Lecture 5(c) ENEE 664 Spring 2004

Constrained Extrema

DEF^N: Let $g_i : X \to \mathbb{R}$ $i = 1, 2, \dots, n$ be Fréchet differentiable functionals. We say that $x_0 \in X$ is a *regular point* of the set $\Omega = \{x : g_i(x) = 0,\ i = 1, 2, \dots, n\}$ if $g_i(x_0) = 0$, $i = 1, 2, \dots, n$ and the Fréchet derivatives

$$f_i = Dg_i(x_0) \quad i = 1, 2, \dots, n$$

are linearly independent linear functionals on X.

Theorem 3 Consider the functional $g : X \to \mathbb{R}$. Let $\Omega = \{x : g_i(x) = 0,\ i = 1, 2, \dots, n\}$ be a *constraint set* defined by the lineal functionals g_i, $i = 1, 2, \dots, n$. If x_0 is an extremum of g subject to the constraints *and* if x_0 is a regular point of Ω then

$$\bigcap_{i=1}^{n} \operatorname{Ker}(Dg_i(x_0)) \subset \operatorname{Ker} Dg(x_0) .$$

Proof Let $h \in \bigcap_{i=1}^{n} \operatorname{Ker} Dg_i(x_0)$

By Theorem 1 (Lecture notes 5(b)) and associated Remark there exist linearly independent vectors $y_1, y_2, \dots y_n \in X$ such that

$$M = [Dg_i(x_0)\, y_j] = \mathbb{1} \quad \text{the identity matrix}$$

Consider the equations (an implicit system)

$$g_i\left(x_0 + \epsilon h + \sum_{i=1}^{n} \varphi_i \, y_i\right) = 0$$

$$i = 1, 2, \cdots, n$$

where x_0, h and all y_i are fixed and the unknowns are ε, $\varphi_1, \varphi_2, \cdots, \varphi_n$.

Let

$$\varphi = \begin{pmatrix} \varphi_1 \\ \varphi_2 \\ \vdots \\ \varphi_n \end{pmatrix} \quad \text{and} \quad \tilde{g} = \begin{pmatrix} g_1 \\ g_2 \\ \vdots \\ g_n \end{pmatrix}$$

The above equations are of the form

$$\tilde{g} : \mathbb{R} \times \mathbb{R}^n \longrightarrow \mathbb{R}$$

$$(\varepsilon, \varphi) \longmapsto \tilde{g}(\varepsilon, \varphi) = 0$$

Since $D_2 \tilde{g}(0,0) = [D g_i (x_0) y_j]$ is nonsingular, by the <u>implicit function theorem</u>, there is a neighborhood U of $\varepsilon = 0$ and a unique vector valued function $\varphi(\varepsilon)$ on U such that

$$\tilde{g}(\varepsilon, \varphi(\varepsilon)) = 0.$$

Let $y(\varepsilon) = \sum_{i=1}^{n} \varphi_i(\varepsilon) y_i$. Then

$$0 = g_i\left(x_0 + \varepsilon h + \sum_{j=1}^{n} \varphi_j(\varepsilon) y_j\right) \qquad i = 1, 2, \cdots n$$

$$= g_i(x_0) + \varepsilon D g_i(x_0) h + D g_i(x_0) y(\varepsilon) + o(\|\varepsilon h + y(\varepsilon)\|)$$

$$= g_i(x_0) + \varepsilon Dg_i(x_0)\overset{0}{h} + Dg_i(x_0)y(\varepsilon)$$
$$+ o(\varepsilon) + o(\|y(\varepsilon)\|).$$

By hypothesis $g_i(x_0) = 0$ and $Dg_i(x_0)\overset{0}{h} = 0$.

Also $Dg_i(x_0)y(\varepsilon) = \varphi_i(\varepsilon)$.

Collecting all n equations above in vector form and taking norms, we get

$$0 = \|\varphi(\varepsilon)\| + o(\varepsilon) + o(\|y(\varepsilon)\|)$$

since $y(\varepsilon) = \sum_{i=1}^{n} y_i \varphi_i(\varepsilon)$, there exist constants c_1, c_2 such that

$$\|y(\varepsilon)\|_X c_1 \le \|\varphi(\varepsilon)\|_{\mathbb{R}^n} \le c_2 \|y(\varepsilon)\|_X$$

Hence $y(\varepsilon) = \varepsilon^2 \tilde{y}(\varepsilon)$

and $x_0 + \varepsilon h + y(\varepsilon) = x_0 + \varepsilon \overset{0}{h} + \varepsilon^2 \tilde{y}(\varepsilon)$

belongs to Ω for every $\varepsilon \in U$.

Thus x_0 is an unconstrained local extremum (i.e. $\varepsilon = 0$ is an <u>unconstrained</u> local extremum) of $g(x_0 + \varepsilon h + \varepsilon^2 \tilde{y}(\varepsilon))$.

By Theorem 1 (Lecture 4 continued)

$$\frac{d}{d\varepsilon} g(x_0 + \varepsilon \overset{0}{h} + \varepsilon^2 \tilde{y}(\varepsilon))\Big|_{\varepsilon=0} = 0$$

$$\Rightarrow \quad Dg(x_0)\overset{0}{h} = 0$$

$\blacksquare$

<u>Corollary</u> (Lagrange Multiplier Theorem)

If x_0 is an extremum of $g : X \to \mathbb{R}$ in Ω and x_0 is a regular point of Ω, then $\exists$ n scalars λ_i such that

$$D(g + \sum_{i=1}^{n} \lambda_i g_i)(x_0) = 0$$

we call λ_i the Lagrange multipliers.

<u>Proof</u>: $f_i \triangleq Dg_i(x_0)$ and $f = Dg(x_0)$

By Theorem 3

$$\bigcap_{i=1}^{n} \operatorname{Ker} f_i \subset \operatorname{Ker} f$$

By Theorem 2

$$f = \sum_{i=1}^{n} \alpha_i f_i$$

$$\Rightarrow \quad Dg(x_0) = \sum_{i=1}^{n} \alpha_i \, Dg_i(x_0)$$

$$\Rightarrow \quad D\left(g(x_0) - \sum_{i=1}^{n} \alpha_i g_i\right)(x_0) = 0$$

Set $\lambda_i = -\alpha_i$ $\quad i = 1, 2, \dots, n$. ☒

<u>Remark</u> From theorem 2 we know that $\alpha_i = f(x_i)$ where $x_1, x_2, \dots, x_n$ are linearly independent vectors satisfying $f_i(x_j) = \delta_j^i$ (i.e. dual basis). So the Lagrange multipliers can be computed from the dual basis ☒