Chapter 7. Review of Differential Calculus

§7.1 First and Second derivatives of a function.

<u>Vector spaces</u>, <u>norms</u>, <u>Linear maps</u>

<u>Defn.</u> Let $f: V \to W$ be a linear map. We say that f is <u>bounded</u> or <u>continuous</u> if

$$\|f\|_{\mathcal{L}(V,W)} = \sup_{\substack{v \neq 0 \\ v \in V}} \frac{\|f(v)\|_W}{\|v\|_V} < \infty$$

operator norm

$\mathcal{L}(V, W)$ is the vector space of all bounded linear maps from V to W.

<u>Note</u> If V is finite dimensional, $f: V \to W$, a linear map, then f is continuous.

<u>Ex.</u> $V = R^n$, $W = R^m$, $\mathcal{L}(V, W)$ is the vector space of all $m \times n$ matrices.

<u>Bilinear maps</u> $B: V \times W \to Y$ is called bilinear if

$$B(\alpha_1 v_1 + \alpha_2 v_2, w) = \alpha_1 B(v_1, w) + \alpha_2 B(v_2, w)$$
$$B(v, \alpha_1 w_1 + \alpha_2 w_2) = \alpha_1 B(v, w_1) + \alpha_2 B(v, w_2)$$

$$\forall \alpha_1, \alpha_2 \in R, \quad \forall v_1, v_2 \in V, \quad \forall w_1, w_2 \in W$$

<u>Ex.</u> Let A be an $m \times n$ matrix

$$V = R^m, \quad W = R^n, \quad Y = R$$

let $\quad B(v,w) = v^T A w \qquad v \in R^m, \ w \in R^n$

<u>Defn.</u> A bilinear map from $B : V \times W \to Y$ is said to be bounded (continuous) if

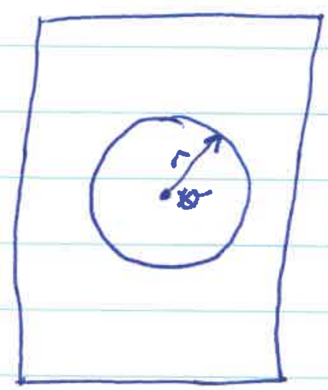
$$\|B\|_{\mathcal{L}(V \times W; Y)} \equiv \sup_{\substack{0 \neq v \in V \\ 0 \neq w \in W}} \frac{\|B(v,w)\|_Y}{\|v\|_V \, \|w\|_W} < \infty$$

<u>Open sets, open balls</u>.

let V be a vector space equipped with a norm $\|\cdot\|$.
let $v \in V$ and $r > 0$

$$B_r(v) \equiv \{ x \in V \mid \|x - v\| < r \}$$

is the <u>open</u> "<u>ball</u>" centered at v with radius r.



<u>Ex.</u> $V = R^n$, $\quad \|v\|_2 \equiv \left(\sum_{i=1}^{n} v_i^2 \right)^{1/2}$ <u>Euclidean</u> norm

<u>Ex.</u> $V = R^n$, $\quad \|v\|_\infty \equiv \underset{1 \le i \le n}{\cancel{\sup}\ \max} |v_i|$ max, sup or infinity norm
uniform

then open balls are cubes !!

<u>Defn.</u> A subset Ω of a normed space $V, \|\cdot\|$ is called <u>open</u> if for every $v \in \Omega$, we can find ~~and~~ some open ball B such that $v \in B \subseteq \Omega$.

Defn. Let $\Omega \subseteq V$ be open and let $f: \Omega \to W$ be a map (not necessarily linear). We say that f is <u>differentiable</u> at some point $a \in \Omega$ if there exists a bdd. linear map $L: V \to W$ such that

$$\lim_{\|h\|_V \to 0} \frac{1}{\|h\|_V} \| f(a+h) - f(a) - Lh \|_W = 0$$

Note $h \to 0 \iff \|h\| \to 0$.

Equivalently, $f(a+h) = f(a) + Lh + \|h\| \varepsilon(h)$ for some function $\varepsilon(h)$ satisfying $\varepsilon(h) \to 0$ as $h \to 0$

This is called differentiability in the sense of Fréchet.

Ex. $V = \mathbb{R}^2$, $W = \mathbb{R}$, $f(v) = 3v_1^2 + 2v_2^2$

In this case $\Omega = \mathbb{R}^2$: f is defined everywhere.

$$\begin{aligned} f(a+h) &= 3(a_1+h_1)^2 + 2(a_2+h_2)^2 \\ -f(a) &= -(3a_1^2 + 2a_2^2) \end{aligned}$$

$$= 6a_1h_1 + 3h_1^2 + 4a_2h_2 + 2h_2^2$$

$$= \underbrace{(6a_1, 4a_2)\cdot(h_1, h_2)}_{f'(a)h} + 3h_1^2 + 2h_2^2$$

$f'(a)$ acting on h.

i.e. $f'(a) = (6a_1, 4a_2)$

$$3h_1^2 + 2h_2^2 \leq 3(h_1^2 + h_2^2) = 3\|h\|_2 \underbrace{\|h\|_2}_{\varepsilon(h)}$$

Ex. $V = W = R^{n\times n}$ vector space of all $n\times n$ real matrices

Ω = set of all $n\times n$ invertible matrices (Can show Ω is open!)

$f: \Omega \to W$, $f(A) = A^{-1}$

Can show $f'(A)H = -A^{-1}HA^{-1}$.

Note f diffble. at $a \Rightarrow f$ continuous at a.

Directional derivative $f: \Omega \subseteq V \to W$, $a \in \Omega$, $h \overset{\neq 0}{\in} V$
we say that f is differentiable at a in the direction of h if ~~the~~ .

$$\partial_h f(a) \equiv \lim_{t\to 0} \frac{f(a+th)-f(a)}{t} \text{ exists}$$

Defn. We say that $f: \Omega \subset V \to W$ is G or Gateaux diffble. at a if $\partial_h f(a)$ exists $\forall\, h \in V$.

Remark (1) $f'(a)$ exists $\Rightarrow$ f is G-diffble. at a and $\partial_h f(a) = f'(a)h$

(2) It is often easier to calculate $f'(a)h$ via the directional derivative defn.

(3) G-differentiability is weaker than F-differentiability. In fact it does not imply continuity.

Second Derivatives Let $f : \Omega \subseteq V \to W$.

Suppose f is diff'ble. at every $a \in \Omega$.

This defines a derivative function $f' : \Omega \subseteq V \to \mathcal{L}(V; W)$

Suppose f' is diff'ble. at $a \in \Omega$, then

$$f''(a) \equiv (f')'(a) \in \mathcal{L}(V; \mathcal{L}(V; W))$$

It can be shown that $\mathcal{L}(V; \mathcal{L}(V; W)) = \mathcal{L}(V \times V; W)$
the space of all bdd. bilinear maps

Ex. $f(x,y) = 4x^3 + 7y^4 - 2x^2 y \qquad f : R^2 \to R$

$$f'(a)h = f'(a_1, a_2) = ((12a_1^2 - 4a_1 a_2)h_1, (28 a_2^3 - 2a_1^2)h_2)$$

$$f'(x,y) \overset{h}{=} (12x^2 - 4xy, 28y^3 - 2x^2) \cdot (h_1, h_2) \;;$$

$f''(a_1, a_2)$ is a bdd. bilinear map $: R^2 \times R^2 \to R$

<u>Directional derivatives</u> $\quad f : \Omega \subset X \to Y$, $a \in \Omega$, $0 \neq h \in X$

If the limit $\quad \lim_{t \to 0} \frac{1}{t}\left[f(a+th) - f(a)\right] \equiv \partial_h f(a) = \frac{d}{dt} f(a+th)\Big|_{t=0}$

exists, we say that f is differentiable at a in the direction of the vector h.

<u>Defn.</u> If $\partial_h f(a)$ exists for all $h \in X$, we say that f is <u>Gateaux</u> differentiable at a.

<u>Remark</u> $f'(a)$ exists $\Rightarrow$ f is G-differentiable at a and $\boxed{f'(a)h = \partial_h f(a).}$

The converse of this is not true. G-differentiability at a does not even imply continuity of f at a.

<u>Ex.</u>
$$f(x,y) = \begin{cases} 0 & \text{if } (x,y) = 0 \\ \dfrac{x^5}{(y-x^2)^2 + x^8} & \text{if } (x,y) \neq 0. \end{cases}$$

$f : R^2 \to R$ is G-diffble. but not diffble. at $(0,0)$.

