REAL ANALYSIS
FUNCTIONS OF SEVERAL VARIABLES

We denote $(x_1 \dots x_2)$ by X and $f(x_1 \dots x_n)$ by $f(X)$. We may think of X as a vector or a point.

If $A = (a_1 \dots a_n)$ $B = (b_1 \dots B - n)$ then

$A - B = (a_1 - b_1 \dots a_n - b_n)$

$A + B = (a_1 + b_1 \dots a_n + b_n)$

$A.B = a_1 b_1 + \dots + a_n b_n$ - a scalar

$||A|| = \sqrt{a_1^2 + \dots + a_n^2}$ norm of A

$||A - B|| = \sqrt{(1_a - a_b)^2 + \dots + (a_n - b_n)^2}$ is the distance AB

$|A.B| \leq ||A||\ ||B||$ Cauchy's inequality.

Suppose we have m functions of n variables ${}^{(1)}f(X),{}^{(2)}f(X)\dots{}^{(m)}F(X)$.

We shall denote by the vector function

$$F(X) = ({}^{(1)}f(x_1 \dots x_n), \dots {}^{(m)}f(x_1 \dots x_n))$$

Theorem 1 If $f(X)$ $g(X)$ are continuous at A relative to S then so are $f(X) \pm g(X), f(X)g(X)$ and, if $g(A) \neq 0$, $\frac{f(X)}{g(X)}$.

Theorem 2 Suppose that the components ${}^{(1)}f(X)\dots{}^{(m)}f(X)$ of the vector function $F(X)$ are continuous at A relative to S. Let $B = F(A)$ and let T be the set of all points $F(X)$ with X in S. Then if $g(Y) = g(y_1 \dots y_m)$ is continuous at B relative to T, it follows that $g(F(x))$ is continuous at A relative to S.

Differentiability $f(X)$ is differentiable at $X{+}A \Leftrightarrow \exists$ a vector $G|\frac{f(X)-f(A)-G(X-A)}{||X-A||} \rightarrow 0$ as $X \rightarrow A$.

If f is differentiable then $\frac{\partial f}{\partial x_1} \dots \frac{\partial f}{\partial x_n}$ all exist and the vector G is $\left(\frac{\partial f}{\partial x_1} \dots \frac{\partial f}{\partial x_n}\right)$.

We call this $(grad\ f(X))_{X=A}$ or $(\nabla f)_{X=A}$.

Thus $f(X)$ is differentiable $\Leftrightarrow \frac{\delta f - \nabla f \delta X}{||\delta X||} \rightarrow 0$ as $\delta X \rightarrow 0$.

Theorem 3 If $\frac{\partial f}{\partial x_1}, \dots \frac{\partial f}{\partial f}\partial x_n$ are continuous at $X = A$, then $f(X)$ is differentiable at $X = A$.

Proof Suppose $H \neq 0$ and $||H||$ is sufficiently small. Consider

$$\begin{aligned}
& \frac{1}{||H||}\left|\sum_{r=1}^{n}\{f(a_1+h_1\ldots,a_r+h_r,a_{r+1}\ldots a_n)\right. \\
& \left.-f(a_1+h_1\ldots a_{r-1}+h_{r-1}a_r\ldots a_n)-h_rf_r(A)\}\right| \\
& (\text{This is } \frac{1}{||H||}\{f(A+h)-f(A)-A\nabla f\}) \\
\leq\ & \frac{1}{||H||}\left|\sum_{r=1}^{n}[\{f_r(a_1+h_1\ldots a_{r-1}+h_{r-1}a_r+\theta_r f_{r_1}a_{r+1}-a_n)-f_r(A)\}\,h_r]\right| \\
& 0<\theta_r<1
\end{aligned}$$

Let V be the vector with components

$$f_r(a_1+h_1,\ldots a_r\theta_f h_r,a_{r+1}\ldots a_n)-f_rA\ (r=1,2,\ldots n)$$

Each component can be made as small as we please, provided only that $||H||$ is sufficiently small since $f_r(X)$ is continuous at $X+A$. Hence we can make $||V||<\varepsilon$ if $||H||$ is sufficiently small. The above inequality then gives

$$\frac{1}{||H||}\left|f(A+h)-f(A)-\sum_{r=1}^{n}h_rf_rA\right|\leq\frac{|V.H|}{||H||}\leq||V||<\varepsilon.$$

Theorem 4 If $f(X)$ and $g(X)$ are both differentiable at $X+A$, then so are $f(X)\pm g(X),\ f(X).g(X)$ and, provided $g(A)|neq0,\ \frac{f(X)}{g(X)}$.

$$\left.\begin{array}{c}\nabla(f\pm g)=\nabla f\pm\nabla g\\ \nabla(fg)=f\nabla g+g\nabla f\\ \nabla\frac{f}{g}=\frac{1}{g}\nabla f-\frac{f}{g^2}\nabla g\end{array}\right\}\ \text{at}X=A$$

