Week 10, Lectures 2-3 **Thursday 10/04** I

The aim of these lectures is to prove the "monster theorem", Theorem 15.3:

Let $\underline{x}_*$ be a feasible point. Then the following statements are true:

- $T_\Omega(\underline{x}_*) \subset \mathcal{F}_\Omega(\underline{x}_*)$
- If the LICQs hold at $\underline{x}_*$, then
$$T_\Omega(\underline{x}_*) = \mathcal{F}_\Omega(\underline{x}_*)$$

Notation: The active set $\mathcal{A}(\underline{x}_*)$ is the set of indices

$$\mathcal{A}(\underline{x}_*) = \mathcal{E} \cup \{ i \in \mathcal{I} \mid c_i(\underline{x}_*) = 0 \}.$$

Label the indices in $\mathcal{A}(\underline{x}_*)$ as $i_1, \dots, i_m$.

Hence:

$$c_{i_1}(\underline{x}_*) = 0, \dots, \quad c_{i_m}(\underline{x}_*) = 0,$$

and

$$\mathcal{A}(\underline{x}_*) = \{ i_1, \dots i_m \}$$

Introduce

$$A(\underline{x}_*) = \begin{pmatrix} \frac{\partial c_{i_1}}{\partial x_1} & \frac{\partial c_{i_1}}{\partial x_2} & \cdots & \frac{\partial c_{i_1}}{\partial x_n} \\ \frac{\partial c_{i_2}}{\partial x_1} & \frac{\partial c_{i_2}}{\partial x_2} & \cdots & \frac{\partial c_{i_2}}{\partial x_n} \\ \vdots & & & \\ \frac{\partial c_{i_m}}{\partial x_1} & \frac{\partial c_{i_m}}{\partial x_2} & \cdots & \frac{\partial c_{i_m}}{\partial x_n} \end{pmatrix} \begin{matrix} \uparrow \\ m \\ \downarrow \end{matrix}$$

$$\longleftarrow \quad n \quad \longrightarrow$$

Hence, $A(\underline{x}_*) \in \mathbb{R}^{m \times n}$.

From now on we drop the $\underline{x}_*$- dependence in $A(\underline{x}_*)$ and assume that $m < n$, such that the # constraints is less than the dimension of the parameter space, to avoid over-constraining the OP.

Kernel of A : We look at the kernel of A.

Hence, we look at:

$$\underbrace{\begin{matrix} \uparrow \\ m \\ \downarrow \end{matrix} \overbrace{\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2n} \\ \vdots & & \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}}^{\in \mathbb{R}^{m \times n}}}_{\substack{\longleftarrow \; n \; \longrightarrow \\ A}} \overbrace{\begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix}}^{\in \mathbb{R}^{n \times 1}} = \overbrace{\begin{pmatrix} a_{11} z_1 + a_{12} z_2 + \cdots + a_{1n} z_n \\ \vdots \\ a_{m1} z_1 + \cdots + a_{mn} z_n \end{pmatrix}}^{\in \mathbb{R}^{m \times 1}}$$

For z to be in the kernel of A, we require:

$$\sum_{i=1}^{n} a_{ji} z_i = 0 \quad \forall \; j \in \{1, \dots, m\}$$

We have:

- $\underline{z} = (z_1, \dots, z_n)^T$, a vector with n variables.
- m constraints on $\underline{z}$:

$$\sum_{i=1}^{n} a_{ji} z_i = 0 \quad \forall j \in \{1, \dots, m\}$$

- Hence, only $n-m$ free variables in $\underline{z}$, hence $\dim(\ker(A)) = n-m$.

Let the basis of $\ker(A)$ be $\{\underline{z}^{(1)}, \dots, \underline{z}^{(n-m)}\}$.

Form the matrix

$$Z = \begin{array}{c} \uparrow \\ n \\ \downarrow \end{array} \underbrace{\begin{pmatrix} | & & | \\ \underline{z}^{(1)} & \cdots & \underline{z}^{(n-m)} \\ | & & | \end{pmatrix}}_{n-m} \in \mathbb{R}^{n \times (n-m)}$$

Hence: $AZ = 0$.

We have the following lemma:

Lemma 15.1 Suppose that the LICQ hold at $\underline{x}_*$.

Then the matrix

$$\begin{pmatrix} A \\ Z^T \end{pmatrix} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \\ a_{m1} & \cdots & a_{mn} \\ z_1^{(1)} & & z_n^{(1)} \\ \vdots & & \\ z_1^{(n-m)} & \cdots & z_n^{(n-m)} \end{pmatrix} \in \mathbb{R}^{n \times n}$$

Lin ind.

has full row rank.

Proof : We have:

- First m rows are lin. independent, by the LICQs.
- Last $n-m$ rows are lin. independent since they form a basis for $\ker(A)$.

So it remains to check that the first m rows and the last $n-m$ rows are mutually lin. independent. We prove this by contradiction.

If these groups of rows are not mutually linearly independent, we can write:

$$\left(z_1^{(1)}, \dots, z_n^{(1)}\right)^T = \left(\sum_{j=1}^{m} \mu_j \cdot (a_{j1}, \dots, a_{jn})\right)^T \quad (*)$$

But $\sum_{i=1}^{n} a_{ji} z_i^{(1)} = 0$ by def^n of $z^{(1)} \in \ker(A)$.

We have:

$$z_i^{(1)} = \sum_{j=1}^{m} \mu_j a_{ji}$$

$$\overset{\text{re-index}}{=} \sum_{k=1}^{n} \mu_k a_{ki}$$

Since $z^{(1)}$ is in the kernel of A: $\longrightarrow$

$$0 = \sum_{i=1}^{n} a_{ji} z_i^{(1)}$$

$$= \sum_{i=1}^{n} a_{ji} \sum_{k=1}^{n} \mu_k a_{ki}$$

$$\Rightarrow \sum_{i=1}^{n} \sum_{k=1}^{n} a_{ji} a_{ki} \mu_k = 0$$

$$\Rightarrow \sum_{ik} (A)_{ki} (A^T)_{ij} \mu_k = 0 \quad \forall j$$

$$\Rightarrow A A^T \underline{\mu} = 0, \quad \underline{\mu} = (\mu_1, \ldots, \mu_m)^T$$

$$\Rightarrow \langle \underline{\mu}, A A^T \underline{\mu} \rangle = 0$$

$$\Rightarrow \langle A^T \underline{\mu}, A^T \underline{\mu} \rangle = 0$$

$$\Rightarrow A^T \underline{\mu} = 0$$

$$\Rightarrow \mu_1 (a_{11}, a_{12}, \ldots, a_{1n}) + \ldots + \mu_m (a_{m1}, a_{m2}, \ldots, a_{mn}) = 0$$

But A has full row rank, hence

$$\mu_1, \ldots, \mu_m = 0,$$

hence, going back up the chain of reasoning, $\underline{z}^{(1)} = 0$, which is a contradiction $\longrightarrow$

Hence $(*)$ is false, so the vectors $\underline{z}^{(1)}, \dots, \underline{z}^{(n-m)}$ are lin. independent from $(a_{11}, \dots, a_{1n})^T, \dots, (a_{m1}, \dots, a_{mn})^T$.

Hence, $\begin{pmatrix} A \\ Z^T \end{pmatrix}$ has full row rank. $\boxtimes$