Constrained Optimization

Note Title 27-06-2011

Implicit Function Theorem:

$f : \underset{\mathbb{R}^{n+k}}{\underset{\cap}{S}} \longrightarrow \mathbb{R}^n$

$y = (y_1 \cdots y_k)$
$x = (x_1 \cdots x_n)$

1) $f = (f_1, \cdots, f_n)$
2) $S \to$ open
3) f has continuous partial derivatives on S

4) Let $(x_0; y_0) \in S$ s.t. $f(x_0; y_0) = 0$

and $\det \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & & \\ \frac{\partial f_n}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_n} \end{bmatrix}_{n \times n} \neq 0$

Then, $\exists$ an open set $Y_0 \subset \mathbb{R}^k$ s.t. $y_0 \in Y_0$

and one and only one $g : Y_0 \to \mathbb{R}^n$
s.t.

1) g has continuous partial derivatives on Y_0
2) $g(y_0) = x_0$
3) $f(g(y), y) = 0 \quad \forall\, y \in Y_0$

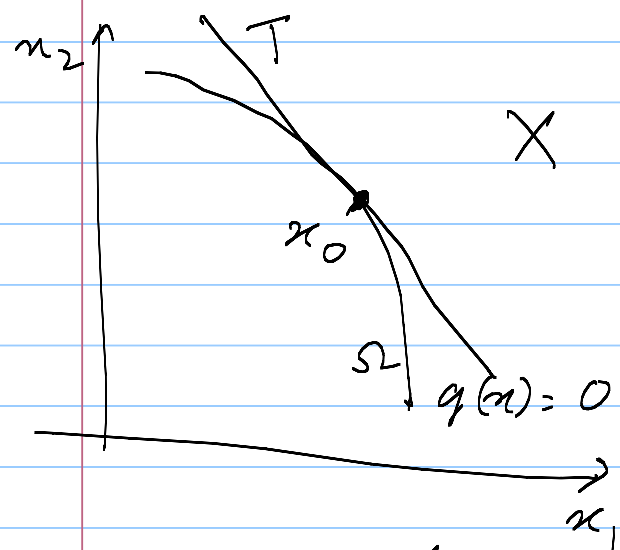
NCS

Problem: $f : X \to \mathbb{R}$; $g_1(x) : X \to \mathbb{R}$
$\vdots$
$g_n(x) : X \to \mathbb{R}$

$$\min_{x \in X} f(x)$$

s.t. $g_1(x) = 0$
$g_2(x) = 0$
$\vdots$
$g_n(x) = 0$

Assumption: f and g are Frechet diff on X.



$\Omega = \{ x : g_1(x) = \cdots g_n(x) = 0 \}$

If f has an extremum along Ω at x_0 it also has an extremum along T at x_0 (FOR small displacement)

→ We need to express T as a function of derivatives of $g(x)$

Def$\stackrel{n}{=}$: A pt. x_0 satisfying $g_1(x_0) = \cdots g_n(x_0) = 0$ is said to be a regular point of these constraints if the n linear functionals $g_1'(x_0), g_2'(x_0) \cdots g_n'(x_0)$ are linearly independent (Frechet derivative)

[Recall $g_1'(x_0) = \delta g_1(x_0; h) \in B(X, \mathbb{R})$]

Thm: If x_0 is an extremum of the functional f s.t. the constraints $g_i(x) = 0$, $i = 1, 2, \ldots, n$ and if x_0 is a regular point of these constraints, then

$$\delta f(x_0; h) = 0$$

for all h satisfying $\delta g_i(x_0; h) = 0$
$i = 1, 2, \cdots, n$

Proof: Choose some $h \in X$ satisfying
$$\delta g_i(x_0; h) = 0 \quad i = 1, 2, \cdots, n$$
$(h \neq 0)$

Let $y_1, y_2, \ldots, y_n \in X$ be n lin. ind vectors chosen s.t.

$$M = \begin{bmatrix} \delta g_1(x_0; y_1) & \delta g_1(x_0; y_2) & \cdots & \delta g_1(x_0; y_n) \\ \vdots & & & \\ \delta g_n(x_0; y_1) & \delta g_n(x_0; y_2) & \cdots & \delta g_n(x_0; y_n) \end{bmatrix}$$

is non-singular

Such $(y_1, \ldots y_n)$ exists since x_0 is regular.

$$\left[\alpha_1 \delta g_1(x_0; h) + \alpha_2 \delta g_2(x_0; h) + \cdots + \alpha_n \delta g_n(x_0; h) = 0 \right.$$
$$\text{for all } h \in X \quad \Leftrightarrow \alpha_i = 0 \ \forall i$$

At y_1, $\underbrace{\alpha_1 \delta g_1(x_0; y_1) + \alpha_2 \delta g_2(x_0; y_1) + \cdots + \alpha_n \delta g_n(x_0; y_1)}_{= 0}$

$\vdots$

At y_n, $\alpha_1 \delta g_1(x_0; y_n) + \cdots + \alpha_n \delta g_n(x_0; y_n) = 0$

$$\begin{bmatrix} \delta g_1(x_0; y_1) & - & \delta g_n(x_0; y_1) \\ \vdots & & \\ \delta g_1(x_0; y_n) & --- & \delta g_n(x_0, y_n) \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} = 0$$

$\Leftrightarrow \alpha_i = 0$

Now let $\varepsilon, \phi_1, \phi_2, \ldots, \phi_n$ be real scalars and $h \neq \theta \in X$.

$$F_1(\varepsilon, \phi_i) = g_1(x_0 + \varepsilon h + \phi_1 y_1 + \cdots + \phi_n y_n)$$
$$F_2 \quad = g_2(x_0 + \varepsilon h + \phi_1 y_1 + \cdots + \phi_n y_n)$$
$$F_n(\varepsilon, \phi_n) = g_n(x_0 + \varepsilon h + \phi_1 y_1 + \cdots + \phi_n y_n)$$

