Multivariable Minimization

$K_1 \subset \mathbb{R}^m$, $K_2 \subset \mathbb{R}^n$ compact sets.

$f : K_1 \times K_2 \longrightarrow \overline{\mathbb{R}}$

$f(x,\cdot)$ is l.s.c on K_2 for each $x \in K_1$,

Then $F_1(x) = \inf\limits_{y \in K_2} f(x,y)$ exists

and inf is attained

If $\hat{x} \in K_1$ minimizes F_1 on K_1 then one

has $F_1(\hat{x}) = \inf\limits_{(x,y) \in K_1 \times K_2} f(x,y)$

Multivariate differentiation

- direction in $\mathbb{R}^n$ is a unit vector of $\mathbb{R}^n$
- $S_1 = \{ x \in \mathbb{R}^n \,/\, \|x\|_2 = 1 \}$ - the set of all directions.

- $U \subset \mathbb{R}^n$, open, $x_0 \in U$, $f : U \to \mathbb{R}$ continuous

f has a derivative at $x_0 \in U$ in the direction h, if $\exists\, d \in \mathbb{R}$ such that.

$$\lim_{t \to 0_+} \frac{f(x_0 + ht) - f(x_0)}{t} = d$$

When $h = e_j$ – j-th canonical base vector, the derivative is the j-th partial derivative of f at x_0, $D_j f(x_0)$.

- f is Gateaux (G) – differentiable at x_0 if $\exists\, v \in \mathbb{R}^n$ such that

$$\lim_{t \to 0_+} \frac{f(x_0 + th) - f(x_0)}{t} = v \cdot h \quad \forall h \in S_1$$

when this holds, $\nabla f(x_0) = v$ is called the gradient of f at x_0 and

$$\nabla f(x_0) = (D_1 f(x_0), \ldots, D_n f(x_0))$$

- $\hat{x} \in U$ is a critical point of f if f is G-differentiable at $\hat{x}$ and

$$\nabla f(\hat{x}) = 0 \in \mathbb{R}^n$$

$\varphi(t) = f(x_0 + th) \quad : \quad (-\delta_1, \delta_2) \longrightarrow \mathbb{R}$

If f is G-differentiable at x_0 then φ is differentiable at zero $\forall h \in S_1$ and

$$\varphi'(0) = \nabla f(x_0) \cdot h$$

- $l(x) = a_0 + \langle a, x \rangle$ – affine function

$\nabla l(x) = a$

$L_c l = \{x,\ l(x) = c\}$ – hyperplanes of $\mathbb{R}^n$

- f G-differentiable at x_0 the tangent hyperplane to the graph of f at x_0 is the graph of

$$z = f(x_0) + \langle \nabla f(x_0), x - x_0 \rangle$$

- $F : U \longrightarrow \mathbb{R}^m$ continuous

$$F(x) = \begin{pmatrix} F_1(x_1, x_2, \dots, x_n) \\ \vdots \\ F_m(x_1, x_2, \dots, x_n) \end{pmatrix}$$

F is G-differentiable at x_0 if

$\exists$ $m \times n$ matrix J such that

$$\lim_{t \to 0_+} \left\| \frac{F(x_0 + th) - F(x_0)}{t} - Jh \right\| = 0 \qquad \forall h \in S_1$$

When this holds we write

$$DF(x_0) = J$$ — Jacobian of F at x_0.

$$DF(x_0) = \begin{pmatrix} D_1 F_1(x_0) & D_2 F_1(x_0) & \cdots & D_n F_1(x_0) \\ \vdots & & & \\ D_1 F_m(x_0) & D_2 F_1(x_0) & \cdots & D_n F_m(x_0) \end{pmatrix}$$

If $F(x) = \nabla f(x) = (D_1 f(x), \ldots, D_n f(x))^T$

$$DF(x) = D^2 f(x) = \begin{pmatrix} D_{11} f(x) & D_{21} f(x) & \cdots & D_{n1} f(x) \\ \vdots & & & \\ D_{1n} f(x) & D_{2n} f(x) & \cdots & D_{nn} f(x) \end{pmatrix}$$