Let $f : \Omega \subset X \longrightarrow Y = Y_1 \times Y_2 \times \cdots \times Y_m$.

We can think of f as having m "components" $f_i : \Omega \subset X \to Y_i$
$i = 1, \ldots, m$

$$f(x) = \begin{bmatrix} f_1(x) \\ \vdots \\ f_m(x) \end{bmatrix}, \quad x \in \Omega.$$

It can be shown that f diff'ble. at $a \in \Omega \iff$ each f_i is diff'ble. at a and that

$$f'(a) = \begin{pmatrix} f_1'(a) \\ \vdots \\ f_m'(a) \end{pmatrix}, \quad f_i'(a) \in \mathcal{L}(X, Y_i)$$

using the identification $\mathcal{L}(X, Y) = \mathcal{L}(X, Y_1) \times \cdots \times \mathcal{L}(X, Y_m)$.
$= \mathcal{L}(X; Y_1 \times Y_2 \times \cdots \times Y_m)$

Now consider $f : \Omega \subset X_1 \times X_2 \times \cdots \times X_n \longrightarrow Y$.

Let $a = (a_1, \ldots, a_n) \in \Omega$ and fix k, $1 \leq k \leq n$.

Then $\exists$ $\Omega_k \subseteq X_k$, open such that $(a_1, \ldots, a_{k-1}, x_k, \ldots, a_n) \in \Omega$ for $x_k \in \Omega_k$. So we may examine the differentiability of the function

$$f_k : \Omega_k \ni x_k \longmapsto f_k(x_k) = f(a_1, \ldots, a_{k-1}, x_k, a_{k+1}, \ldots, a_n) \in Y.$$

If f_k is differentiable at $a_k \in \Omega_k$, we denote its derivative by $\partial_k f_k(a_k) \in \mathcal{L}(X_k, Y)$. We call this the k-th partial derivative of f at a and denote it by $\partial_k f(a)$.

If f is diff'ble. at a, then all its partial derivatives exist and

$$f'(a)h = \sum_{k=1}^{n} \partial_k f(a) h_k , \quad \text{with} \quad h = \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix} \in X .$$

The converse of this fact is not true as evidenced by

$$f : x = (x_1, x_2) \in R^2 \longrightarrow \begin{cases} 0 & \text{if} \quad x_1 x_2 = 0 \\ 1 & \text{if} \quad x_1 x_2 \neq 0 . \end{cases}$$

not continuous at $(0,0)$, hence not diff'ble. there.

If $X = X_1 \times \cdots \times X_n$ and $Y = Y_1 \times Y_2 \times \cdots \times Y_m$, then

$$k = f'(a) h \quad \text{with} \quad h = \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix} \in X \text{ and } k = \begin{pmatrix} k_1 \\ \vdots \\ k_m \end{pmatrix} \in Y$$

is equivalent to

$$k_i = \sum_{j=1}^{n} \partial_j f_i(a) h_j , \quad 1 \leq i \leq m$$

which can be expressed as

$$\begin{pmatrix} k_1 \\ \vdots \\ k_m \end{pmatrix} = \begin{pmatrix} \partial_1 f_1(a) & \cdots & \partial_n f_1(a) \\ \vdots & & \vdots \\ \partial_1 f_m(a) & \cdots & \partial_n f_m(a) \end{pmatrix} \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix} .$$

<u>Note</u> If $X = R^n = R \times \cdots \times R$ and $Y = R^m = R \times \cdots \times R$, then

$\partial_j f_i(a)$ can be interpreted as a real number and the $m \times n$ matrix $(\partial_j f_i(a))$ represents the linear transformation $f'(a) \in \mathcal{L}(R^n, R^m)$.

Derivative of a composite function

Thm. Let $f:\Omega\subset X\to Y$ be diff'ble. at $a\in\Omega$ and let $g:\Omega'\subset Y\to Z$ be differentiable at $b=f(a)\in\Omega'$. Suppose that $f(\Omega)\subset\Omega'$. Then the composite function $h=g\circ f:\Omega\subset X\to Z$ is diff'ble. at a and

$$h'(a)=g'(b)f'(a)\in\mathcal{L}(X,Z).$$

For the special case of $f:\Omega\subset R^n\to R^m$ and $g:\Omega'\subset R^m\to R^\ell$,

$$\begin{bmatrix}\partial_1 h_1(a) & \cdots & \partial_n h_1(a)\\ \vdots & & \\ \partial_1 h_\ell(a) & \cdots & \partial_n h_\ell(a)\end{bmatrix}=\begin{bmatrix}\partial_1 g_1(b) & \cdots & \partial_m g_1(b)\\ \vdots & & \\ \partial_1 g_\ell(b) & \cdots & \partial_m g_\ell(b)\end{bmatrix}\begin{bmatrix}\partial_1 f_1(a) & \cdots & \partial_n f_1(a)\\ \vdots & & \\ \partial_1 f_m(a) & \cdots & \partial_n f_m(a)\end{bmatrix}$$

or $$\partial_j h_i(a)=\sum_{k=1}^m \partial_k g_i(b)\,\partial_j f_k(a),\qquad \begin{matrix}1\le j\le n\\ 1\le i\le \ell\end{matrix}$$

The Implicit Function Theorem

If X and Y are two normed vector spaces, then

$$\mathrm{Isom}(X;Y) \quad \text{or simply} \quad \mathrm{Isom}(X) \text{ if } Y=X$$

denotes the set of all functions which are continuous, linear and bijective, i.e. 1-1 and onto Y, with continuous inverse functions.

Ex. $\mathrm{Isom}(\mathbb{R}^n;\mathbb{R}^n)$ = set of all $n \times n$ invertible matrices (identification).

Theorem 7.1-3 (Implicit function theorem) Let $\varphi : \Omega \subset X_1 \times X_2 \to Y$, $\varphi \in C^1(\Omega)$ and $(a_1, a_2) \in \Omega$, $b \in Y$ be such that

$$\varphi(a_1, a_2) = b \;, \quad \partial_2 \varphi(a_1, a_2) \in \mathrm{Isom}(X_2; Y).$$

Suppose that X_2 is complete. Then there exists an open subset $O_1 \subset X_1$, an open subset $O_2 \subset X_2$ and a continuous function f, called the implicit function $f : O_1 \subset X_1 \to X_2$, such that $(a_1, a_2) \in O_1 \times O_2$ and

$$\{(x_1, x_2) \in O_1 \times O_2 : \varphi(x_1, x_2) = b\} = \{(x_1, x_2) \in O_1 \times \overset{O_2}{X_2} : x_2 = f(x_1)\}.$$

Moreover, the function f is differentiable at a_1 and

$$f'(a_1) = -\{\partial_2 \varphi(a_1, a_2)\}^{-1} \partial_1 \varphi(a_1, a_2).$$

Remarks (i) For a given $x_1 \in O_1$, it is possible that there exists $x_2' \in X_2$, $x_2' \neq x_2$, $(x_1, x_2') \in \Omega$ and $\varphi(x_1, x_2') = b$ but $x_2' \notin O_2$.

(ii) $\exists$ an open set $O_1' \subset O_1$ such that $a_1 \in O_1' \subset \Omega_1$ and such that f is differentiable on O_1'.

Corollary (Inverse function theorem). Let $g : \Omega_2 \subset X_2 \to X_1$ be a continuously differentiable function and let $a_1 \in X_1$, $a_2 \in \Omega_2$ be such that

$$a_1 = g(a_2), \quad g'(a_2) \in \mathrm{Isom}(X_2; X_1).$$

Suppose that X_2 is complete. Then there exists open $O_1 \subseteq X_1$ and open $O_2 \subset X_2$ and a continuous function $f : O_1 \subset X_1 \to X_2$ such that $a_2 \in O_2 \subset \Omega_2$ and

$$\{(x_1, x_2) \in O_1 \times O_2 : x_1 = g(x_2)\} = \{(x_1, x_2) \in O_1 \times \overset{O_2}{X_2} : x_2 = f(x_1)\},$$

i.e. f and g are inverses of each other. Moreover, f is diff'ble. at a_1 (in fact on an open set $O_1' \subset O_1$) and

$$f_1'(a_1) = \{g'(a_2)\}^{-1}. \quad \text{as operators}$$

$a_1 = 1, \quad a_2 = (1,-1)$

Ex. The point $(1,1,-1)$ satisfies the equation

$$\cos(x-y) + e^{x+z} - 2 = 0 \qquad \equiv \varphi_1(x,y,z)$$
$$xz + \sin(x-y) + 1 = 0 \qquad \equiv \varphi_2(x,y,z)$$