$F_i(0,0) = 0$

$\dfrac{\partial F_i}{\partial \phi_j}$ and $\dfrac{\partial F_i}{\partial \varepsilon}$ exists and are continuous fn. of (ε, ϕ_i) in some nbd of $(0,0)$

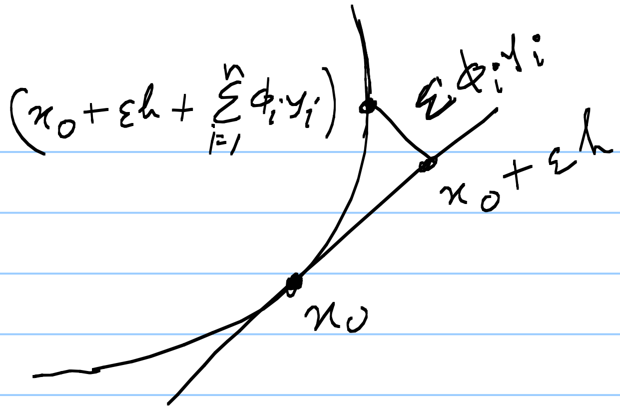
$\left[\dfrac{\partial F_i}{\partial \phi_j}\right]$ has rank n. for (ε, ϕ) in some nbd $(0,0)$

$$\left.\frac{\partial F_i}{\partial \varepsilon}\right|_{\substack{\varepsilon = 0 \\ \phi_i = 0}} = 0$$

$\lim\limits_{\varepsilon \to 0} \dfrac{g_1(x_0 + \varepsilon h) - g_1(x_0)}{\varepsilon}$

$= \delta g_1(x_0; h) = 0$

by hypothesis



$$\begin{bmatrix} \delta g_1(x_0,y_1) & \cdots & \delta g_1(x_0,y_n) \\ & & \\ \delta g_n(x_0,y_1) & \cdots & \end{bmatrix}$$

Consider the Jacobian //

$$\det\begin{bmatrix} \frac{\partial g_1}{\partial \phi_1} & \frac{\partial g_1}{\partial \phi_2} & \cdots & \frac{\partial g_1}{\partial \phi_n} \\ & & & \\ \frac{\partial g_n}{\partial \phi_1} & \cdot & \cdot & \frac{\partial g_n}{\partial \phi_n} \end{bmatrix}_{\varepsilon=0,\ \phi=0} = \det M \neq 0$$

Hence by implicit function thm:

$\exists$ $\underline{n \text{ functions}}$

$$\phi_i = \phi_i(\varepsilon)$$

defined in some nbd of $\varepsilon = 0$.

satisfying, for each i,

1> $0 = g_i\left(x_0 + \varepsilon h + \sum_{j=1}^{n} \phi_j(\varepsilon) y_j\right)$ — $(*)$

2> Also, $\phi_i(0) = 0$ and
Each $\phi_i(\varepsilon)$ is continuously diff.
in a nbd of $\varepsilon = 0$.

3> Moreover, since $\frac{\partial F_i}{\partial \varepsilon}(0,0) = 0 \quad \forall i = 1, \ldots, p$

$$\Rightarrow \quad \phi_i'(0) = 0 \quad \forall i = 1, \ldots, p$$

[Proof: $\phi_i(\varepsilon) = \varepsilon\,\phi_i'(\theta\varepsilon)$ $\quad 0<\theta<1$, ε small enough]

$$\lim_{\varepsilon\to 0}\frac{\phi_i(\varepsilon)}{\varepsilon} = \phi_i'(0) = 0$$

Now by (**),

$\left(x_0 + \varepsilon h + \sum_{j=1}^{n}\phi_j(\varepsilon)y_j\right) \in$ the constraint set

$$f\left(x_0+\varepsilon h + \sum_{j=1}^{n}\phi_j(\varepsilon)y_j\right) - f(x_0)$$

$$= \varepsilon\left[\sum_{j=1}^{n}\delta f(x_0;y_j)\,\phi_j'(0) + \delta f(x_0;h) + o(\varepsilon)\right]$$

$$= \varepsilon\left[\delta f(x_0;h) + o(\varepsilon)\right]$$

If $\delta f(x_0;h) = -a^2 \neq 0$, $\varepsilon>0$ produce contra.

$\delta f(x_0;h) = +a^2\neq 0$, $\varepsilon<0$ "

Hence $\delta f(x_0;h)=0$.

Lemma: Let $f_0, f_1, \dots f_n$ be linearly ind. linear functional on X (NLS) s.t

For any x satisfying $f_i(x)=0$, $(i=1,\dots,n)$
$f_0(x)=0$.

Then $\exists$ $\lambda_1,\dots\lambda_n$ in $\mathbb{R}$ s.t.

$$f_0 + \lambda_1 f_1 + \cdots + \lambda_n f_n = 0$$

Proof: Consider $S = \{ (f_1(x), f_2(x), \ldots, f_n(x)) : x \in X \}$

S is a subspace of E^n

Thm: If x_0 is an extremum of the functional f s.t. the constraints

$$g_i(x) = 0 \qquad i = 1, 2, \ldots, n$$

and x_0 is a regular pt. of these constraints, Then $\exists\ \lambda_1, \ldots \lambda_n$ s.t.

$$f(x) + \sum_{i=1}^{n} \lambda_i g_i(x)$$

is stationary at x_0.

Proof: By last thm $\delta f(x_0; h) = 0$ for every h satisfying $\delta g_i(x_0; h) = 0$ $(i = 1, \ldots, n)$