Proof of (iii) Take $f\equiv 1$ and suppose $g(A)\neq 0$. Consider

$$\begin{aligned}
& \frac{1}{||H||}\left|\frac{1}{g(A+h)}-\frac{1}{g(A)}-\left\{\frac{\nabla g_A}{g^2(A)}\right\}.H\right| \\
=\ & \frac{1}{||H||}\left|\frac{g(A)\{g(A)-g(A+H)+\nabla g_A.H\}+\{g(A+H)-g(A)\}\{\nabla g_A.H\}}{g^2(A)g(A+H)}\right| \\
\leq\ & \left|\frac{g(A+h)-g(A)-\nabla g.H}{||H||}\right|\frac{1}{g(A)g(A+H)}+\frac{|g(A+H)-g(A)|}{|g^2(A).g(A+H)|}||\nabla g||
\end{aligned}$$

using $|\nabla g.H$
$leq||\nabla g||\ ||H|| \to 0$ as $||H|| \to 0$.

Theorem 5 Function of a function rule.

Let ${}^{(1)}f(X), {}^{(2)}f(X), \ldots {}^{(n)}f(X)$ be differentiable at $X = A$. Let $g(Y) = g(y_1 \ldots y_m)$ be differentiable at $Y = B$ where $B = F(A) = ({}^{(1)}f(A) \ldots {}^{(m)}f(A))$.

Then $h(X) = g(F(x))$ is differentiable at $X + A$ and

$$\begin{pmatrix} h_1(x) \\ \vdots \\ h_n(X) \end{pmatrix}_{X=A} = \begin{pmatrix} {}^{(1)}f_1(X) \ldots {}^{(m)}f_1(X) \\ \\ {}^{1}f_n(X) \ldots {}^{m}f_n(X) \end{pmatrix}_{X=A} \begin{pmatrix} g_1(Y) \\ \vdots \\ g_m(Y) \end{pmatrix}_{Y=B}$$

Proof Let RHS of the expression be D. Let $(g_1(Y) \ldots g_m(Y))^T = G'$.

We have the following results

1. Since $f(X)$ is differentiable at $X + A$ $\frac{F(A+H)-f(A)}{||H||}$ is bounded for $0 < ||H|| < \delta$.
   Thus, if each component of $F(X)$ is differentiable at $X = A$ $\frac{||F(A+H)-F(A)||}{||H||}$ is bounded in $0 < ||H|| < \delta$
2. Since $g(Y)$ is differentiable at $Y + B$

$$g(B + \Omega) = g(B) - G'.\Omega = \varepsilon(\omega)||\Omega||$$

   where $\varepsilon(\Omega) \to 0$ as $||\omega||to0$.

Consider

$$\begin{aligned} & \frac{1}{||H||}|g(F(A+H)) - g(F(A)) - D.H| \\ = & \frac{1}{||H||}|g(F(A+H)) - g(F(A)) - G'.(F(A+H) - F(A)) \\ & + G'.(F(A+H) - F(A)) - D.H| \\ \leq & \frac{1}{||H||}|g(B+\Omega) - g(B) - G'.\Omega| + \frac{1}{||H||}|G'.(F(A+H) - F(A)) - D.H| \end{aligned}$$

(writing $F(A) + B$ $F(A+H) - F(A) = \Omega$.

First term $= \varepsilon(\Omega)\frac{||\Omega||}{||H||}$ by (2) where $\varepsilon \to 0$ as $\Omega \to 0$.

By (1) $\frac{||\Omega||}{||H||}$ is bounded for $0 < ||H|| < \delta$ also since $\Omega \to 0$ as $H \to 0$, $\varepsilon(\Omega) \to 0$ as $H \to 0$.

Second term

$$
\begin{aligned}
&= \frac{1}{||H||}\left|\sum_{r=1}^{m} g_r(B)\left\{{}^{(r)}f(A+H) -^{(r)} f(A) - H.\nabla^r f(A)\right\}\right| \\
&\leq \sum_{r=1}^{n} |g_r(B)|\frac{1}{||H||}\left|{}^{(r)}f(A+H) -^{(r)} f(A) - H.\nabla^r f(A)\right|
\end{aligned}
$$

$\to 0$ as $||H|||to0$.

Hence $\frac{1}{||H||}\,|g(F(A+H)) - g(F(A) - D.H| \to 0$ as $||H|||to0$. Hence the result.

Corollary In The special case when $n = 1$ we get, when $h(x) = g(F(X))$ that

$$h'(x) = F'(a).(\nabla g)_B$$

Theorem 6 First Mean Value Theorem Suppose $d(X)$ is differentiable at all points of the open line segment $(A, A+H)$ and continuous on the closed segment. Then for some

$$g(A+H) - g(A) = H.\nabla g(A+\theta H)$$

Proof Suppose $0 < t_0 < 1$. Then $h(t) = g(A+tH)$ is differentiable at $t = t_0$, since $g(X)$ is differentiable at $X = A + t_0 H$ and so are $a_r + t_0 h_r$ $r = 1, \ldots, n$.

