5. Approximations (207)

5.1. Mean value theorem (208)

- **Theorem 5.1.1** (Mean value theorem) Let $U \subseteq \mathbb{R}^n$ be open and let $a, b \in U$. Let f be a differentiable real-valued function on U. If U contains the line segment L from a to b then there exists $c \in L$ such that
$$f(b) - f(a) = \nabla f(c) \cdot (b - a).$$
(208)
- **Theorem 5.1.4** Let $U \subseteq \mathbb{R}^n$ be open and C^1 path-connected. Let $f : U \to \mathbb{R}^m$ be differentiable. The Jacobian $Df(x)$ is the $m \times n$ zero matrix for all $x \in U$ if and only if f is a constant map. (209)
- **Corollary 5.1.5** Let $U \subseteq \mathbb{R}^n$ be open and C^1 path connected. Let $F : U \to \mathbb{R}^m$ and $G : U \to \mathbb{R}^m$ be differentiable. If $DF(x) = DG(x)$ for all $x \in U$, then there exists a constant $C \in \mathbb{R}^m$ such that $F(x) = G(x) + C$ for all $x \in U$. (209)

5.2. Second order derivatives and the Hessian (210)

- ### 5.2.1 Definition (210)
  - **Definition 5.2.1** Let $U \subseteq \mathbb{R}^n$ be open. Let $f : U \to \mathbb{R}^m$ be C^1. Fix $i, j \in \{1, \ldots, n\}$ and $a \in U$. The **second order partial derivative** $\partial_i \partial_j f$ at a is defined by $\partial_i \partial_j f(a) = \partial_i(\partial_j f)(a)$, provided i^{th} partial of $\partial_j f$ exists at a. If $i \neq j$ then the partial is **mixed** and if $i = j$ then the partial is **pure**. (210)
    - $f_{xx} = (f_x)_x = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial x^2}$ (210)
    - $f_{xy} = (f_x)_y = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial y \partial x}$ (210)
- ### 5.2.2 Clairaut's theorem (212)
  - **Definition 5.2.7** Let $U \subseteq \mathbb{R}^n$ be an open set. A function $f : U \to \mathbb{R}^m$ is **twice continuously differentiable** (or C^2) provided for all $i, j \in \{1, \ldots, n\}$, the second partials $\partial_i \partial_j f$ exist and are continuous everywhere in U. (212)
  - **Theorem 5.2.8** (Clairaut) Let $U \subseteq \mathbb{R}^n$ be an open set and let $f : U \to \mathbb{R}^m$. If f is C^2, then
$$\forall i, j \in \{1, \ldots, n\}, \quad \partial_i \partial_j f = \partial_j \partial_i f.$$
(212)
  - **Lemma 5.2.12** Let $U \subseteq \mathbb{R}^n$ be open. If $f : U \to \mathbb{R}^m$ is C^2, then for $h = (h_1, \ldots, h_n) \in \mathbb{R}^n$,
$$H_f^2(x) := \partial_h(\partial_h f)(x) = \sum_{i=1}^{n} h_i^2 (\partial_i^2 f)(x) + \sum_{i=1}^{n} \sum_{j=i+1}^{n} 2h_i h_j (\partial_i \partial_j f)(x).$$
(213)
- ### 5.2.3 Hessian matrix (213)
  - **Definition 5.2.13** Let f be a real-valued function which is C^2 at $a \in \mathbb{R}^n$. The **Hessian of f at a** is the $n \times n$ matrix $Hf(a)$ defined by
$$Hf(a) = \left[\partial_i \partial_j f(a)\right]_{i,j}.$$
(213)
    - By Clairaut's theorem, the Hessian is a symmetric matrix. (213)