$\hookrightarrow$ Hessian of f

$$D_{jk} f(x) = D_j (D_k f)(x)$$

Thus

$$\lim_{t \to 0_+} \left\| \frac{\nabla f(x_0+th) - \nabla f(x_0)}{t} - D^2 f(x_0)\cdot h \right\| = 0 \quad \forall h \in S_1$$

$$\Downarrow$$

$$\left(\frac{\nabla f(x_0+th) - \nabla f(x_0)}{t}\right)\cdot h \xrightarrow[t \to 0]{} \langle D^2 f(x_0) h, h\rangle$$

For.

$\varphi(t) = f(x_0 + th)$

$$\varphi''(0) = \langle D^2 f(x_0) h, h\rangle$$

Multivariate Minimization

First variation of f at x_0 in the direction h is

$$\delta f(x_0, h) = \lim_{t \to 0_+} \frac{f(x_0+th - f(x_0)}{t}$$

If f is G-differentiable at x_0

$$\delta f(x_0, h) = \langle \nabla f(x_0), h\rangle$$

First Order Necessary optimality Conditions

$f: U \longrightarrow R$ continuous and $\hat{x}$ is a local minimizer of f on U,

If the first variation of f at $\hat{x}$ exists

then $\delta f(\hat{x}, h) \geq 0. \quad \forall h \in S_1$.

If f is G-differentiable at $\hat{x}$ then

$\nabla f(\hat{x}) = 0$

Proof $\hat{x}$ local minimizer

$f(\hat{x} + th) \geq f(\hat{x}) \quad \forall h \in S_1$ and t small enough

This implies $\delta f(\hat{x}, h) \geq 0 \quad \forall h \in S_1$

If f is G-differentiable at $\hat{x}$ we have $\delta f(\hat{x}, h) = \nabla f(\hat{x}) \cdot h \geq 0 \quad \forall h \in S_1$

Thake $h = \pm e_j$ when e_j are vectors of the canonical basis, and get $\partial_j f = 0 \quad \forall j$

- A critical point $\hat{x}$ of f is a degenerate critical point provided either

i) $D^2 f(\hat{x})$ ~~does~~ not exists. or

ii) $D^2 f(\hat{x})$ is a singular matrix

If $D^2 f(\hat{x})$ exists and is non-singular $\hat{x}$ is said to be a non-degenerate critical point of f

2-nd order Necessary Optimality Condition

- Let $\hat{x}$ be a local minimizer of f on U with f being C^1 in an open neighborhood of $\hat{x}$ in U. If $D^2 f(\hat{x})$ exists then

$$\langle D^2 f(\hat{x})h, h\rangle \geq 0 \quad \text{for all } h \in \mathbb{R}^n.$$

<u>Proof</u> Consider $\varphi(t) = f(\hat{x} + th)$ for some $h \in S_1$

Sufficient optimality condition

$f : U \to \mathbb{R}$.
$f \in C^1$ in a neighborhood of a critical point $\hat{x}$. Also suppose $D^2 f(\hat{x})$ exists and that there is a $c_1 > 0$ such that

$$\langle D^2 f(\hat{x}) h, h \rangle \geq c_1 \|h\|^2 \quad \forall h \in \mathbb{R}^n$$

Then $\hat{x}$ is an isolated, strict local minimizer of f on U.

<u>Proof</u> Let $\varphi(t) = f(x_0 + th)$

$$\varphi''(0) = \langle D^2 f(\hat{x}) h, h \rangle \geq c_1 > 0$$
$$\forall h \in S_1$$

If $\hat{x}$ is a critical point of f the 0 is a strict isolated local minimum of φ and this implies the conclusion.