$X_1 = R, \quad X_2 = R^2, \quad Y = R^2, \quad \varphi : R^3 \to R^2.$ ✓

$$\partial_2\varphi = \begin{pmatrix} \partial_y\varphi_1 & \partial_z\varphi_1 \\ \partial_y\varphi_2 & \partial_z\varphi_2 \end{pmatrix} = \begin{pmatrix} \sin(x-y) & e^{x+z} \\ -\cos(x-y) & x \end{pmatrix}.$$

$$\partial_2\varphi(a_1,a_2) = \begin{pmatrix} 0 & 1 \\ -1 & 1 \end{pmatrix} \in \mathscr{L}(R^2,R^2) = \mathscr{L}(X_2,Y).$$

All the hypotheses of the Imp. function thm are satisfied.
Hence, $\exists$ an open nbhd. O_1 of 1 such that
an " " O_2 of $(1,-1)$
such that

$$y = f_1(x) \qquad x \in O_1$$
$$z = f_2(x) \qquad x \in O_1, \ (y,z) \in O_2.$$

$$f : O_1 \subset R \to O_2 \subset X_2 = R^2$$

$$\left.\begin{pmatrix} \frac{dy}{dx} \\ \frac{dz}{dx} \end{pmatrix}\right|_{x=1} = \left.\begin{pmatrix} f_1'(x) \\ f_2'(x) \end{pmatrix}\right|_{x=1} = -\left.\begin{pmatrix} 0 & 1 \\ -1 & 1 \end{pmatrix}^{-1}\begin{pmatrix} -\sin(x-y) + e^{x+z} \\ z + \cos(x-y) \end{pmatrix}\right|_{(1,1,-1)}$$

$$= \begin{pmatrix} -1 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} -1 \\ -1 \end{pmatrix}.$$

Mean Value Theorem

Let $f : [a, b] \to R$ be continuous on $[a, b]$ and diff'ble. on (a, b). Then $\exists\, c \in (a, b)$ such that

$$f(b) - f(a) = f'(c)(b - a).$$

This extremely useful result does not hold in this form in the more general setting $f : X \to Y$![(+)]

Often, one is only interested in obtaining a bound for $f(b) - f(a)$.

$$|f(b) - f(a)| \leq \Big(\sup_{t \in (a,b)} |f'(t)|\Big) |b - a| .$$

This inequality can be generalized. First, for $a, b \in X$, introduce the <u>open</u> and <u>closed</u> line segments

$$(a, b) = \{x = ta + (1-t)b : \ 0 < t < 1\}$$

$$[a, b] = \{x = ta + (1-t)b : \ 0 \leq t \leq 1\}.$$

<u>Theorem 7.1-2</u> (Mean-Value theorem) Let $f : \Omega \subset X \to Y$ and suppose the segment $[a, b] \in \Omega$. Suppose f is continuous on $[a, b]$ and differentiable on (a, b). Then

$$\| f(b) - f(a) \|_Y \leq \sup_{x \in (a,b)} \| f'(x) \|_{\mathcal{L}(X, Y)} \| b - a \|_X .$$

It follows from Taylor's theorem that the MVT can be generalized if $Y = R$.

Theorem 7.1-4 Let $f: \Omega \subset X \to Y$ and let $[a, a+h] \subset \Omega$.

(1) If f is diff'ble. at a, then

$$(*) \qquad f(a+h) = f(a) + f'(a)h + \|h\| \,\varepsilon(h), \qquad \lim_{h \to 0} \varepsilon(h) = 0$$

(2) Mean-value theorem: If $f \in C^0(\Omega)$ and f is diff'ble. on $(a, a+h)$ then

$$\|f(a+h) - f(a)\| \le \sup_{x \in (a, a+h)} \|f'(x)\| \, \|h\| .$$

(3) The Taylor-Maclaurin formula: If $f \in C^0(\Omega)$ and f is diff'ble. on (a,h) and $Y = \mathbb{R}$, then $\exists\ \theta \in (0,1)$ such that

$$(**) \qquad f(a+h) = f(a) + f'(a+\theta h)h .$$

(4) Taylor's formula with integral remainder: If $f \in C^1(\Omega)$ and Y is complete, then

$$f(a+h) = f(a) + \int_0^1 \{f'(a+th)h\}\, dt .$$

Remark (i) $(*)$ is precisely the definition of differentiability at a.

(ii) If $Y = \mathbb{R}^n$, then $(**)$ holds with

$$f_i(a+h) = f_i(a) + f_i'(a+\theta_i h)h, \qquad 0 < \theta_i < 1, \ i = 1, \dots, n$$

(iii) In (4), Y must be complete in order for the integral to make sense.

The Second derivative

Suppose $f: \Omega \subset X \to Y$ is differentiable on Ω. If the derivative $f': \Omega \subset X \to Y$ is differentiable at $a \in \Omega$, its derivative, denoted by

$$f''(a) \equiv (f')'(a) \in \mathcal{L}(X; \mathcal{L}(X;Y))$$

is called the <u>second derivative</u> of f at a.

In practice, $f''(a)$ can be calculated from

$$(f''(a)h)k = \lim_{t \to 0} \frac{1}{t}\{f'(a+th) - f'(a)\}k \quad , \quad h,k \in X.$$

It can be shown that $f''(a)$ is <u>symmetric</u>, i.e.

$$(f''(a)h)k = (f''(a)k)h \qquad \forall\, h,k \in X.$$

Hence, we may view $f''(a)$, for fixed a, as a <u>bilinear</u> map on X, which is also continuous, i.e. $f''(a) \in \mathcal{L}_2(X;Y)$, the vector space of all maps $B: X \times X \to Y$ such that the maps $x_1 \mapsto B(x_1, x_2)$ and $x_2 \mapsto B(x_1, x_2)$ are linear on X and which are continuous in the sense that

$$\|B\|_{\mathcal{L}_2(X;Y)} = \|B\| \equiv \sup_{0 \neq x_1, x_2 \in X} \frac{\|B(x_1,x_2)\|_Y}{\|x_1\|_X \|x_2\|_X} < \infty.$$

Ex. Let A be an $n \times n$ matrix. The map

$B: R^n \times R^n \to R$ given by $B(x,y) = x^T A y$

is a continuous bilinear map on R^n and hence belongs to $\mathcal{L}_2(R^n; R)$.

Ex. $\int_0^1 fg\,dx$ defines a continuous bilinear form on $L^2(0,1)$.

Remarks (i) If $Y = R$, we call an element of $\mathcal{L}_2(X;R)$ a bilinear form.

(ii) In a similar fashion, $\mathcal{L}_n(X;Y)$ denotes the space of all continuous n-linear maps from X to Y. Indeed, the n-th derivative of a function $f: \Omega \subset X \to Y$ at a (if it exists) can be shown to be a symmetric, continuous, n-linear map.

H.W. Compute the 2nd derivative of the map in 7.1-3.

Hint: Use the identity $(I-A)^{-1} = I - A\,(I-A)^{-1}$
$= I - (I-A)^{-1}A$

If $X = R^n$ and $Y = R$, then $f''(a)(h,k) = \sum_{i,j=1}^{N} h_i k_j f''(a)(e_i, e_j)$ since $f''(a)$ is a bilinear form. Also, $f''(a)(e_i, e_j) = f''(a)(e_j, e_i)$

Theorem 7.1-5 (Taylor's formulae for twice diffble. fns.)

Let $f : \Omega \subset X \to Y$ and let $[a, a+h] \subset \Omega$.

(1) If f is diffble in Ω and twice diffble. at a, then

$$\circledast \quad f(a+h) = f(a) + f'(a)h + \frac{1}{2} f''(a)(h,h) + \|h\|^2 \varepsilon(h), \quad \lim_{h \to 0} \varepsilon(h) = 0.$$

(2) If $f \in C^1(\Omega)$ and f is twice diffble. on $(a, a+h)$, then

$$\|f(a+h) - f(a) - f'(a)h\| \leq \frac{1}{2} \sup_{x \in (a, a+h)} \|f''(x)\| \, \|h\|^2.$$

(3) If $f \in C^1(\Omega)$ and f is twice diffble on $(a, a+h)$ and $Y = \mathbb{R}$, then

$$f(a+h) = f(a) + f'(a)h + \frac{1}{2} f''(a + \theta h)(h,h), \quad 0 < \theta < 1.$$

(4) If $f \in C^2(\Omega)$ and Y is complete, then

$$f(a+h) = f(a) + f'(a)h + \int_0^1 (1-t)\{f''(a+th)(h,h)\}\, dt.$$

Remarks (i) Condition $\circledast$ is not equivalent to f being twice diffble. at a. cf pb. 7.1-1.

