SCHOOL OF MATHEMATICS, D.A.V.V., Indore.

Some useful results on functions of several variables

Book: Chapter 9: Principles of Mathematical Analysis, Walter Rudin

9.7 **Theorem:**

(a) If $A \in L(\mathbb{R}^n, \mathbb{R}^m)$, then $\|A\| < \infty$ and A is a uniformly continuous mapping of $\mathbb{R}^n$ into $\mathbb{R}^m$.

(b) If $A, B \in L(\mathbb{R}^n, \mathbb{R}^m)$ and c is a scalar, then

$$\|A + B\| \leq \|A\| + \|B\|, \qquad \|cA\| = |c| \, \|A\|.$$

With the distance between A and B defined as $\|A - B\|$, $L(\mathbb{R}^n, \mathbb{R}^m)$ is a metric space.

(c) If $A \in L(\mathbb{R}^n, \mathbb{R}^m)$ and $B \in L(\mathbb{R}^m, \mathbb{R}^k)$, then

$$\|BA\| \leq \|B\| \, \|A\|.$$

9.8 **Theorem:**

Let Ω be the set of all invertible linear operators on $\mathbb{R}^n$.

(a) If $A \in \Omega$, $B \in L(\mathbb{R}^n)$, and

$$\|B - A\| \, \|A^{-1}\| < 1,$$

then $B \in \Omega$.

(b) Ω is an open subset of $L(\mathbb{R}^n)$, and the mapping $A \to A^{-1}$ is continuous on Ω.

(This mapping is also obviously a one-one mapping of Ω onto Ω, which is its own inverse.)

9.9 **Theorem:** If S is a metric space, if $a_{11}, \ldots, a_{mn}$ are real continuous functions on S, and if, for each $p \in S$, A_p is the linear transformation of $\mathbb{R}^n$ into $\mathbb{R}^m$ whose matrix has entries $a_{ij}(p)$, then the mapping $p \to A_p$ is a continuous mapping of S into $L(\mathbb{R}^n, \mathbb{R}^m)$.

9.10 **Differentiation**

If f is a real function with domain $(a, b) \subset \mathbb{R}$ and if $x \in (a, b)$, then $f'(x)$ is usually defined to be the real number

$$\lim_{h \to 0} \frac{f(x + h) - f(x)}{h}, \tag{1}$$

provided, of course, that limit exists. Thus

$$f(x+h) - f(x) = f'(x)h + r(h) \tag{2}$$

where the "remainder" $r(h)$ is small, in the sense that

$$\lim_{h \to 0} \frac{r(h)}{h} = 0. \tag{3}$$

Note that equation (2) expresses the difference $f(x+h) - f(x)$ as the sum of the linear function that takes h to $f'(x)h$, plus a small remainder.

We can therefore regard the derivative of f at x, not as a real number, but as the linear operator on $\mathbb{R}$ that takes h to $f'(x)h$.

Observe that every real number α gives rise to a linear operator on $\mathbb{R}$, simply a multiplication by α. Conversely, every linear function that carries $\mathbb{R}$ to $\mathbb{R}$ is multiplication by some real number. It is a natural correspondence between $\mathbb{R}$ and $L(\mathbb{R})$ which motivates the preceding statements.

Let us next consider a function $\mathbf{f}$ that maps $(a, b) \subset \mathbb{R}$ into $\mathbb{R}^m$. In that case, $\mathbf{f}'(x)$ is defined to be the vector $\mathbf{y} \in \mathbb{R}^m$ (if there is one) for which

$$\lim_{h \to 0} \left\{ \frac{\mathbf{f}(x+h) - \mathbf{f}(x)}{h} - \mathbf{y} \right\} = 0 \tag{4}$$

We can again rewrite this in the form

$$\mathbf{f}(x+h) - \mathbf{f}(x) = h\mathbf{y} + \mathbf{r}(h), \tag{5}$$

where $\mathbf{r}(h)/h \to \mathbf{0}$ as $h \to 0$. The main term on the right side of (5) is again a linear function of h.

Note that every $\mathbf{y} \in \mathbb{R}^m$ induces a linear transformation of $\mathbb{R}$ into $\mathbb{R}^m$, by associating to each $h \in \mathbb{R}$ the vector $h\mathbf{y} \in \mathbb{R}^m$. This identification of $\mathbb{R}^m$ with $L(\mathbb{R}, \mathbb{R}^m)$ allows us to regard $\mathbf{f}'(x)$ as a member of $L(\mathbb{R}, \mathbb{R}^m)$.

