Differentiation in $\mathbb{R}^n$

Let U be an open subset of $\mathbb{R}^n$ and let $f : U \to \mathbb{R}^m$ be a map defined on U with values in $\mathbb{R}^m$.

1. f is differentiable at $p \in U$ if there exists a linear function $L : \mathbb{R}^n \to \mathbb{R}^m$ such that
$$\lim_{x \to p} \frac{\|f(x) - f(p) - L(x-p)\|}{\|x-p\|} = 0.$$
2. If f is differentiable at p, then it is (locally) Lipschitz near p.
3. If the linear function exists at p, then it is unique.
4. If f is differentiable at p, then each of the partial derivatives $\dfrac{\partial f_i}{\partial x_j}(p)$, $1 \le i \le m$ and $1 \le j \le n$, exists and
$$\frac{\partial f_i}{\partial x_j}(p) = \text{ the } i\text{-th compnent of } L(e_j) \quad \forall\, 1 \le i \le m,\ 1 \le j \le n.$$
5. If f is differentiable at p and if u is any element of $\mathbb{R}^n$, then the partial derivative $D_u f(p)$ of f at p with respect to u exists. Moreover,
$$D_u f(p) = Df(p)(u) = Df(p)(\sum_{i=1}^{n} u_i e_i) = \sum_{i=1}^{n} u_i Df(p)(e_i)$$
and it is called the directional derivative of f at p in the direction of u provided that u is a unit vector of $\mathbb{R}^n$.

Properties of the Derivative

Let U be an open subset of $\mathbb{R}^n$, let f, $g : U \to \mathbb{R}^m$ be maps defined on U with values in $\mathbb{R}^m$, and let $\varphi : U \to \mathbb{R}$ a function defined on U.

1. If f and g are differentiable at $p \in U$, and if α, $\beta \in \mathbb{R}$, then $h = \alpha f + \beta g$ is differentiable at p and
$$Dh(p) = \alpha Df(p) + \beta Dg(p).$$
2. If φ and f are differentiable at $p \in U$, then $k = \varphi f : E \to \mathbb{R}^m$ is differentiable at p and
$$Dk(p)(u) = \big[D\varphi(p)(u)\big] f(p) + \varphi(p)\big[Df(p)(u)\big] \quad \text{for } u \in \mathbb{R}^n.$$
3. If f and g are differentiable at $p \in U$, then $h = \langle f, g\rangle = \displaystyle\sum_{i=1}^{m} f_i g_i : E \to \mathbb{R}$ is differentiable at p and
$$Dh(p) = g(p) \cdot Df(p) + f(p) \cdot Dg(p) = \sum_{i=1}^{m} Df_i(p) g_i(p) + \sum_{i=1}^{m} f_i(p) Dg_i(p).$$

Chain Rule Let f have domain $E \subseteq \mathbb{R}^n$ and range in $\mathbb{R}^m$, and let g have domain $B \subseteq \mathbb{R}^m$ and range in $\mathbb{R}^\ell$. Suppose that f is differentiable at p and that g is differentiable at $q = f(p)$. Then the composition $h = g \circ f$ is differentiable at p and
$$Dh(p) = Dg(q) \circ Df(p).$$
Alternatively, we write
$$D(g \circ f)(p) = Dg(f(p)) \circ Df(p).$$

Mean Value Theorems

1. Let f be defined on an open subset U of $\mathbb{R}^n$ and have values in $\mathbb{R}$. Suppose that the set U contains the points a, b and the line segment S joining them and that f is differentiable at every point of this segment. Then there exists a point c on S such that

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

2. Let $U \subseteq \mathbb{R}^n$ be an open set and let $f : U \to \mathbb{R}^m$. Suppose that the set U contains the points a, b and the line segment S joining them and that f is differentiable at every point of this segment. Then there exists a point c on S such that

$$\|f(b) - f(a)\| \leq \|Df(c)(b - a)\|.$$

Higher Derivatives

Let U be an open subset in $\mathbb{R}^n$ and let $f : U \to \mathbb{R}$ be a differentiable function on U.