Convex Minimization

Assumptions for this section

$f: C \to \mathbb{R}$, C open, non-empty and convex set in $\mathbb{R}^n$

f G-differentiable in C.

Then $\nabla f : C \to \mathbb{R}^n$ is defined.

- f is convex on C iff

(1) $f(y) \geq f(x) + \langle \nabla f(x), y - x \rangle \quad \forall x, y \in C$

Proof. f convex, $x, y \in C$ and

let $z = (1-t)x + ty$, $\quad 0 < t < 1$

$$f(z) \leq (1-t) f(x) + t f(y)$$

$$\Downarrow$$

$$\frac{f(z) - f(x)}{t} \leq f(y) - f(x)$$

$$\lim_{t \to 0^+} \frac{f(z) - f(x)}{t} = \langle \nabla f(x), y - x \rangle \leq f(y) - f(x)$$

If (1) holds. $x, y \in C$, $x \neq y$,

Let $z = (1-t)x + ty$, $t \in (0,1)$

$f(y) \geq f(z) + \langle \nabla f(z), y - z \rangle$ $/ \cdot t$

$f(x) \geq f(z) + \langle \nabla f(z), x - z \rangle$. $/ (1-t)$ $\oplus$

$(1-t) f(x) + t f(y) \geq f(z)$.

- $z = l(x) = f(x_0) + \langle a, x - x_0 \rangle$ is a support hyperplane for the graph of f at $x_0 \in C$. If f is convex on C then $f(y) \geq l(y)$ $\forall y \in C$.

Proof.

epi f is a convex set that lies above the graph of f.

- f is convex on C iff

(2) $< \nabla f(y) - \nabla f(x), y-x > \geq 0 \quad \forall x, y \in C$

It is strictly convex on C if strict inequality holds when $y \neq x$

<u>Proof</u> let $x, y \in C$, $x \neq y$,

$$\varphi(t) = f((1-t)x + ty).$$

$\varphi(t)$ is convex and differentiable function of t and $\varphi'(t) = < \nabla f((1-t)x+ty), y-x >$

Then

$$\varphi'(1) \geq \varphi'(0)$$ and this imply

the statement. (2)

If (2) holds. then $\varphi'(t) \geq \varphi'(0) \quad \forall t > 0$

$$\varphi(t) = \varphi(0) + \int_0^t \varphi'(s) ds \geq \varphi(0) + \varphi'(0) t$$

$$f((1-t)x + ty) \geq f(x) + t < \nabla f(x), y-x >$$

If (2) holds then $\nabla f(x)$ is called monotone mapping of C into $\mathbb{R}^n$

- C – non-empty compact convex set in $\mathbb{R}^n$. $f : C \longrightarrow \mathbb{R}$ l.s.c and quasi-convex.

i). The set of all minimizers of f on C is a non-empty closed subset of C

ii) If f is strictly convex on C then this set consists of exactly one point

- $f : \mathbb{R}^n \longrightarrow \overline{\mathbb{R}}$ proper, l.s.c., quasi-convex, coercive. Then

$$\alpha(f) = \inf_{x \in \mathbb{R}^n} f(x)$$ is finite and

the set of minimizers of f on $\mathbb{R}^n$ is non-empty, bounded, closed and convex

- C non-empty open convex set in $\mathbb{R}^n$ and $f : C \longrightarrow \mathbb{R}$ is convex and G-differentiable on C. A vector $\hat{x} \in C$ minimises f on C iff $\hat{x}$ is a critical point of f.

Proof

$\hat{X}$ critical point of f, then (1) implies $f(y) \geq f(x) \quad \forall y \in C$.

If $\hat{X}$ minimises f on C then

$$f(\hat{x} + td) \geq f(\hat{x}) \quad \forall t > 0, d \in S,$$

$$\Downarrow$$

$$\langle \nabla f(\hat{x}), d \rangle \geq 0 \quad \forall d \in S$$

$$\Downarrow$$

critical point.