The Gradient and the Hessian

Let $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$.

Suppose f is diff'ble at $a \in \Omega$. We introduce the gradient vector $\nabla f(a) \in \mathbb{R}^n$ as follows:

$$(\nabla f(a), h) = f'(a)h \qquad \forall h \in \mathbb{R}^n,$$

where $(\cdot,\cdot)$ denotes the Euclidean inner product in $\mathbb{R}^n$, i.e.

$$(x,y) = \sum_{i=1}^{n} x_i y_i .$$

It is easy to see that, with this choice of inner product,

$$\nabla f(a) = \begin{pmatrix} \partial_1 f(a) \\ \vdots \\ \partial_n f(a) \end{pmatrix}.$$

Let us note that changing the inner product will change $\nabla f(a)$.

Similarly, if f is twice differentiable at a, then we define the (symmetric) Hessian matrix $\nabla^2 f(a) \in \mathbb{R}^{n\times n}$ by

$$(\nabla^2 f(a) h, k) \equiv (h, \nabla^2 f(a) k) = f''(a)(h,k), \quad \forall h, k \in \mathbb{R}^n.$$

It follows that

$$\nabla^2 f(a) = \begin{bmatrix} \partial_{11} f(a) & \cdots & \partial_{1n} f(a) \\ \vdots & & \vdots \\ \partial_{n1} f(a) & \cdots & \partial_{nn} f(a) \end{bmatrix}.$$

Thus, the Gradient vector and the Hessian matrix are particular representations of $f'(a)$ and $f''(a)$ respectively corresponding to $(\cdot,\cdot)_2$.

Since $f''(a)$ is symmetric, the numbers

$$\partial_{ij}f(a) \equiv \partial_i(\partial_j f)(a) = \partial_j(\partial_i f)(a) = f''(a)(e_i, e_j)$$

are the <u>second partial derivatives</u> of f at a.

18

(+) A neighborhood of a point $x \in W$ is a set that contains an open set containing x.

§7.2 Extrema of real functions: Lagrange multipliers.

Defn. Let $J : W \to \mathbb{R}$ be a function defined over a topological ~~(vector)~~ space W. We say that J has a relative minimum (or a relative maximum) at $u \in W$ if there exists a nbhd.(+) $\mathcal{O}$ of u such that

$$J(u) \leq J(v) \quad (\text{or } J(u) \geq J(v)) \quad \forall\, v \in \mathcal{O}.$$

If there is no need to distinguish between a rel. min. and a rel. max., we say that f has a relative extremum at u.

Theorem 7.2-1 (Necessary condition for a relative extremum) Let Ω be an open subset of a normed vector space V and $J : \Omega \subset V \to \mathbb{R}$ be given. If J has a relative extremum at $u \in \Omega$ and if $J'(u)$ exists, then

$$J'(u) = 0.$$

Proof.

Let $v \in V$. Since Ω is open, $\exists$ open interval I containing 0 such that the (composite) map

$$\varphi : t \in I \longmapsto J(u + tv)$$

is well-defined. Indeed, $\|u + tv\| \leq \|u\| + |t|\,\|v\|$. So for $|t|$ sufficiently small, $u + tv \in$ Ball centered at $u \subset \Omega$.

φ is a composite fn.; By Thm. 7.1-1, φ is diff'ble. at 0 and

$$\varphi'(0) = J'(u)\, v.$$

Assume w.l.o.g. that u is a relative minimum. Then

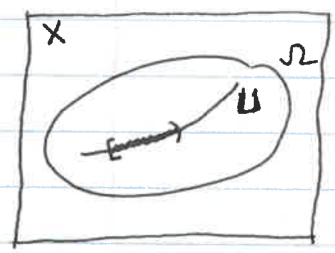
$$0 \geq \lim_{t \to 0^-} \frac{\varphi(t) - \varphi(0)}{t} = \varphi'(0) = \lim_{t \to 0^+} \frac{\varphi(t) - \varphi(0)}{t} \geq 0$$

we have thus shown that $\varphi'(0) = J'(u)v = 0 \quad \underline{\forall v \in V}$.

Obviously this implies that $J'(u) = 0$. $\square$

<u>Remarks</u> (i) The fact that Ω is open is <u>essential</u>.
Consider for example $f: [0,1] \to R$, $f(x) = x$.

Let U be a (not necessarily open) <u>subset</u> of Ω.
A <u>neighborhood</u> of a point $u \in U$ (in the relative topology induced by the topology of X) is a subset of U that <u>contains</u> the intersection of U with an <u>open</u> subset of X



<u>Defn.</u> Let $J: \Omega \subset X \to R$ and U a subset of Ω. We say that J has a relative minimum (rel. maximum) at a point $u \in U$ with respect to the set U, if the restriction of J to U, endowed with the induced topology, has a relative minimum (rel. maximum) at u.

This means that with J_U denoting the restriction of J to U, $\exists$ a nbhd. N_u of u <u>in U</u> such that

$$J_U(u) \leq J_U(v) \quad \forall v \in N_u : \text{ rel. minimum}$$

or

$$J_U(u) \geq J_U(v) \quad \forall v \in N_u : \text{ rel. maximum.}$$

This is ~~an~~ example of a <u>constrained</u> relative extremum the set U "supplying" the constraint(s).

Theorem 7.2-2 (Necessary condition for a constrained relative minimum)

Let $\Omega \subset V_1 \times V_2$ be open, V_2 complete. Let

$\varphi : \Omega \to V_2$, $\varphi \in C^1(\Omega)$ and let $u = (u_1, u_2)$ be a pt. of the set

$$U = \{(v_1, v_2) \in \Omega : \varphi(v_1, v_2) = 0\} \subset \Omega$$

at which

$$\partial_2 \varphi(u_1, u_2) \in \mathrm{Isom}(V_2).$$

Let $J : \Omega \to \mathbb{R}$ be differentiable at u. If J has a relative extremum at u with respect to the set U, then, there exists an element $\Lambda(u) \in \mathcal{L}(V_2, \mathbb{R})$ such that

$$J'(u) + \Lambda(u)\varphi'(u) = 0.$$

proof. The conditions above enable us to use the Implicit function theorem. Hence, $\exists$ open $O_1 \subset V_1$, open $O_2 \subset V_2$ and a continuous fn. $f : O_1 \to O_2$ such that $(u_1, u_2) \in O_1 \times O_2$ and

$$(O_1 \times O_2) \cap U = \{(v_1, v_2) \in O_1 \times \overset{\text{or } V_2}{O_2} : v_2 = f(v_1)\}.$$

Further, f is differentiable at $u_1 \in O_1$ and

$$f'(u_1) = -\{\partial_2 \varphi(u_1, u_2)\}^{-1} \partial_1 \varphi(u_1, u_2).$$

$\Rightarrow$ The restriction of J to U can be viewed as a function of a "single" variable as follows

$$G : v_1 \in O_1 \to G(v_1) \equiv J(v_1, f(v_1)) \in \mathbb{R}.$$

Now G has a rel. extremum at u_1 in the open set O_1. Thus, by Theorem 7.2-1,

Proof. $\{\varphi_i'(u)\}_{i=1}^m$ lin. indep. $\Rightarrow$ There exists a collection of m columns of an $m\times n$ matrix $\{\partial_j\varphi_i(u)\} \equiv \mathcal{M}$ $= \{\partial_1\varphi_i(u), \dots, \partial_n\varphi_i(u)\}$ that are lin. independent, Each column corresponding to one variable of $\{\varphi_i\}_{i=1}^m$. Reordering these variables, we may assume w.l.o.g. that these are the last m columns of $\mathcal{M}$. Letting $V_1 = R^{n-m}$ and $V_2 = R^m$, and writing

$$x = (\underbrace{x_1, \dots, x_{n-m}}_{v_1}, \underbrace{x_{n-m+1}, \dots, x_n}_{v_2}), \text{ (reordered variables)}$$

and

$$\varphi = \begin{pmatrix} \varphi_1 \\ \vdots \\ \varphi_m \end{pmatrix} : R^{n-m} \times R^m \to R^m,$$

we have that $\partial_2 \varphi(u_1, u_2)$ is invertible as an element of $\mathcal{L}(R^m, R^m)$.