1. If $Df : U \to \mathbb{R}^n$ is differentiable, we denote the derivative of Df by $D^2f : U \to \mathbb{R}^{n\times n}$, and we shall refer the n^2 functions (entries) of D^2f as the second partial derivatives of f denoted by

$$D_{ij}f \quad \text{or} \quad \frac{\partial^2 f}{\partial x_i\, \partial x_j}, \quad i,\, j = 1,\, 2, \ldots, n.$$

   Similarly, if $D^2f : U \to \mathbb{R}^{n^2}$ is differentiable, we denote the derivative of D^2f by $D^3f : U \to \mathbb{R}^{n^3}$, and we shall refer the n^3 functions (entries) of D^3f as the third partial derivatives of f denoted by

$$D_{ijk}f \quad \text{or} \quad \frac{\partial^3 f}{\partial x_i\, \partial x_j\, \partial x_k}, \quad i,\, j,\, k = 1,\, 2, \ldots, n.$$

2. Note that if f is differentiable at a point $p \in U$, then $Df(p)$ is a linear map on $\mathbb{R}^n$ to $\mathbb{R}$ defined by

$$Df(p)(z) = \sum_{i=1}^{n} D_i f(p) z_i = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(p) z_i \quad \text{for } z = (z_1, \ldots, z_n) \in \mathbb{R}^n.$$

   If Df is differentiable at a point $p \in U$, then $D^2f(p)$ is a linear map on $\mathbb{R}^{n^2}$ to $\mathbb{R}$ defined by

$$D^2f(p)(y,\, z) = \sum_{i,\,j=1}^{n} D_{ij}f(p) y_j z_i = \sum_{i,\,j=1}^{n} \frac{\partial^2 f}{\partial x_i\, \partial x_j}(p) y_j z_i \quad \text{for } y,\, z \in \mathbb{R}^n.$$

   Similarly, if D^2f is differentiable at a point $p \in U$, then $D^3f(p)$ is a linear map on $\mathbb{R}^{n^3}$ to $\mathbb{R}$ defined by

$$D^3f(p)(y,\, z,\, w) = \sum_{i,\,j,\,k=1}^{n} D_{ijk}f(p) y_k z_j w_i = \sum_{i,\,j,\,k=1}^{n} \frac{\partial^3 f}{\partial x_i\, \partial x_j\, \partial x_k}(p) y_k z_j w_i \quad \text{for } y,\, z,\, w \in \mathbb{R}^n.$$

3. **Taylor's Theorem** Suppose that f has continuous partial derivatives of order m in a neighborhood of every point of a line segment S joining two point a, $b = a + u$ in U. Then there exists a point c on S such that

$$f(a+u) = f(a) + \sum_{k=1}^{m-1} \frac{1}{k!} D^k f(a)(u)^k + \frac{1}{m!} D^m f(c)(u)^m$$

where $D^k f(p)(u)^k = D^k f(p)(u, u, \ldots, u)$ for $k = 1, \ldots, m$.

The Class $C^1(U)$

Let U be an open subset in $\mathbb{R}^n$, let $f : U \to \mathbb{R}^m$ be differentiable on U and let $D_j f_i = \dfrac{\partial f_i}{\partial x_j}$.

1. A function $L : \mathbb{R}^n \to \mathbb{R}^m$ is said to be linear if

$$L(ax + by) = aL(x) + bL(y) \text{ for all } a,\, b \in \mathbb{R} \text{ and } x,\, y \in \mathbb{R}^n.$$

   Note that if $L : \mathbb{R}^n \to \mathbb{R}^m$ is linear, then $L(0) = 0 \in \mathbb{R}^m$.
2. Let $\mathscr{L}(\mathbb{R}^n, \mathbb{R}^m) = \{L \mid L : \mathbb{R}^n \to \mathbb{R}^m \text{ is linear}\}$. Define

$$(*) \quad \|L\|_{nm} = \sup_{z \in \mathbb{R}^n,\, \|z\| \le 1} \|L(z)\| \quad \text{for each } L \in \mathscr{L}(\mathbb{R}^n, \mathbb{R}^m).$$

   Since $L \in \mathscr{L}(\mathbb{R}^n, \mathbb{R}^m)$ is linear,

$$\|L\|_{nm} = \sup_{z \in \mathbb{R}^n,\, \|z\| \le 1} \|L(z)\| = \sup_{z \in \mathbb{R}^n,\, \|z\| = 1} \|L(z)\|.$$

   If $L \in \mathscr{L}(\mathbb{R}^n, \mathbb{R}^m)$ is represented by an $m \times n$ matrix, then

$$\|L\|_{nm} = \sup_{z \in \mathbb{R}^n,\, \|z\| = 1} \|L(z)\| = \text{the maximum length of column vectors of } L$$

   and the derivative of L equals itself, i.e. $DL = L$.
3. $f \in C^1(U)$ if the derivative $Df(x)$ exists and is continuous under the norm $(*)$ for all $x \in U$, i.e.

$$\lim_{y \in U,\, y \to x} \|Df(y) - Df(x)\|_{nm} = 0 \quad \text{for all } x \in U.$$