Furthermore

$$
\begin{aligned}
h'(t) &= \left\{\frac{d}{dt}A + tH\right\}.\nabla g(A+tH) \text{ at } t = t_0 \\
&= H.\nabla g(A+tH)\ t = t_0
\end{aligned}
$$

And $h(t)$ is continuous for t in [01] hence by MVT $h(1) - h(0) - h'(\theta$ for some $\theta|0 < \theta < 1$. Hence

$$g(A+H) - g(A) = H.\nabla g(A+\theta H)$$

Theorem 7 Taylor's Theorem Suppose that the function $f(X)$ is such that all its partial derivatives of (total) order $u-1$ are continuous on the closed line segment $[A, A+H]$, and differentiable on the open line segment $(A, A+H)$, then for some θ with $0 < \theta < 1$, we have

$$\begin{aligned} f(A+H) &= \left\{\sum_{r=0}^{u-1} \frac{1}{r!}\Omega^r f(X)\right\}_{X=A} + \left\{\frac{1}{u!}\Omega^U f(X)\right\}_{X=A+\theta H} \\ \Omega &= H.\nabla = h_1\frac{\partial}{\partial x_1} + \ldots + h_n\frac{\partial}{\partial x_n} \end{aligned}$$

Proof Write $h(t) = f(A+tH)$

By induction on r we have

$$\frac{d^r}{dt^r}h(t) = [\Omega^r f(X)]_{X=A+tH} \begin{cases} r = 0, 1, \ldots u-1 \\ 0 \leq t \leq 1 \\ r = u \ 0 < t < 1 \end{cases}$$

By Taylor's theorem for a function of one variable

$$h(1) = \sum_{r=0}^{u-1} \frac{1}{r!}h^{(r)}(0) + \frac{1}{U!}h^{(n)}\theta \ 0 < \theta < 1$$

Hence the result.

Maxima and Minima $f(X)$ has a strict maximum at $X = A$ means $\exists \varepsilon > 0|$ in $0 < (X - A| < \varepsilon$ we have $f(A) > f(X)$.

By a weak minimum we mean that in $0 < |X_A| < \varepsilon| \ f(A) \geq f(X)$.

Theorem 8 Suppose $f(X)$ has a maximum or a minimum at $X+A$. If $f(X)$ has first order partial derivatives at $X = A$ then $(\nabla f)_A = 0$ i.e. $f_r(A) = 0 \ r = 1, 2, \ldots, n$. If $f(X)$ has second order derivatives continuous in a neighbourhood of A, then the quadratic form $\sum_{ij} h_i h_j f_{ij}(A)$ in $h_1 \ldots h_m$ is negative or positive semi-definite.

Proof Consider $f(x_1, x_2 \ldots a_n) = \phi(x_1)$. This, as a function of x_1, has a maximum or minimum at $x_1 = a_1$ therefore by the theorem for a function of one variable $\phi'(x_1) = \frac{\partial f}{\partial ; x_1} = 0$. Similarly for other variables therefore $\nabla f_A = 0$.

Suppose that the quadratic form is not semi-definite. Then $\exists U = (u_1 \ldots u_n)$ such that

$$\sum_{ij} f_{ij} u_i u_j > 0$$

and $V = (v_1 \dots v_n)$ such that

$$\sum_{ij} f_{ij} A v_i v_j < 0$$

Let $H^1 = (u_1 h \dots u_n h)$.

Using Taylor's Theorem

$$f(A + H@) = f(A) + \frac{h^2}{2!} \sum_{ij} f_{ij}(A + \theta^1 H^1) u_i u_j$$

The linear terms vanishing as $\nabla f = 0$

Let $H^2 = (v_1 h \dots v_n h)$

$$f(A + H^2) = f(A) + \frac{h^2}{2!} \sum_{ij} f_{ij}(A + \theta^1 H^2) v_i v_j$$

Since the second derivatives are continuous, $\exists \delta > 0|$ for $0 < h < \delta$

$$\begin{aligned} f(A + H^1) &= f(A) + (\phi^1)^2 \\ f(A + H^2) &= f(A) - (\phi^2)^2 \end{aligned}$$

We get neither a maximum nor a minimum since in any sphere of radius ε about A, we can choose h such that $|H^1| < \varepsilon$ $|H^2| < \varepsilon$.

item[Theorem 9] Suppose that $f(x)$ has second derivatives which are continuous in the neighbourhood of A. Suppose that $\nabla f_A = 0$ and that the quadratic form $\sum_{ij} h_i h_j f_{ij}(A)$ is negative/positive definite in $h_1 \dots h_n$. Then $f(X)$ has maximum/minimum at $X = A$.

Proof Given $\sum_{ij} f_{ij}(A) h_i h_j$ positive definite, and the second derivatives are continuous. Then $\exists \delta > 0$ for each X satisfying $|X - A| < \delta$ the quadratic form $\sum_{ij} f_{ij}(X) h_i h_j$ is also a positive definite form.

[Using the determinant test, as all the determinants are continuous functions of X.]

Using Taylor's Theorem,

$$f(X) = f(A) + \frac{1}{2} \sum_{Ij} f_{ij}(A + \theta H) h_i h_j > f(A)$$

since the quadratic form is positive definite at each point in the sphere.