Thus, we may apply theorem 7.2-2 to get

$$J'(u) + \Lambda(u)\varphi'(u) = 0, \quad \Lambda(u) \in \mathcal{L}(R^m, R)$$

or

$$(*) \quad J'(u) + \sum_{i=1}^m \lambda_i(u)\varphi_i'(u) = 0. \quad \square$$

Remark The componentwise form of $(*)$ is

$$\partial_j J(u) + \sum_{i=1}^m \lambda_i(u)\,\partial_j\varphi_i(u) = 0, \quad j = 1, \dots, n$$

which may be written as

$$\begin{bmatrix} \partial_1 J(u) \\ \vdots \\ \partial_n J(u) \end{bmatrix} + \underbrace{\begin{bmatrix} \partial_1\varphi_1(u) & \cdots & \partial_1\varphi_m(u) \\ \vdots & & \\ \partial_n\varphi_1(u) & \cdots & \partial_n\varphi_m(u) \end{bmatrix}}_{\{\varphi'(u)\}^T} \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_m \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$$

or

$$\nabla J(u) + \sum_{i=1}^m \lambda_i(u)\,\nabla\varphi_i(u) = 0$$

(†) The converse of this result is not true as evidenced by the following examples: $x \in R \to f(x) = x^3 \in R$. Here, $f'(0)=0$ but f does not have a local extremum at 0.
Also, the property of Ω open is crucial as evidenced by the example of $x \in R \mapsto f(x) = x$, with $\Omega = [0,1]$.

$$0 = G'(u_1) = \partial_1 J(u) + \partial_2 J(u) f'(u_1)$$
$$= \partial_1 J(u) - \partial_2 J(u) \{\partial_2 \varphi(u_1,u_2)\}^{-1} \partial_1 \varphi(u)$$

$$= \cancel{\partial_1 J(u) - \partial_2 J(u) \{\partial_2 \varphi(u)\}^{-1} \partial_1 \varphi(u)} .$$

Using the trivial identity
this together with

$$0 = \partial_2 J(u) - \partial_2 J(u) \{\partial_2 \varphi(u)\}^{-1} \partial_2 \varphi(u),$$

we obtain

$$0 = J'(u) - \partial_2 J(u) \{\partial_2 \varphi(u)\}^{-1} \varphi'(u).$$

Thus the theorem follows with $\Lambda(u) = -\partial_2 J(u) \{\partial_2 \varphi(u)\}^{-1}$. □†

We next consider the application of this result to the important special case of

$$J : \Omega \subset R^n \to R, \quad \varphi_i : (\Omega) \subset R^n \to R, \quad i = 1, \dots, m$$

where $1 \le m < n$.

<u>Theorem 7.2-3</u> Let $\varphi_i : \Omega \subset R^n \to R$, $i = 1, \dots, m < n$ be m functions of class C^1 over Ω and let u belong to the set

$$U = \{ v \in \Omega : \varphi_i(v) = 0, \quad i = 1, \dots, m \} \subset \Omega,$$

and suppose that the derivatives $\underbrace{\varphi_1'(u), \dots, \varphi_m'(u)}_{\text{(vectors in } R^n)}$ are linearly independent.
Suppose $J : \Omega \to R$ is differentiable at u.
If J has a relative extremum with respect to U at u, then $\exists$ m numbers $\lambda_1(u), \dots, \lambda_m(u)$ such that

$$J'(u) + \sum_{i=1}^{m} \lambda_i(u) \varphi_i'(u) = 0.$$

<u>Ex.</u> let A be a given $n\times n$ symmetric matrix and let $b\in R^n$ be a given vector. Define the functional

$$J: v\in R^n \longrightarrow J(v) = \frac{1}{2}(Av, v) - (b, v).$$

J is diff'ble. and $\nabla J(u) = Au - b$.

Thus, if we wish to solve the linear system $Au = b$, then we must find the local extrema of J.

<u>Ex.</u> Suppose we wish to find the local extrema of the functional J above, subject to the constraint $Cu = d$ where C is an $m\times n$ matrix and $d\in R^m$.

Suppose $m < n$ and that C has rank m.
If u has a relative extremum
at the point $u\in U$ with respect to the set $U = \{v\in R^n : Cv - d = 0\}$, then by theorem 7.2-3, $\exists\ \Lambda\in R^m$ such that

$$0 = \nabla J(u) + C^T\Lambda \quad = Au - b + C^T\Lambda = 0$$

combining this with $Cu = d$, we get the system

$$\begin{pmatrix} A & C^T \\ C & 0 \end{pmatrix}\begin{pmatrix} u \\ \Lambda \end{pmatrix} = \begin{pmatrix} b \\ d \end{pmatrix};$$

which is an $n+m$ system of linear equations in the unknowns u and Λ.

Cone of admissible (feasible) directions

Defn. A Cone S is a subset of a vector space V if whenever $v \in S$, $\lambda v \in S$ $\forall \lambda \geq 0$.

Note that a cone contains 0 by definition. On the other hand, we can generalize the definition into cones that do not contain the origin by translation:

let $w \in V$, $w \neq 0$, S a cone centered at the origin, then

$w + S$ is a cone centered at w.

Defn. Let U be a nonempty subset of a vector space V. For every point u in U, the cone $C(u)$ of admissible directions is the union of $\{0\}$ and the set of nonzero vectors w in V for which there exists a sequence $\{u_k\}_{k\geq 0}$ such that

$$u_k \in U, \quad u_k \neq u, \quad \lim_{k\to\infty} u_k = u$$

$$\lim_{k\to\infty} \frac{u_k - u}{\|u_k - u\|} = \frac{w}{\|w\|}.$$

Equivalently, there exists sequence $\{u_k\}_{k\geq 0}$ in U and $\{\delta_k\}_{k\geq 0}$ in V

$$u_k = u + \|u_k - u\| \frac{w}{\|w\|} + \|u_k - u\| \delta_k, \quad \lim_{k\to\infty} \delta_k = 0, \quad w \neq 0.$$

Equivalently, $w \neq 0$ is a feasible direction at u if

$\exists \{w_k\}$ with $\lim_{k\to\infty} w_k = w$, $\exists \{\epsilon_k\}$, $\epsilon_k > 0$ with $\lim_{k\to\infty} \epsilon_k = 0$ such that $u + \epsilon_k w_k \in U$.

Note that the cone of feasible directions $C(u)$ is indeed a cone centered at 0. The translated cone of feasible directions is $u + C(u)$.

Remark (i) It is possible to have $C(u) = \{0\}$.
(ii) If $u \in \Omega \subseteq U$, Ω open in the topology of V,

Then $C(u) = V$.

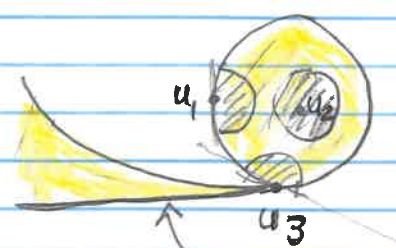
Ex. U is a singleton or a finite set of points. Then

$$C(u) = \{0\}.$$

$C(u_1)$ is "half-space" at u_1 tangent to the boundary

$C(u_2)$ is R^2

$C(u_3)$ is the union of the half-space tangent to the boundary at u_3 and the ray



Lemma (Euler Inequality, general case). Let u be a local minimum of J with respect to U. If J is differentiable at u, then

$$J'(u)w \geq 0 \qquad \forall\, w \in C(u).$$

proof

If $w = 0$, then inequality holds. Let $w \neq 0$ and consider

$$u_k = u + \|u_k - u\| \frac{w}{\|w\|} + \|u_k - u\| \delta_k, \quad u_k \neq u, \quad \delta_k \to 0, \quad u_k \to u$$

From Taylor's Thm,

$$J(u_k) = J(u) + J'(u)\left\{\|u_k - u\|\left[\frac{w}{\|w\|} + \delta_k\right]\right\} + \|u_k - u\|\left[\frac{w}{\|w\|} + \delta_k\right]\eta$$

where $\eta \to 0$ as $\|u_k - u\| \to 0$

Hence

$$\frac{J(u_k)-J(u)}{\|u_k-u\|} = J'(u)\left[\frac{w}{\|w\|}+\delta_k\right]+\left[\frac{w}{\|w\|}+\delta_k\right]\eta.$$

Since $u_k \in U$ and $u_k \to u$, we will have $J(u_k) \geq J(u)$ for k large since u is a local minimum of J. Taking limits, we have, since $\delta_k \to 0$ and $\eta \to 0$

$$0 \leq J'(u)\frac{w}{\|w\|} \Rightarrow J'(u)w \geq 0. \quad \boxtimes$$

Remark If u is a local maximum of J with respect to U, then

$$J'(u)w \leq 0 \quad \forall w \in C(u).$$

proposition 7.2-1 Let $\Omega \subseteq V = V_1 \times V_2$ be open, V_2 complete. Let $\varphi : \Omega \to V_2$, $\varphi \in C^1(\Omega)$ and let u be a point in the set

$$U = \{v \in \Omega,\ \varphi(v) = 0\},$$ assumed to be nonempty.