4. For all x, $y \in U$,

$$\|Df(x) - Df(y)\|_{nm} \le \left\{ \sum_{i=1}^{m} \sum_{j=1}^{n} |D_j f_i(x) - D_j f_i(y)|^2 \right\}^{1/2} \le \sqrt{m} \|Df(x) - Df(y)\|_{nm}.$$

5. $f \in C^1(U)$ if and only if $D_j f_i$ is continuous on U for each $1 \le i \le m$, $1 \le j \le n$.
6. If $f \in C^1(U)$, then $\det Df = \det[D_j f_i] : U \to \mathbb{R}$ is continuous on U.
7. Suppose that U contains the points a, b and the line segment S joining them, and let $p \in U$. Then we have

$$\|f(b) - f(a) - Df(p)(b - a)\| \le \|b - a\| \sup_{x \in S} \{\|Df(x) - Df(p)\|_{nm}\}.$$

8. If $f \in C^1(U)$, $p \in U$ and $\varepsilon > 0$, then there exists $\delta(\varepsilon) > 0$ such that if $\|x_k - p\| \leq \delta(\varepsilon)$, $k = 1, 2$, then $x_k \in U$ and

$$\|f(x_1) - f(x_2) - Df(p)(x_1 - x_2)\| \leq \varepsilon \|x_1 - x_2\|.$$

9. If $L = Df(p)$ is an injection for some $p \in U$, then there exists $r > 0$ such that

$$(\dagger) \quad r\|u\| \leq \|Df(p)(u)\| = \|L(u)\| \quad \text{for all } u \in \mathbb{R}^n.$$

10. If $f \in C^1(U)$ and if $L = Df(p)$ is an injection for some $p \in U$, then there exists a $\delta > 0$ such that

$$(\dagger\dagger)\tfrac{1}{2}r\,\|x_1 - x_2\| \leq \|f(x_1) - f(x_2)\| \quad \text{for all } x_1,\, x_2 \in \bar{B}_\delta(p) \subset U.$$

This prove that the restriction of f to $\bar{B}_\delta(p)$ is an injection; hence this restriction has an inverse function which we shall denote by $g = (f_{\bar{B}_\delta(p)})^{-1}$. Note that g is uniformly continuous on $f(\bar{B}_\delta(p))$ by using $(\dagger\dagger)$.

11. If $f \in C^1(U)$ and if $L = Df(p)$ is a surjection for some $p \in U$, then there exists a bounded linear function $M : \mathbb{R}^m \to \mathbb{R}^n$ such that $L \circ M(y) = y$ for all $y \in \mathbb{R}^m$.
12. If $f \in C^1(U)$ and if $L = Df(p)$ is a surjection for some $p \in U$, then there exists numbers $d > 0$ and $\alpha > 0$ such that if $y \in \bar{B}_{\alpha/2d}(f(p)) \subset \mathbb{R}^m$, then there exists an $x \in U \cap \bar{B}_\alpha(p)$ such that $f(x) = y$.
13. If $f \in C^1(U)$ and if $L = Df(p)$ is a surjection for all $p \in U$, then $f(U)$ is open in $\mathbb{R}^m$.
14. If $m = n$, $f \in C^1(U)$ and if $L = Df(p)$ is a bijection for some $p \in U$, then there exist open neighborhoods V and W of p and $f(p)$, respectively, such that $f : V \to W$ has a differentiable inverse $g = f^{-1} : W \to V$ such that

$$Df^{-1}(y) = Dg(y) = [Df(g(y))]^{-1} = [Df(f^{-1}(y))]^{-1} \quad \text{for } y \in W.$$

15. Let W be an open subset of $\mathbb{R}^{n+m}$, $f : W \to \mathbb{R}^m \in C^1(W)$ and let $(a, b) \in W$ be a point at which $f(a, b) = 0$. Suppose that $a \in \mathbb{R}^n$, $b \in \mathbb{R}^m$ and the $m \times m$ matrix

$$D_y f|_{(a,b)} = \left[\frac{\partial f_i}{\partial y_j}(a, b)\right]_{1 \leq i,j \leq m} \quad \text{is invertible.}$$

Then there exist open neighborhoods $A \subset \mathbb{R}^n$ and $B \subset \mathbb{R}^m$ of a and b, respectively, and a unique $g : A \to \mathbb{R}^m \in C^1(A)$ such that

$$b = g(a) \quad \text{and} \quad F(x, g(x)) = 0 \quad \text{for all } x \in A$$

and

$$\begin{aligned} & \{(x, y) \in A \times B \subset W \mid f(x, y) = 0\} \\ = \; & \{(x, g(x)) \mid x \in A\} \\ = \; & \text{the graph of } g \text{ over } A. \end{aligned}$$