Thus, if $\mathbf{f}$ is a differentiable mapping of $(a,b) \subset \mathbb{R}$ into $\mathbb{R}^m$, and if $x \in (a,b)$, then $\mathbf{f}'(x)$ is the linear transformation of $\mathbb{R}$ into $\mathbb{R}^m$ that satisfies

$$\lim_{h \to 0} \frac{\mathbf{f}(x+h) - \mathbf{f}(x) - \mathbf{f}'(x)h}{h} = \mathbf{0} \tag{6}$$

or equivalently,

$$\lim_{h \to 0} \frac{|\mathbf{f}(x+h) - \mathbf{f}(x) - \mathbf{f}'(x)h|}{|h|} = 0 \tag{7}$$

Now we define differentiation for the case $n > 1$.

9.11 **Definition:**

Suppose E is an open set in $\mathbb{R}^n$, $\mathbf{f}$ maps E into $\mathbb{R}^m$, and $x \in E$. If there exists a linear transformation A of $\mathbb{R}^n$ into $\mathbb{R}^m$ such that

$$\lim_{h \to 0} \frac{|\mathbf{f}(x+h) - \mathbf{f}(x) - Ah|}{|h|} = 0 \tag{8}$$

then we say that $\mathbf{f}$ is *differentiable* at x, and we write

$$\mathbf{f}'(x) = A \tag{9}$$

Note:

Note that in (8) $h \in \mathbb{R}^n$. If $|h|$ is small enough, then $x + h \in E$, since E is open. Thus $\mathbf{f}(x+h)$ is defined, $\mathbf{f}(x+h) \in \mathbb{R}^m$, and since $A \in L(\mathbb{R}^n, \mathbb{R}^m)$, $Ah \in \mathbb{R}^m$. Thus

$$\mathbf{f}(x+h) - \mathbf{f}(x) - Ah \in \mathbb{R}^m$$

The norm in the numerator of (8) is that of $\mathbb{R}^m$. In denominator we have the $\mathbb{R}^n$ norm of h

9.12 **Theorem:**

Suppose E and $\mathbf{f}$ are as in the definition 9.11, $x \in E$, and (8) holds with $A = A_1$ and with $A = A_2$. Then $A_1 = A_2$.

9.13 **Remarks:**

(a) The relation (8) can be rewritten in the form

$$\mathbf{f}(x+h)-\mathbf{f}(x)=\mathbf{f}'(x)h+\mathbf{r}(h) \tag{10}$$

where the remainder $\mathbf{r}(h)$ satisfies

$$\lim_{h\to 0}\frac{|\mathbf{r}(h)|}{|h|}=0 \tag{11}$$

we may interpret (10) by saying that for fixed x and small h, the left side of (10) is approximately equal to $\mathbf{f}'(x)h$, that is, to the value of a linear transformation applied to h.

(b) $\mathbf{f}'(x)$ is then a function, namely, a linear transformation of $\mathbb{R}^n$ into $\mathbb{R}^m$. But $\mathbf{f}'$ is also a function: $\mathbf{f}'$ maps E into $L(\mathbb{R}^n,\mathbb{R}^m)$.

(c) The equation (10) shows that $\mathbf{f}$ is continuous at any point at which $\mathbf{f}$ is differentiable.

(d) The derivative defined by (8) or by (10) is called the *differential* of $\mathbf{f}$ at x, or the *total derivative* of $\mathbf{f}$ at x.

9.14

If $A\in L(\mathbb{R}^n,\mathbb{R}^m)$ then $\mathbf{A}'(x)=A$.

9.15 **Theorem:(Chain rule)**

Suppose E is an open set in $\mathbb{R}^n$, $\mathbf{f}$ maps E into $\mathbb{R}^m$, $\mathbf{f}$ is differentiable at $x_0\in E$, $\mathbf{g}$ maps an open set containing $\mathbf{f}(E)$ into $\mathbb{R}^k$, and $\mathbf{g}$ is differentiable at $\mathbf{f}(x_0)$. Then the mapping $\mathbf{F}$ of E into $\mathbb{R}^k$ defined by

$$\mathbf{F}(x)=\mathbf{g}(\mathbf{f}(x))$$

is differentiable at x_0, and

$$\mathbf{F}'(x_0)=\mathbf{g}'(\mathbf{f}(x_0))\mathbf{f}'(x_0).$$

Note that on the right hand side we have the product of two linear transformations (matrices).

9.16 **Partial Derivatives**

Consider a function $\mathbf{f}$ that maps an open set $E\subset\mathbb{R}^n$ into $\mathbb{R}^m$. Let $\{e_1,e_2,\ldots,e_n\}$ and $\{u_1,u_2,\ldots,u_m\}$ be the standard bases of $\mathbb{R}^n$ and $\mathbb{R}^m$. The components of $\mathbf{f}$ are the real functions $f_1,f_2,\ldots,f_m$ defined by

$$\mathbf{f}(x) = \sum_{i=1}^{m} f_i(x) u_i \tag{12}$$

or equivalently, by $f_i(x) = \mathbf{f}(x) \cdot u_i$, $1 \leq i \leq m$.
For $x \in E$, $1 \leq i \leq m$, $1 \leq j \leq n$, we define

$$(D_j f_i)(x) = \lim_{t \to 0} \frac{f_i(x + te_j) - f_i(x)}{t} \tag{13}$$

provided the limit exists. We see that $D_j f_i$ (called partial derivative) is the derivative of f_i with respect to x_j, keeping the other variables fixed.