Suppose $\partial_2\varphi(u)$ belongs to $Isom(V_2)$. Then, for the cone $C(u)$ of admissible directions at u

$$C(u) = \{w \in V,\ \varphi'(u)w = 0\}.$$

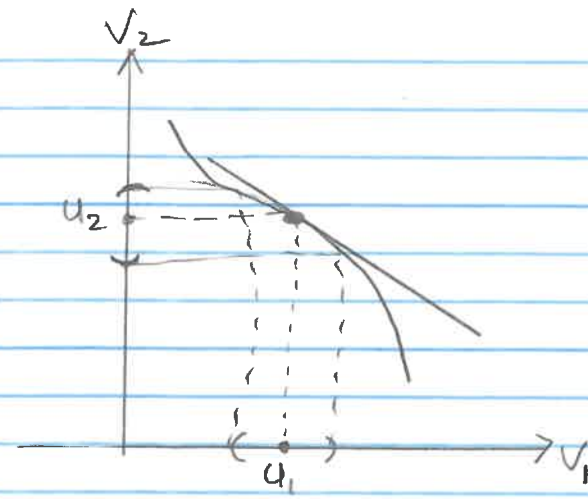
proof. Let $w \in C(u)$. Then $\exists \{w_k\}$, $\lim_{k\to\infty} w_k = w$, $\exists \{\epsilon_k\}$, $\epsilon_k > 0$ with $\lim_{k\to\infty} \epsilon_k = 0$ such that $u + \epsilon_k w_k \in U$. We have

$$0 = \varphi(u+\epsilon_k w_k) = \overset{0}{\cancel{\varphi(u)}} + \epsilon_k \varphi'(u)w_k + \epsilon_k \|w_k\| \eta_k, \quad \lim_{k\to\infty}\eta_k = 0$$

$$\Rightarrow 0 = \varphi'(u)w_k + \|w_k\|\eta_k.$$ Taking limits, we see that

$$\varphi'(u)w = 0.$$

Conversely, let w ($\|w\|=1$) be given with $\varphi'(u)w = 0$. We want to show that $w \in C(u)$, i.e. w is a feasible direction.



Let $w = (w_1, w_2)$, $w_1 \in V_1$, $w_2 \in V_2$.

$\varphi'(u)w = 0$ means

$$\partial_1 \varphi(u) w_1 + \partial_2 \varphi(u) w_2 = 0.$$

Since $\partial_2 \varphi(u)$ is invertible with bounded inverse, we can write

$$① \qquad w_2 = -\{\partial_2 \varphi(u)\}^{-1} \partial_1 \varphi(u) w_1 = f(u_1) w_1$$

where $f : O_1 \to O_2$ is the (implicit) function of the Implicit Function Theorem. Now for $t > 0$ small enough $u_1 + t w_1$ belongs to O_1, so we define the points $u(t)$ by

$$u(t) = (u_1 + t w_1, f(u_1 + t w_2)) \neq u \text{ for } t > 0.$$

We will show that $\lim\limits_{t \to 0} \dfrac{u(t) - u}{\|u(t) - u\|} = \dfrac{w}{\|w\|}$ proving that w is a feasible direction. We have using Taylor's Thm. and ①

$$u(t) - u = (u_1 + t w_1, f(u_1 + t w_1)) - (u_1, u_2)$$

$$= (t w_1, \cancel{f(u_1)} + t f'(u_1) w_1 + t \|w_1\| \varepsilon(t) - \cancel{u_2})$$

$$= t^2 (w_1, w_2 + \|w_1\| \varepsilon(t)), \quad \varepsilon(t) \to 0 \text{ as } t \to 0.$$

Hence

$$\frac{u(t) - u}{\|u(t) - u\|} = \frac{w + \|w_1\| (0, \varepsilon(t))}{\|w + \|w_1\| (0, \varepsilon(t))\|} \longrightarrow \frac{w}{\|w\|} \text{ as } t \to 0. \blacksquare$$

§7.3 Extrema of real functions: Consideration of the 2nd derivatives

$\to$ on 2nd derivative.

Theorem 7.3-1 (Necessary condition for a relative minimum)

Ω open

Let $J : \Omega \subset V \to \mathbb{R}$ be diff'ble. on Ω and twice differentiable at $u \in \Omega$. If J has a local minimum at u, then

$$J''(u)(w,w) \geq 0 \qquad \forall\, w \in V.$$

proof.

Let $w \in V$. Then we can find an open interval $0 \in I \subseteq \mathbb{R}$ such that

$$t \in I \Rightarrow u + tw \in \Omega \quad \text{and} \quad J(u+tw) \geq J(u).$$

Using Taylor's formula, by thm. 7.2-1

$$(*) \qquad 0 \leq J(u+tw) - J(u) = \cancel{J'(u)}^{\,0} + \frac{t^2}{2}\left(J''(u)(w,w) + 2\varepsilon(t)\underbrace{\vphantom{x}}_{\|w\|^2}\right),$$

where $\lim_{t \to 0} \varepsilon(t) = 0$. Now if $J''(u)(w,w) < 0$, then for t sufficiently small, $J''(u)(w,w) + 2\varepsilon(t) < 0$, which would contradict (*). $\square$

Remark (i) If J has a local maximum at u, then

$$J''(u)(w,w) \leq 0 \qquad \forall\, w \in V.$$

(ii) The converse of the above result is not true as shown by the example of the function $f(x) = x^3$.

Defn. Let $J : W \to \mathbb{R}$ be defined over the topological vector space W. J is said to have a strict local minimum, resp. strict local maximum, at the point $u \in W$ if $\exists$ nbhd. O of u such that

$$J(u) < J(v) \qquad \forall\, v \in O - \{u\}$$

or $\quad J(u) > J(v) \quad \forall\, v \in O - \{u\}$ respectively for a strict local maximum.

Theorem 7.3-2 (sufficient conditions for a relative minimum)

let $J : \Omega \subset V \to R$ be differentiable at $u \in \Omega$ and assume that $J'(u) = 0$.

open

(1) If $J''(u)$ exists and $\exists\, \alpha > 0$ such that

$$J''(u)(w,w) \geq \alpha \|w\|^2 \qquad \forall\, w \in V,$$

then J has a strict local minimum at u.

(2) If J is twice diffble. on Ω and if $\exists$ a ball $B \subseteq \Omega$ centered at u such that

$$J''(v)(w,w) \geq 0 \qquad \forall v \in B, \ \forall w \in V,$$

then J has a relative minimum at u.

proof

(1) From Taylor's formula, for $w \in V$ suff. small,

$$J(u+w) - J(u) = \tfrac{1}{2}\left(J''(u)(w,w) + 2\|w\|^2 \varepsilon(w)\right)$$

$$\geq \tfrac{1}{2}\left(\alpha - 2\varepsilon(w)\right)\|w\|^2, \quad \lim_{w \to 0} \varepsilon(w) = 0,$$

let B be a ball centered at u whose radius r is chosen small so that $2\varepsilon(w) < \alpha$. Then $J(u+w) - J(u) > 0$ for all $u \neq w \in B$.

(2) Using the Taylor-Maclaurin formula, we get

$$J(u+w) = J(u) + \frac{1}{2} J''(v)(w,w)$$

$$\geq J(u) \quad , \quad v \in (u, u+w), \quad u+w \in B,$$

which shows that J has a relative minimum at u. $\square$

Ex. Solution to 7.3-1 (1) $f(x) = x^4$

(1)

has strict minimum at $x=0$ but $f''(x)\big|_{x=0} = 12x^2\big|_{x=0} = 0$.