5.3. Higher order derivatives (215)

- ### 5.3.1 Generalized Clairaut's theorem (215)
  - **Definition 5.3.1** Let $k \in \mathbb{N}^+$. Let $U \subseteq \mathbb{R}^n$ be an open set. A function $f : U \to \mathbb{R}^m$ is **k-times continuously differentiable** (or C^k) provided all of its kth order partials exist and are continuous everywhere in U. That is, for all $i_1, \ldots, i_k \in \{1, \ldots, n\}$, the kth **order partial derivative** $\partial_{i_1} \partial_{i_2} \cdots \partial_{i_k} f$ exists and is continuous everywhere in U. (215)
    - A function f is **smooth** (or C^∞) if it is C^k for every $k \in \mathbb{N}^+$. (215)
  - **Lemma 5.3.3** Every polynomial is C^∞. (215)
  - **Theorem 5.3.4** (Generalized Clairaut) Let $U \subseteq \mathbb{R}^n$ be open. If $f : U \to \mathbb{R}^m$ is C^k, then the mixed partial derivatives of f up to order k commute. That is, for any integers $1 \leq i_1, i_2, \ldots, i_k \leq n$ and any reordering $j_1, \ldots, j_k$ of $i_1, \ldots, i_k$,
$$\partial_{i_1} \partial_{i_2} \cdots \partial_{i_k} f = \partial_{j_1} \partial_{j_2} \cdots \partial_{j_k} f.$$
(215)
- ### 5.3.2 Chain rule with higher derivatives (216)
- ### 5.3.3 Multi-index notation and polynomials (218)
  - **Definition 5.3.8** An element $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$ is a **multi-index**. The **degree of** α is the non-negative integer $|\alpha| = \alpha_1 + \cdots + \alpha_n$. The **factorial** $\alpha!$ is the positive integer $\alpha_1! \alpha_2! \cdots \alpha_n!$. (218)
  - **Definition 5.3.10** Let $U \subseteq \mathbb{R}^n$ be an open set. Let $f : U \to \mathbb{R}^m$ be a C^k function. For any multi-index $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$ with degree $|\alpha| \leq k$, define the **α-partial derivative** by
$$\partial^\alpha f := \partial_1^{\alpha_1} \cdots \partial_n^{\alpha_n} f.$$
(218)
  - **Definition 5.3.12** Let $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$ be a multi-index and let $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$. The **monomial** x^α in the variables $x_1, \ldots, x_n$ is defined by
$$x^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n}.$$
The monomial x^α has **degree** equal to $|\alpha| = \alpha_1 + \cdots + \alpha_n$ in the variables $x_1, \ldots, x_n$. (219)
  - **Definition 5.3.14** A **polynomial** is a finite linear combination of monomials. The **degree of a polynomial** is the maximum of the degree of its monomials (with non-zero coefficients). (219)
  - **Lemma 5.3.16** Let $\alpha, \beta \in \mathbb{N}^n$ be any multi-indices. Define $f : \mathbb{R}^n \to \mathbb{R}$ by $f(x) = x^\beta$. Then
    (a) If $\alpha = \beta$ then $\partial^\beta f(x) = \alpha!$ for $x \in \mathbb{R}^n$.
    (b) If $|\alpha| > |\beta|$ then $\partial^\alpha f(x) = 0$ for $x \in \mathbb{R}^n$.
    (c) If $\alpha \neq \beta$ then $\partial^\beta f(0) = 0$. (219)
  - **Lemma 5.3.18** Let $P : \mathbb{R}^n \to \mathbb{R}$ be a polynomial in n variables of degree $\leq k$, so
$$P(x) = \sum_{\substack{\alpha \in \mathbb{N}^n \\ |\alpha| \leq k}} C_\alpha x^\alpha$$
for some constants $C_\alpha \in \mathbb{R}$ with $\alpha \in \mathbb{N}^n$ and $|\alpha| \leq k$. Then $C_\alpha = \frac{\partial^\alpha P(0)}{\alpha!}$ for every such α. (220)

5.4. Taylor polynomials (221)

- ### 5.4.1 Explicit formula (221)
  - **Definition 5.4.1** Let f be a real-valued function of class C^N on an open ball centred at $a \in \mathbb{R}^n$. The **Nth Taylor polynomial of f at a** is
$$P_N(x) = \sum_{|\alpha| \leq N} \frac{\partial^\alpha f(a)}{\alpha!}(x - a)^\alpha.$$
More explicitly,
$$P_N(x) = \sum_{k=0}^{N} \sum_{\substack{\alpha_1, \ldots, \alpha_n \geq 0 \\ \alpha_1 + \cdots + \alpha_n = k}} \frac{\partial^\alpha f(a)}{\alpha_1! \cdots \alpha_n! \, \partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}}(x_1 - a_1)^{\alpha_1} \cdots (x_n - a_n)^{\alpha_n}.$$
(221)
  - **Lemma 5.4.4** Let f be of class C^2 on an open ball centred at $a \in \mathbb{R}^n$. For $x \in \mathbb{R}^n$,
$$P_0(x) = f(a),$$
$$P_1(x) = f(a) + \nabla f(a) \cdot (x - a),$$
$$P_2(x) = f(a) + \nabla f(a) \cdot (x - a) + \frac{1}{2}(x - a)^T Hf(a)(x - a).$$
(222)
- ### 5.4.2 Matching derivatives (223)
  - **Lemma 5.4.6** Let f be a real-valued function of class C^N on an open ball at $a \in \mathbb{R}^n$. Then a polynomial P is the Nth Taylor polynomial of f at a if and only if P is a polynomial of degree $\leq N$ such that for all multi-indices $\alpha \in \mathbb{N}^n$ with $|\alpha| \leq N$,
$$\partial^\alpha f(a) = \partial^\alpha P(a).$$
(223)
- ### 5.4.3 Higher order approximations (224)
  - **Lemma 5.4.8** Let $\alpha \in \mathbb{N}^n$. If $|\alpha| \geq N + 1$ then $\lim_{x \to 0} \frac{x^\alpha}{\|x\|^N} = 0$. (224)
  - **Definition 5.4.9** Let f and g be real-valued functions defined on an open ball centred at $a \in \mathbb{R}^n$. The function g is an **Nth order approximation of f at $a \in \mathbb{R}^n$** if
$$\lim_{x \to a} \frac{f(x) - g(x)}{\|x - a\|^N} = 0.$$
(225)
  - **Theorem 5.4.11** (Taylor) Let f be a real-valued function of class C^{N+1} on an open ball at $a \in \mathbb{R}^n$. A polynomial P is the Nth Taylor polynomial of f at a if and only if P is the unique degree $\leq N$ polynomial which is an Nth order approximation of f at a. (225)