Extremum problems

1. Let $U \subseteq \mathbb{R}^n$ be open and let $f : U \to \mathbb{R}$ have continuous second partial derivatives on U. If $p \in U$ is a point of relative minimum [respectively, maximum] of f, then

$$D^2 f(p)(w)^2 = w^t \left[\frac{\partial f_i}{\partial x_j}(p)\right]_{1\le i,j\le n} w = \sum_{i,j=1}^{n} D_{ij} f(p) w_i w_j \ge 0 \quad \text{for all } w \in \mathbb{R}^n$$

[respectively, $D^2 f(p)(w)^2 = \sum_{i,j=1}^{n} D_{ij} f(p) w_i w_j \le 0$ for all $w \in \mathbb{R}^n$].

2. **Lagrange's Theorem** Let $U \subseteq \mathbb{R}^n$ be open and suppose that f and $g : U \to \mathbb{R}$ are real-valued functions in $C^1(U)$. Suppose $p \in U$ is such that $g(p) = 0$ and that there exists an $r > 0$ such that

$$f(x) \le f(p) \quad \big[\text{or } f(x) \ge f(p)\big] \quad \text{for all } x \in B_r(p) \subset U \text{ satisfying } g(x) = 0.$$

Then there exist real numbers μ, λ, not both zero, such that

$$\mu Df(p) = \lambda Dg(p).$$

Moreover, if $Dg(p) \ne 0$, we can take $\mu = 1$.

Riemann Integrability

1. (Riemann Criterion for Integrability) f is Riemann integrable over K, i.e. $\underline{\int_K} f = \overline{\int_K} f$, if and only if for each $\varepsilon > 0$, there exists a partition P_ε of K such that

$$|U(P_\varepsilon, f) - L(P_\varepsilon, f)| < \varepsilon.$$

2. (Cauchy Criterion for Integrability) f is Riemann integrable over K if and only if for each $\varepsilon > 0$, there exists a partition P_ε of K such that if P and Q are refinements of P_ε and $S_P(f, K)$ and $S_Q(f, K)$ are any corresponding Riemann sums, then

$$|S_P(f, K) - S_Q(f, K)| < \varepsilon.$$

Convergence and Differentiation

1. Let $f_n : [a, b] \to \mathbb{C}$ be a sequence of continuously differentiable functions defined on $I = [a, b]$. If $f_n(x_0)$ converges for some $x_0 \in [a, b]$ and if $f_n' : [a, b] \to \mathbb{C}$ converges uniformly on $[a, b]$, then f_n converges uniformly on $[a, b]$ to a function f.
2. **Term-by-Term Differentiation** For each $n \in \mathbb{N}$, let f_n be a real-valued function on $K = [a, b]$ which has a derivative f_n' on K. Suppose that
   (i) $\sum f_n$ converges at $x_0 \in K$,
   (ii) $\sum f_n'$ converges uniformly on K,
   there exists a real-valued function f on K such that
   (a) $\sum f_n$ converges uniformly on K to f,
   (b) f is differentiable on K and $f' = \sum f_n'$ on K.

Convergence and Integration

1. Let $\{f_n\}$ be a sequence of integrable functions that converges uniformly on a closed cell $K \subset \mathbb{R}^p$ to a function f. Then f is integrable and
$$\int_K f = \lim_{n\to\infty} \int_K f_n.$$

2. **Term-by-Term Integration** For each $n \in \mathbb{N}$, let f_n be a real-valued integrable function on $K = [a, b]$. Suppose that the series $\sum f_n$ converges to f uniformly on K.
Then f is integrable on K and
$$\int_K f = \sum_{j=1}^{\infty} \int_K f_n.$$

3. **Bounded Convergence Theorem** Let $\{f_n\}$ be a sequence of integrable functions on a closed cell $K \subset \mathbb{R}^p$. Suppose that there exists $B > 0$ such that $\|f_n(x)\| \leq B$ for all $n \in \mathbb{N}$, $x \in K$. If the function $f(x) = \lim\limits_{n\to\infty} f_n(x)$, $x \in K$, exists and is integrable, then
$$\int_K f = \lim_{n\to\infty} \int_K f_n.$$

4. **Monotone Convergence Theorem** Let $\{f_n\}$ be a monotone sequence of integrable functions on a closed cell $K \subset \mathbb{R}^p$. If the function $f(x) = \lim f_n(x)$, $x \in K$, exists and is integrable, then
$$\int_K f = \lim_{n\to\infty} \int_K f_n.$$