9.17 **Theorem:**

Suppose $\mathbf{f}$ maps an open set $E \subset \mathbb{R}^n$ into $\mathbb{R}^m$, and $\mathbf{f}$ is differentiable at a point $x \in E$. Then the partial derivatives $(D_j f_i)(x)$ exist, and

$$\mathbf{f}'(x) e_j = \sum_{i=1}^{m} (D_j f_i)(x) u_i \quad (1 \leq j \leq n) \tag{14}$$

Remark:

$\mathbf{f}'(x) e_j$ is the j^{th} column vector of the matrix that represents $\mathbf{f}'(x)$ i.e., the number $(D_j f_i)(x)$ occupies the spot in the i^{th} row and j^{th} column of the matrix of $\mathbf{f}'(x)$. Thus

$$\mathbf{f}'(x) = \begin{bmatrix} (D_1 f_1)(x) & \cdots & (D_n f_1)(x) \\ \cdots & \cdots & \cdots \\ (D_1 f_m)(x) & \cdots & (D_n f_m)(x) \end{bmatrix}.$$

If $h = \sum h_j e_j$ is any vector in $\mathbb{R}^n$, then (14) implies that

$$\mathbf{f}'(x) h = \sum_{i=1}^{m} \left\{ \sum_{j=1}^{n} (D_j f_i)(x) h_j \right\} u_i. \tag{15}$$

9.18 **Example:**

Let γ be a differentiable mapping of the segment $(a, b) \subset \mathbb{R}$ into an open set $E \subset \mathbb{R}^n$, in other words, γ is a differentiable curve in E. Let f be a real valued differentiable function with domain E. Thus f is a differentiable mapping of E into $\mathbb{R}$. Define

$$g(t) = f(\gamma(t)) \qquad (a < t < b). \tag{16}$$

The chain rule asserts that

$$g'(t) = f'(\gamma(t))\gamma'(t) \qquad (a < t < b). \tag{17}$$

Since $\gamma'(t) \in L(\mathbb{R}, \mathbb{R}^n)$ and $f(\gamma'(t)) \in L(\mathbb{R}^n, \mathbb{R})$, (17) defines $g'(t)$ as a linear operator on $\mathbb{R}$. However $g'(t)$ can also be regarded as a real number. This number can be computed in terms of the partial derivatives of f and the derivatives of the components of γ.

With respect to the standard basis $\{e_1, e_2, \ldots, e_n\}$ of $\mathbb{R}^n$, $\gamma'(t)$ may be viewed as the n by 1 matrix which has $\gamma_i'(t)$ in the ith row, where $\gamma_1, \ldots, \gamma_n$ are the components of γ. For every $x \in E$, $f'(x)$ may be viewed as the 1 by n matrix which has $(D_j f)(x)$ in the jth column. Hence $g'(t)$ is the 1 by 1 matrix whose only entry is the real number

$$g'(t) = \sum_{i=1}^{n} (D_i f)(\gamma(t))\gamma_i'(t) \tag{18}$$

We can look this formula in following manner.
Associate with each $x \in E$ a vector, the so called "gradient" of $\mathbf{f}$ at x, defined by

$$(\nabla \mathbf{f})(x) = \sum_{i=1}^{n} (D_i f)(x) e_i. \tag{19}$$

Since

$$\gamma'(t) = \sum_{i=1}^{n} \gamma_i'(t) e_i, \tag{20}$$

(18) can be written in the form

$$g'(t) = (\nabla \mathbf{f})(\gamma(t)) \cdot \gamma'(t) \tag{21}$$

the scaler product of the vectors $(\nabla \mathbf{f})(\gamma(t))$ and $\gamma'(t)$.

Directional Derivative:

Fix an $x \in E$, let $u \in \mathbb{R}^n$ be a unit vector and specialize γ so that

$$\gamma(t) = x + tu \qquad (-\infty < t < \infty) \tag{22}$$

Then $\gamma'(t) = u$ for every t. Hence (21) shows that

$$g'(0) = (\nabla \mathbf{f})(x) \cdot u \tag{23}$$

On the other hand, (22) shows that

$$g(t) - g(0) = f(x + tu) - f(x)$$

Hence (23) gives

$$\lim_{t \to 0} \frac{f(x + tu) - f(x)}{t} = (\nabla \mathbf{f})(x) \cdot u \tag{24}$$

the limit in (24) is called the *directional derivative* of f at x, in the direction of the unit vector u, and may be denoted by $(D_u f)(x)$.