5.5. Classification of critical points (228)

- ### 5.5.1 Quadratic forms (228)
  - **Definition 5.5.1** Let f be a real-valued function that is C^2 on an open ball containing $a \in \mathbb{R}^n$. The **quadratic form of f at a** is the function $q : \mathbb{R}^n \to \mathbb{R}$ defined by
$$\forall v \in \mathbb{R}^n, \quad q(v) = v^T Hf(a)v.$$
(228)
  - **Definition 5.5.2** Let A be a real $n \times n$ symmetric matrix. The **quadratic form associated to** A is the function $q : \mathbb{R}^n \to \mathbb{R}$ given by $q(v) = v^T Av$ for $v \in \mathbb{R}^n$. (228)
  - **Lemma 5.5.4** Let A be a $n \times n$ real symmetric matrix. Let $q : \mathbb{R}^n \to \mathbb{R}$ be its associated quadratic form. If $v \in \mathbb{R}^n$ is an eigenvector of A with eigenvalue $\lambda \in \mathbb{R}$, then
$$q(v) = \lambda \|v\|^2.$$
(229)
  - **Theorem 5.5.5** (Spectral Theorem) Every $n \times n$ real symmetric matrix has an orthonormal[2] basis of eigenvectors with real eigenvalues. (229)
  - **Theorem 5.5.6** Let A be a $n \times n$ real symmetric matrix. Let $q : \mathbb{R}^n \to \mathbb{R}$ be its quadratic form defined by $q(v) = v^T Av$ for $v \in \mathbb{R}^n$. The maximum and minimum values of q on the unit sphere S^{n-1} are respectively equal to the maximum and minimum eigenvalue of A. (231)
- ### 5.5.2 Second derivative test (231)
  - **Theorem 5.5.9** (Second derivative test) Let f be a real-valued function of class C^2 on an open ball at $a \in \mathbb{R}^n$. If a is a critical point of f, then f has
    (a) a local minimum if all the eigenvalues of the Hessian $Hf(a)$ are positive.
    (b) a local maximum if all the eigenvalues of the Hessian $Hf(a)$ are negative.
    (c) a saddle point if the Hessian $Hf(a)$ has a positive and a negative eigenvalue. (231)
  - **Corollary 5.5.12** Let f be a two-variable real-valued C^2 function on an open ball at $p \in \mathbb{R}^2$. If p is a critical point of f, then f has
    (a) a local minimum if $f_{xx}(p)f_{yy}(p) - (f_{xy}(p))^2 > 0$ and $f_{xx}(p) > 0$.
    (b) a local maximum if $f_{xx}(p)f_{yy}(p) - (f_{xy}(p))^2 > 0$ and $f_{xx}(p) < 0$.
    (c) a saddle point if $f_{xx}(p)f_{yy}(p) - (f_{xy}(p))^2 < 0$. (233)

5.6. Proof of Taylor's theorem (235)

- ### 5.6.1 Another explicit formula (235)
  - **Definition 5.6.1** Let f be a real-valued function of class C^k. The **kth iterated directional derivative of f** is a map defined by
$$D_h^k f := \underbrace{D_h(D_h(\cdots(D_h f)))}_{k \text{ times}}.$$
(235)
  - **Lemma 5.6.4** *Let f be of class C^k on an open ball at $a \in \mathbb{R}^n$. For all $h \in \mathbb{R}^n$,*
$$\frac{D_h^k f(a)}{k!} = \sum_{\substack{\alpha \in \mathbb{N}^n \\ |\alpha| = k}} \frac{\partial^\alpha f(a)}{\alpha!} h^\alpha.$$
(236)
  - **Corollary 5.6.5** Let f be of class C^N on an open ball at $a \in \mathbb{R}^n$. If P_k is the Nth Taylor polynomial of f at a given by (5.6.1) then for $h \in \mathbb{R}^n$,
$$P_N(a + h) = \sum_{k=0}^{N} \frac{D_h^k f(a)}{k!}.$$
(237)
- ### 5.6.2 Lagrange's remainder theorem (237)
  - **Theorem 5.6.6** (Lagrange's remainder theorem) Let $N \in \mathbb{N}$. If f is of class C^{N+1} on an open set B containing the line segment L from $a \in \mathbb{R}^n$ to $a + h \in \mathbb{R}^n$, then there exists a point[4] $c \in L$ such that the Nth remainder of f at a satisfies
$$R_N(a + h) = \frac{D_h^{N+1} f(c)}{(N + 1)!}.$$
(237)
- ### 5.6.3 Completing the proof (238)
  - **Lemma 5.6.8** Let Q be a polynomial in n variables of degree $\leq N$. Then Q is the zero polynomial if and only if
$$\lim_{x \to 0} \frac{Q(x)}{\|x\|^N} = 0. \qquad (5.6.3)$$
(238)