(2)

let
$$f(x) = \begin{cases} x^5 \sin\frac{1}{x} + |x|^5 + x^6, & x \neq 0 \\ 0 & x = 0 \end{cases}$$

It is clear that $f(x) \geq x^6$, $\forall x$. Indeed,

$$x^5 \sin\frac{1}{x} + |x|^5 \geq -|x|^5 + |x|^5 = 0.$$

Hence f has a strict relative and global minimum at 0.

$$f'(x) = \begin{cases} 5x^4 \sin\frac{1}{x} - x^3 \cos\frac{1}{x} + 5x^4 + 6x^5, & x > 0 \\ \quad '' \qquad\qquad '' \qquad - 5x^4 + 6x^5, & x < 0 \\ 0, \quad x = 0 \end{cases}$$

$$f''(x) = \begin{cases} 20x^3 \sin\frac{1}{x} - 8x^2 \cos\frac{1}{x} - x \sin\frac{1}{x} + 20x^3 + 30x^4, & x > 0 \\ \quad '' \qquad\qquad '' \qquad\qquad '' \qquad - 20x^3 + 30x^4, & x < 0 \\ 0, \quad x = 0. \end{cases}$$

It is clear that as $x \to 0$, the term $x \sin\frac{1}{x}$ dominates $f''(x)$; and this term is obviously oscillatory. QED

Ex. Let Ω be a bdd. domain in $\mathbb{R}^n$.

Let $H_0^1(\Omega) = \{ v \in L^2(\Omega), \frac{\partial v}{\partial x_j} \in L^2(\Omega), j = 1, \dots, n \ , \ v/_{\partial\Omega} = 0 \}$.

$H_0^1(\Omega)$ is a complete normed space with respect to the norm

$$\|v\|_1 = \int_\Omega \Big[|v|^2 + \sum_{j=1}^n \Big|\frac{\partial v}{\partial x_j}\Big|^2 \Big] dx .$$

For given $f \in L^2(\Omega)$, define $J : H_0^1 \to \mathbb{R}$ by

$$J(u) = \frac{1}{2}\int_\Omega |\nabla u|^2 dx - \int_\Omega f u \, dx = \frac{1}{2}\int_\Omega \sum_{j=1}^n \Big|\frac{\partial u}{\partial x_j}\Big|^2 - \int_\Omega f u \, dx .$$

J is differentiable on H_0^1. Indeed,

$$J(u+v) - J(u) = \int_\Omega [\nabla u \cdot \nabla v - f v] dx + \frac{1}{2}\int_\Omega |\nabla v|^2 dx .$$

Using Cauchy-Schwarz, we can show that the map

$$v \longmapsto \int_\Omega [\nabla u \cdot \nabla v - f v] dx \quad \text{is an element of} \quad \mathcal{L}(H_0^1; \mathbb{R})$$

Also, since $\frac{1}{2}\int_\Omega |\nabla v|^2 dx \le \frac{1}{2} \|v\|_1^2$, J is diffble. $\forall u \in H_0^1$ and

$$J'(u) v = \int_\Omega [\nabla u \cdot \nabla v - f v] dx .$$

If J has a rel. minimum at some u, then by Thm. 7.1-1,

$$0 = J'(u) v = \int_\Omega [\nabla u \cdot \nabla v - f v] dx \qquad \forall v \in H_0^1 .$$

Now assume that $u \in H^2(\Omega)$. Then, integration by parts yields

$$0 = J'(u) v = \int_\Omega [-\Delta u - f] v \, dx = 0 \qquad \forall v \in H_0^1(\Omega),$$

Since $H_0^1(\Omega)$ is dense in $L^2(\Omega)$, it follows that any smooth minimizer of J is a solution to the 2nd order Elliptic pb. $\begin{cases} -\Delta u = f & \text{in } \Omega \\ u = 0 & \text{on } \partial\Omega . \end{cases}$

§7.4 Extrema of real functions: considerations of convexity

"Line segment" $[a,b] = \{v \in V : v = t\,a + (1-t)\,b,\ 0 \le t \le 1\}$

Defn. A subset U of a vector space V is said to be convex if $[a,b] \subset U$ whenever $a, b \in U$.

Ex. An open, or closed ball is convex.

Defn. A function $J : U \subset V \to \mathbb{R}$, where U is convex, is said to be convex over U if

$$u, v \in U \implies J(\theta u + (1-\theta) v) \le \theta J(u) + (1-\theta) J(v), \ \forall\ 0 \le \theta \le 1,$$

and is strictly convex over U if

$$u, v \in U \implies J(\theta u + (1-\theta) v) < \theta J(u) + (1-\theta) J(v), \ \forall\ 0 < \theta < 1.$$

Defn A function $G : U \subset V \to \mathbb{R}$, where U is convex, is said to be concave (or strictly concave) over U if the function $-G$ is convex (or strictly convex) over U.

Theorem 7.4-1 (Necessary cond. for a relative minimum over a convex set)

Let $J : \Omega \subset V \to \mathbb{R}$ be given and let U be a convex subset of Ω. If J is diff'ble. at $u \in U$ and has a relative minimum at u with respect to U, then,

$$J'(u)(v-u) \ge 0 \qquad \forall\ v \in U.$$

proof.

Let $v \in U$ and write $w = v - u$. U convex implies

$$u + \theta w \in U \qquad \forall\ 0 \le \theta \le 1.$$ Since J is diff'ble. at u,

$$0 \le \frac{J(u+\theta w) - J(u)}{\theta} = J'(u)(v-u) + \varepsilon(\theta); \quad \lim_{\theta \to 0} \varepsilon(\theta) = 0.$$

This shows that $J'(u)(v-u) \ge 0$.

Remarks (i) If U is a vector subspace, then

$$J'(u)(v-u) = 0 \qquad \forall v \in U.$$

To see this, let $\theta \to 0^-$.

(ii) If $U = V$, or if U is open in V, then $J'(u) = 0$. Thus, we recover the result of thm. 7.2-1.

(iii) If U is a convex cone centered at the origin, i.e.

If $x, y \in U$, then so are cx, $c>0$ and $x+y$;

then the condition of Thm. 7.4-1 is equivalent to

$$J'(u)u = 0 \quad \text{and} \quad J'(u)v \geq 0 \quad \forall\, v \in U.$$

To see this, let $v = cu$, for $c > 0$. Then

$$J'(u)(v-u) = (1-c)\, J'(u)\, u \geq 0 \qquad \forall c > 0.$$

taking $c = 1/2$ and $c = 3/2$, we clearly have $J'(u)u = 0$.
Also,

$$0 \leq J'(u)(v-u) = J'(u)\, v.$$

Theorem 7.4-2 (convexity and first derivative) let $J : \Omega \subset V \to R$ be diffble. on Ω and let $U \subset \Omega$ be convex. Then

(1) J is convex on U $\iff$ $J(v) \geq J(u) + J'(u)(v-u), \quad \forall u, v \in U.$

(2) J is strictly convex on U $\iff$ $J(v) > J(u) + J'(u)(v-u), \quad \forall u, v \in U,\ u \neq v$

proof.

(1) let u, v be distinct pts. of U and $\theta \in (0,1)$.

If J is convex, then

$$J(u+\theta(v-u)) = J((1-\theta)u+\theta v) \le (1-\theta)\,J(u) + \theta\,J(v)$$

$$\Rightarrow \quad \frac{J(u+\theta(v-u)) - J(u)}{\theta} \le J(v) - J(u).$$

$$J'(u)(v-u) = \lim_{\theta\to 0^+} \frac{J(u+\theta(v-u))-J(u)}{\theta} \le J(v) - J(u).$$

↑ argument fails here for strict convexity.

Conversely, suppose that $J(v) \ge J(u) + J'(u)(v-u)$, $\forall u, v \in U$.
Let $u, v \in U$ and $\theta \in (0,1)$, then $v + \theta(u-v) = \theta u + (1-\theta)v \in U$, and

$$J(v) \ge J(v+\theta(u-v)) - \theta J'(v+\theta(u-v))(u-v)$$

and

$$J(u) \ge J(v+\theta(u-v)) + (1-\theta) J'(v+\theta(u-v))(u-v)$$

multiplying the first equation by $(1-\theta)$, the 2nd by θ and adding, we get

$$J(v+\theta(u-v)) = J(\theta u + (1-\theta)v) \le \theta J(u) + (1-\theta)J(v)$$

which shows that J is convex on U.

