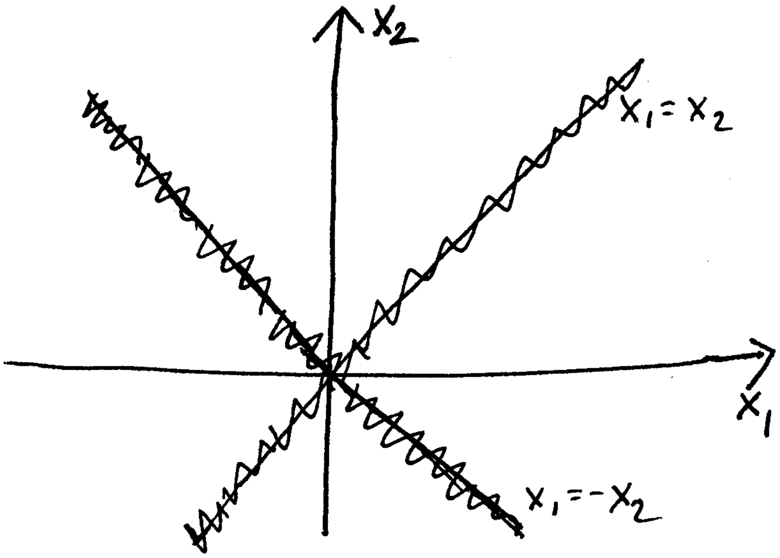
<u>REMARK</u>

Usually $(*)$ holds for (NEP).
But there are also cases when

$$T^0 \text{ at } x_* \subset \text{Null}(J(x_*))$$

<u>EXAMPLE</u>

Let $c(x) = x_1^2 - x_2^2$ and $x_* = (0,0)$



$$x_1^2 - x_2^2 = 0$$
$$\Longleftrightarrow$$
$$x_1 = x_2 \quad \text{OR} \quad x_1 = -x_2$$

Tangent Cone T^0 at x_*

$$= \left\{ \alpha \begin{bmatrix} 1 \\ 1 \end{bmatrix} : \alpha \in \mathbb{R} \right\} \cup \left\{ \alpha \begin{bmatrix} 1 \\ -1 \end{bmatrix} : \alpha \in \mathbb{R} \right\}$$

Null Space of Jacobian ($J(x) = [2x_1 \; -2x_2]$)

$$\text{Null}(J(x_*)) = \{ p \in \mathbb{R}^2 : J(x_*)p = 0 \}$$
$$= \mathbb{R}^2$$

Consequently

$$T^0 \text{ at } x_* \subset \text{Null}(J(x_*))$$

DEFN (Constraint Qualification)

We say that the constraint qualification holds at x_* if

$$\text{Null}(J(x_*)) = T^0 \text{ at } x_*$$

Assuming constraint qualification holds at x_*

OPTIMALITY CHARAC 4

x_* is a local minimizer

$$\Longrightarrow$$

$\nabla f(x_*)^T p \geq 0$ for all $p \in \text{Null}(J(x_*))$

Null Space of a matrix $A \in \mathbb{R}^{m \times n}$

$$\mathrm{Null}(A) = \{x : Ax = 0\}$$

and the Range of A^T

$$\mathrm{Range}(A^T) = \{A^T y : y \in \mathbb{R}^m\}$$

are related. $\left(= \mathrm{span}\{\bar{a}_1^T, \bar{a}_2^T, \dots, \bar{a}_m^T\} \text{ where } \bar{a}_1, \dots, \bar{a}_m \text{ denote rows of } A \right)$

