Theorem 1 **(Lagrange)** *Let $f : \mathbb{R}^n \to \mathbb{R}$ and $1 \leq i \leq k$, $g_i : \mathbb{R}^n \to \mathbb{R}$ be C^1 functions. Suppose x^* is a local maximum of f on*

$$D = U \cup \{x \in \mathbb{R}^n : g_i(x) = 0, i = 1, \dots l\},$$

where U is an open set in $\mathbb{R}^n$. Suppose rank($Dg(x^*) = k$, then there exists a vector $\lambda^* = (\lambda_1^*, \dots, \lambda_l^*)$, $\lambda_i \in \mathbb{R}$ such that the following conditions are met:

$$Df(x^*) + \sum_{i=1}^{k} \lambda_i^* Dg_i(x^*) = 0 \tag{0.1}$$

Proof. WLOG, assume that the principal $k \times k$ matrix of $Dg(x^*)$ is invertible. Let $x \in \mathbb{R}^n$ be thought of as $x = \begin{bmatrix} w \\ z \end{bmatrix}$, with $w \in \mathbb{R}^k$. So if $A = Dg(x^*)$ then A_w is invertible (why ?).

By Implicit Function Theorem,

$$\exists W \text{ open in } \mathbb{R}^{n-k} : z^* \in W, \text{ and } h : W \to \mathbb{R}^k : h(z^*) = w^* \text{ and } g(h(z^*), z^*) = 0$$

$$Dh(z^*) = -A_w^{-1} A_z \tag{0.2}$$

Let $B = Df(x^*)$. We need a $\lambda^* = \begin{bmatrix} \lambda_1^* \\ \lambda_2^* \\ \dots \\ \dots \\ \lambda_k^* \end{bmatrix}$, such that

$$B_w + (\lambda^*)^T A_w = 0 \tag{0.3}$$

$$B_z + (\lambda^*)^T A_y = 0. \tag{0.4}$$

$$\tag{0.5}$$

Lets define λ^* from (0.3) and show that (0.4) is satisfied. Hence take $\lambda^* = (-B_w A_w^{-1})^T$. Now, taking this definition and using (0.2),

$$(\lambda^*)^T A_y + B_z = -B_w A_w^{-1} A_y + B_z = B_w Dh(z^*) + B_z = DF(z^*),$$

where $F : W \to \mathbb{R}$, such that $F(z) = f(h(z), z)$. But $DF(z^*) = 0$ (why ?), so we are done. □