(2) It is an easy exercise to show that the function

$$t : (0,1] \longmapsto \frac{J(tv+(1-t)u) - J(u)}{t} \qquad \leftarrow \text{H.W. Exercise}$$

is increasing (strictly increasing) if J is convex (strictly convex) on the segment $[u,v]$.
Let $0<\theta<\omega<1$, with ω fixed. Then

$$\frac{J(u+\theta(v-u))-J(u)}{\theta} = \frac{J(\theta v + (1-\theta)u)-J(u)}{\theta} < \frac{J(\omega v + (1-\omega)u) - J(u)}{\omega}$$

$$< J(v) - J(u).$$

Taking the limit $\theta\to 0^+$, we get as before

$$J'(u)(v-u) < J(v) - J(u).$$

The proof of the converse in this case is exactly the same as that of part (1), with $<$ replacing $\leq$ in the appropriate places. $\square$

<u>Theorem 7.4-3</u> (Convexity and the second derivative)

Let $J : \Omega \subset V \to \mathbb{R}$ be twice diff'ble. on Ω and let U be a convex subset of Ω.

~~(1) the function J is convex over $U \iff J''(u)(v-u, v-u) \geq 0 \quad \forall u, v \in$~~ U.

(1) $J''(u)(v-u, v-u) \geq 0 \quad \forall u, v \in U \iff J$ is convex over U

(2) $J''(u)(v-u, v-u) > 0 \quad \forall u \neq v \in U \implies J$ is strictly convex over U.

<u>Proof.</u> (1) ($\Rightarrow$)
(2) ($\Rightarrow$) Applying the Taylor-MacLaurin formula,

$$J(v) - J(u) - J'(u)(v-u) = \frac{1}{2} J''(w)(v-u, v-u)$$

where $w = u + \theta(v-u)$ for some $0 < \theta < 1$. Since $v - u = \frac{1}{\theta}(w-u)$, we have

$$J(v) - J(u) - J'(u)(v-u) = \frac{1}{2\theta^2} J''(w)(w-u, w-u).$$

(1) ($\Rightarrow$) and (2) ($\Rightarrow$)
The convexity or strict convexity of J follows from this and the previous theorem.

(1) ($\Leftarrow$)
To prove the converse of (1), let $u \in U$ be arbitrary but fixed. Introduce the function

$$G : v \in \Omega \to G(v) = J(v) - J'(u)\, v.$$

Now

$$G(v) - G(u) = J(v) - J(u) - J'(u)(v-u).$$

If J is convex on U, then by theorem 7.4-2,

$J(v) - J(u) - J'(u)(v-u) \geq 0 \quad \forall v \in U$, hence G has a minimum at u with respect to the set U.

Since J is twice diffble. on Ω, G is twice diffble. on Ω. Also,

$$G'(v) = J'(v) - J'(u) \Rightarrow G'(u) = 0;$$

$$G''(v) = J''(v).$$

Thus, for every $v = u + w \in U$ and every $t \in [0,1]$,

$$0 \leq G(u+tw) - G(u) = t\,\overset{0}{\cancel{G'(u)}}\,w + \frac{t^2}{2}\left[J''(u)(w,w) + \varepsilon(t)\right]$$

with $\lim_{t \to 0} \varepsilon(t) = 0$. This clearly shows that $J''(u)(w,w) \geq 0$. □

Remarks (1) The example of the strictly convex function $J(t) = t^4$ shows that the converse of (2) is false.

(2) The converse of (2) is true for a "quadratic" functional. Indeed, in that case $J'''(u) = 0$ and

$$J(v) - J(u) - J'(u)(v-u) = J''(u)(v-u, v-u) \quad \forall u, v \in \Omega.$$

~~Thus, if~~ $\cancel{J''(u)(v-u, v-u) > 0 \quad \forall u, v \in U \text{ with } v \neq u}$, ~~then~~

$$\cancel{J(v) - J(u) - J'(u)(v-u) > 0 \quad \forall u, v \in U \text{ with } v \neq u},$$

~~and hence~~

$$\cancel{J \text{ is convex by theorem } 7.4\text{-}2 - (2)}.$$

Thus, if J is strictly convex, then by Thm. 7.4-2
$J(v) - J(u) - J'(u)(v-u) > 0 \quad \forall u, v \in U$ with $u \neq v$.
$\Rightarrow J''(u)(v-u, v-u) > 0 \quad \forall u, v \in U$ with $u \neq v$.

Now, if $J(v) = \frac{1}{2}(Av, v) - (b, v)$ with $A = A^T$, then

$$J''(u)(v-u, v-u) = \big(A(v-u), v-u\big).$$

It is clear that
(i) J is convex over $R^n \iff A$ is positive semidefinite
(ii) J is strictly convex over $R^n \iff A$ is positive definite.

Let $J : W \to R$, where W is a set. We say that J has a minimum (or a maximum) at a pt. $u \in W$ if

$$J(u) \le J(v) \quad (\text{or } J(u) \ge J(v)) \quad \forall v \in W;$$

or that has a strict minimum (or a strict maximum) at a point $u \in W$ if

$$J(u) < J(v) \quad (\text{or } J(u) > J(v)) \quad \forall v \in W,\ v \ne u.$$

Theorem 7.4-4 Let U be a convex subset of a normed vector space V.

(1) If a convex function $J : U \subset V \to R$ has a relative minimum at a pt. of U, then it has a minimum there.

(2) A strictly convex function $J : U \subset V \to R$ has at most one minimum and that minimum is strict.

(3) Let $J : \Omega \subset V \to R$ be a convex fn defined on an open set Ω that contains U and suppose it is diff'ble. at some $u \in U$. Then,

J has a minimum at u with respect to U

$$\Updownarrow$$

$$J'(u)(v-u) \geq 0, \qquad \forall v \in U.$$

(4) If U is open, the preceding condition is equivalent to Euler's equation $J'(u)=0$; i.e.

proof. (1) Let $v = u + w \in U$ be given. By convexity of J

$$J(u+\theta w) = J(\theta v + (1-\theta)u) \leq (1-\theta)J(u) + \theta J(v), \quad 0 \leq \theta \leq 1$$

$\Rightarrow$

$$J(u+\theta w) - J(u) \leq \theta[J(v) - J(u)], \qquad 0 \leq \theta \leq 1.$$

Since u is a relative minimum, $\exists\ \theta_0 > 0$ such that

$$J(u+\theta_0 w) - J(u) \geq 0,$$

which shows that $J(v) \geq J(u)$.

(2) Reasoning as in (1), we obtain for $v \neq u$

$$0 \leq J(u+\theta_0 w) - J(u) < \theta_0[J(v) - J(u)]$$

which establishes that the minimum is strict and unique.

(3) In Theorem 7.4-1 it was shown that $J'(u)(v-u) \geq 0 \quad \forall v \in U$ is necessary for a relative minimum ~~without~~ in fact using the convexity of J. The same proof also works for a minimum. To prove the sufficiency, observe that, from Theorem 7.4-2, since J is convex,

Thm. 7.4-2

$$\circledast \quad J(v) - J(u) \geq J'(u)(v-u) \geq 0 \qquad \forall v \in U,$$

$$\Rightarrow J(v) \geq J(u) \qquad \forall v \in U.$$

(Note that in Thm. 7.4-2 J was diff'ble. on Ω. We really don't need that since $\circledast$ ~~was~~ is needed only at u).

(4) Immediate. $\square$

Example Consider the following minimization problem: (least squares) $B \in R^{m\times n}$ and $c \in R^m$ be given, find $u \in R^n$ such that

$$\|Bu - c\|_m = \inf_{v\in R^n} \|Bv - c\|_m \qquad (LS)$$

where $\|\cdot\|_m$ denotes the Euclidean norm in R^m. Introduce the quadratic functional $J: R^n \to R$,

$$J(v) \equiv \frac{1}{2}\|Bv - c\|_m^2 - \frac{1}{2}\|c\|_m^2$$

$$= \frac{1}{2}(Bv, Bv)_m - (c, Bv)_m$$

$$= \frac{1}{2}(B^TBv, v)_n - (B^Tc, v)_n$$

where $(\cdot,\cdot)_m$ and $(\cdot,\cdot)_n$ denote the scalar products in R^m and R^n respectively.

It is easy to show that B^TB is nonnegative definite. Thus, by Theorem 7.4-3 and a previous example, J is convex on R^n.

It is clear that the original problem (LS) is equivalent to

$$J(u) = \inf_{v\in R^n} J(v).$$

Thus, by Theorem 7.4-4, the set of solutions of (LS) coincides with the set of solutions of

$$J'(u) = B^TBu - B^Tc = 0,$$

the so-called normal equations.

If $m > n$ and B has full rank, then B^TB is positive definite and the (LS) pb. has a unique solution. In this case, J is strictly convex on R^n.

If $m = n$ and B has full rank, then

$\inf_{v\in R^n} \|Bv - c\|_m = 0.$