EXAMPLE

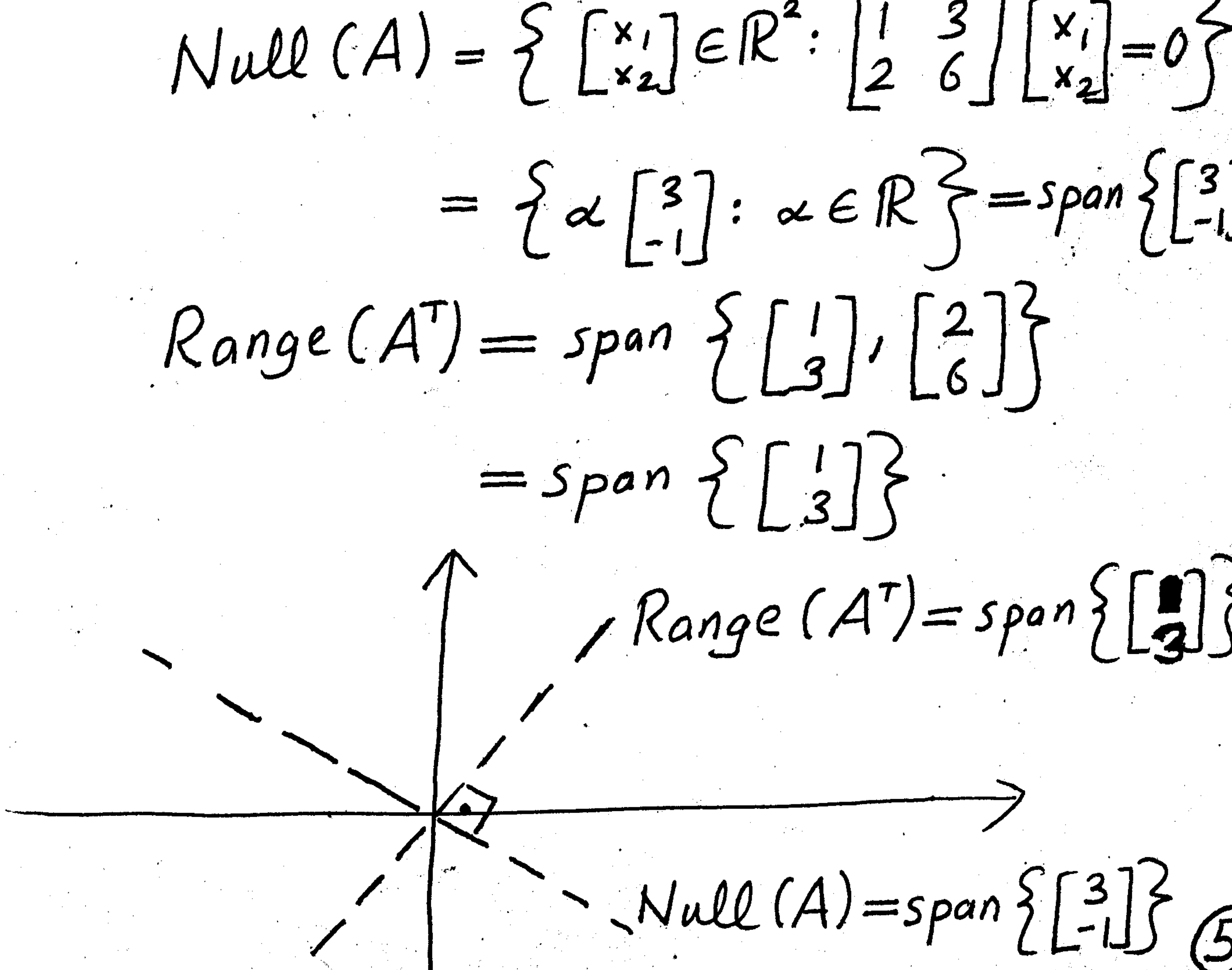
Let $A = \begin{bmatrix} 1 & 3 \\ 2 & 6 \end{bmatrix}$.