If f and x are fixed, but u varies, then (24) shows that $(D_u f)(x)$ attains its maximum when u is a positive scalar multiple of $(\nabla \mathbf{f})(x)$.

If $u = \sum u_i e_i$ then (24) shows that $(D_u f)(x)$ can be expressed in terms of the partial derivatives of f at x by the formula

$$(D_u f)(x) = \sum_{i=1}^{n} (D_i f)(x) u_i \tag{25}$$

9.20 **Definition:**

A differentiable mapping $\mathbf{f}$ of an open set $E \subset \mathbb{R}^n$ into $\mathbb{R}^m$ is said to be *continuously differentiable* in E if $\mathbf{f}'$ is a continuous mapping of E into $L(\mathbb{R}^n, \mathbb{R}^m)$.

9.21 **Theorem:**

Suppose $\mathbf{f}$ maps an open set $E \subset \mathbb{R}^n$ into $\mathbb{R}^m$. Then $\mathbf{f} \in \mathscr{C}'(E)$ if and only if the partial derivatives $D_j f_i$ exist and are continuous on E for $1 \leq i \leq m$, $1 \leq j \leq n$.

9.24 **Inverse Function Theorem:**

Suppose $\mathbf{f}$ is a $\mathscr{C}'$-mapping of an open set $E \subset \mathbb{R}^n$ into $\mathbb{R}^n$, $\mathbf{f}'(a)$ is invertible for some $a \in E$, and $b = \mathbf{f}(a)$. Then

(a) there exist open sets U and V in $\mathbb{R}^n$ such that $a \in U$, $b \in V$, $\mathbf{f}$ is one to one on U, and $\mathbf{f}(U) = V$;

(b) if $\mathbf{g}$ is the inverse of $\mathbf{f}$ (which exists , by (a)), defined in V by

$$\mathbf{g}(\mathbf{f}(x)) = x \qquad (x \in U)$$

then $\mathbf{g} \in \mathscr{C}'(V)$.

9.25 **Theorem: (Open Mapping Theorem)**

If $\mathbf{f}$ is a $\mathscr{C}'$-mapping of an open set $E \subset \mathbb{R}^n$ into $\mathbb{R}^n$ and if $\mathbf{f}'(x)$ is invertible for every $x \in E$, then $\mathbf{f}(W)$ is an open subset of $\mathbb{R}^n$ for every open set $W \subset E$.

9.26 **Notation:**

For $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^m$, write

$$(x, y) = (x_1, \ldots, x_n, y_1, \ldots, y_m) \in \mathbb{R}^{n+m}.$$

Every $A \in L(\mathbb{R}^n, \mathbb{R}^m)$ can be split into two linear transformations A_x and A_y, defined by

$$A_x h = A(h, 0), \qquad A_y k = A(0, k)$$

for any $h \in \mathbb{R}^n$, $k \in \mathbb{R}^m$. Then $A_x \in L(\mathbb{R}^n)$, $A_y \in L(\mathbb{R}^m, \mathbb{R}^n)$, and

$$A(h, k) = A_x h + A_y k.$$

9.27 **Linear Version of the Implicit Function Theorem:**

If $A \in L(\mathbb{R}^{n+m}, \mathbb{R}^n)$ and if A_x is invertible, then there correponds to every $k \in \mathbb{R}^m$ a unique $h \in \mathbb{R}^n$ such that $A(h, k) = 0$. This h can be computed from k by the formula

$$h = -(A_x)^{-1} A_y k.$$

9.28 **Implicit Function Theorem:**

Let $\mathbf{f}$ be a $\mathscr{C}'$-mapping of an open set $E \subset \mathbb{R}^{n+m}$ into $\mathbb{R}^n$ such that $\mathbf{f}(a, b) = 0$ for some point $(a, b) \in E$. Put $A = \mathbf{f}'(a, b)$ and assume that A_x is invertible.

Then there exist open sets $U \subset \mathbb{R}^{n+m}$ and $W \subset \mathbb{R}^m$, with $(a, b) \in U$ and $b \in W$, having the following property:

To every $y \in W$ corresponds a unique x such that

$$(x, y) \in U \quad \text{and} \quad \mathbf{f}(x, y) = 0. \tag{26}$$

If this x is defined to be $\mathbf{g}(y)$, then $\mathbf{g}$ is a $\mathscr{C}'$-mapping of W into $\mathbb{R}^n$, $\mathbf{g}(b) = a$,

$$\mathbf{f}(\mathbf{g}(y), y) = 0 \qquad (y \in W) \tag{27}$$

and

$$\mathbf{g}'(b) = -(A_x)^{-1} A_y.$$

Note:

The function $\mathbf{g}$ is implicitly defined by (27). Hence the name of the theorem.