$$\mathrm{Null}(A) = \left\{ \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{R}^2 : \begin{bmatrix} 1 & 3 \\ 2 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0 \right\}$$

$$= \left\{ \alpha \begin{bmatrix} 3 \\ -1 \end{bmatrix} : \alpha \in \mathbb{R} \right\} = \mathrm{span}\left\{ \begin{bmatrix} 3 \\ -1 \end{bmatrix} \right\}$$

$$\mathrm{Range}(A^T) = \mathrm{span}\left\{ \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \begin{bmatrix} 2 \\ 6 \end{bmatrix} \right\}$$

$$= \mathrm{span}\left\{ \begin{bmatrix} 1 \\ 3 \end{bmatrix} \right\}$$



Notice that

ORTHOGONAL (1) $\mathrm{Range}(A^T) \perp \mathrm{Null}(A)$

i.e. $x^T y = 0$ for all $x \in \mathrm{Range}(A^T)$, $y \in \mathrm{Null}(A)$

COMPLEMENT (2) Any $v \in \mathbb{R}^2$ can be written as

$$v = \underbrace{\alpha_1 \begin{bmatrix} 1 \\ 3 \end{bmatrix}}_{\substack{v_R \\ \text{COMPONENT} \\ \text{ALONG RANGE}(A^T)}} + \underbrace{\alpha_2 \begin{bmatrix} 3 \\ -1 \end{bmatrix}}_{\substack{v_N \\ \text{COMPONENT} \\ \text{ALONG NULL}(A)}}$$

<u>DEFN</u> (Orthogonal Complement)

Two subspaces S_1 and S_2 of $\mathbb{R}^n$ are called orthogonal complement of each other if

(i) $x^T y = 0$ for all $x \in S_1$ and $y \in S_2$

(ii) $S_1 \oplus S_2 = \{ x + y : x \in S_1, y \in S_2 \} = \mathbb{R}^n$

<u>THM</u> (Range and Null Spaces)

Let $A \in \mathbb{R}^{m \times n}$. The subspaces

$\mathrm{Null}(A)$ and $\mathrm{Range}(A^T)$

are orthogonal complements of each other.

THM (Lagrange Multipliers)

① $\nabla f(x_*)^T p \geqslant 0$ for all $p \in Null(J(x_*))$

$$\Longleftrightarrow$$

② $\nabla f(x_*) = J(x_*)^T \lambda$ for some $\lambda \in \mathbb{R}^{r}$

PROOF

② $\Longrightarrow$ ①

Suppose ② holds. Let $p \in Null(J(x_*))$

Then

$$\nabla f(x_*)^T p = \left(J(x_*)^T \lambda\right)^T p$$

$$= \lambda^T \underbrace{J(x_*) p}$$

$$= 0$$

$\sim$② $\Longrightarrow$ $\sim$①

Suppose ② does not hold meaning

$$\nabla f(x_*) \notin Range(J(x_*)^T)$$

Then (since $Null(J(x_*)) \oplus Range(J(x_*)^T) = \mathbb{R}^{n}$)

$$\nabla f(x_*) = v_R + v_N$$

where

$$v_R \in Range(J(x_*)^T)$$

$$v_N \neq 0 \in Null(J(x_*))$$

Choose p (in ①) as $p = -v_N$

$$\nabla f(x_*)^T p = \nabla f(x_*)^T(-v_N)$$
$$= (v_R + v_N)^T(-v_N)$$
$$= \underbrace{v_R^T v_N}_{\substack{0,\text{ since} \\ \text{Range}(J(x_*)^T) \perp \text{Null}(J(x_*))}} - v_N^T v_N$$
$$= -\|v_N\|^2 < 0$$

Therefore there exists a $p = -v_N \in \text{Null}(J(x_*))$ such that $\nabla f(x_*)^T p < 0$.

$\square$

COROLLARY (First Order Optimality Conditions for NEP)

Suppose $f: \mathbb{R}^n \to \mathbb{R}$, $c_j: \mathbb{R}^n \to \mathbb{R}$ $j=1,\dots,m$ are twice continuously differentiable. If $x_* \in \mathbb{R}^n$ is a local minimizer of (NEP) where constraint qualification holds, then there exists a $\lambda \in \mathbb{R}^m$ (called the Lagrange multiplier) such that

$$\nabla f(x_*) = J(x_*)^T \lambda.$$