1 Multi-indices and higher order partials

1.1 Second-Order Partial Derivatives

Theorem 1.1. ***Clairut's Theorem*** *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a function and $a \in \mathbb{R}^n$ a point. Let $i, j \in \{1, ..., n\}$ with $i \neq j$. If $\partial_{ij} f(a)$ and $\partial_{ji} f(a)$ both exist and are continuous in a neighbourhood of a, then $\partial_{ij} f(a) = \partial_{ji} f(a)$*

Definition 1.1. C^2 **Functions** Let $U \in \mathbb{R}^n$ be an open set. We define $C^2(U, \mathbb{R})$ to be the collection of $f : \mathbb{R}^n \to \mathbb{R}$ whose second partial derivatives exist and are continuous at every point in U

Remark. Therefore, if f is a C^2 function, Clairut's theorem immediately imply that it's mixed partials exists, continuous, and hence ar equal.

An example in using high-order partial derivatives in conjunction with the chain rule.

Let $u = f(x, y)$ and suppose x, y are functions of (s, t), i.e. $x(s, t)$, $y(s, t)$. Compute $\dfrac{\partial^2 y}{\partial s^2}$

Solution.
Using the chain rule we have first order partials

$$\frac{\partial u}{\partial s} = \frac{\partial u}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial u}{\partial y}\frac{\partial y}{\partial s}$$

Then we take partials again with respect to s

$$\frac{\partial^2 u}{\partial s^2} = \frac{\partial}{\partial s}\left[\frac{\partial u}{\partial s}\right] = \frac{\partial}{\partial s}\left[\frac{\partial u}{\partial x}\frac{\partial x}{\partial s}\right] + \frac{\partial}{\partial s}\left[\frac{\partial u}{\partial y}\frac{\partial y}{\partial s}\right]$$

Note here $\dfrac{\partial u}{\partial s}$ is a function of (x, y). Thus to differentiate this function with respect to s, we must once again use the chain rule.

$$\begin{aligned}
\frac{\partial}{\partial s}\left[\frac{\partial u}{\partial x}\frac{\partial x}{\partial s}\right] &= \left[\frac{\partial}{\partial s}\frac{\partial u}{\partial x}\right]\frac{\partial x}{\partial s} + \frac{\partial u}{\partial x}\frac{\partial^2 x}{\partial s^2} && \text{(product rule)} \\
&= \left[\frac{\partial^2 u}{\partial x^2}\frac{\partial x}{\partial s} + \frac{\partial^2 u}{\partial x \partial y}\frac{\partial y}{\partial s}\right]\frac{\partial x}{\partial s} + \frac{\partial u}{\partial x}\frac{\partial^2 x}{\partial s^2} && \text{(chain rule)} \\
&= \frac{\partial^2 u}{\partial x^2}\left[\frac{\partial x}{\partial s}\right]^2 + \frac{\partial^2 u}{\partial x \partial y}\frac{\partial y}{\partial s}\frac{\partial x}{\partial s} + \frac{\partial u}{\partial x}\frac{\partial^2 x}{\partial s^2}
\end{aligned}$$

Similar computation can be applied to the latter term. Then

$$\frac{\partial^2 u}{\partial s^2} = \frac{\partial^2 u}{\partial x^2}\left[\frac{\partial x}{\partial s}\right]^2 + \frac{\partial^2 u}{\partial y^2}\left[\frac{\partial y}{\partial s}\right]^2 + 2\frac{\partial^2 u}{\partial x \partial y}\frac{\partial y}{\partial s}\frac{\partial x}{\partial s} + \frac{\partial u}{\partial x}\frac{\partial^2 x}{\partial s^2} + \frac{\partial u}{\partial y}\frac{\partial^2 y}{\partial s^2}$$

$\square$

Definition 1.2. Higher Order Partials If $U \subseteq \mathbb{R}^n$ is an open set, then for $k \in \mathbb{N}$ we define $C^k(U, \mathbb{R})$ to be the collection of functions $f : \mathbb{R}^n \to \mathbb{R}$ such that the k-th order partial derivatives of f all exist and are continuous on U. If the partials exist and are continuous for all k, we say that f is of type $C^\infty(U, \mathbb{R})$

Theorem 1.2. ***Generalized Clairuit's Theorem*** *If $f : U \subseteq \mathbb{R}^2 \to \mathbb{R}$ is of type C^k, then*

$$\partial_{i_1 \ldots i_k} f = \partial_{j_1 \ldots j_k} f$$

whenever $(i_1, \ldots, i_k)$ and $(j_1, \ldots, j_k)$ are re-orderings of each other.

Definition 1.3. Multi-index notation A multi-index α is a tuple of non-negative integers

$$\alpha = (\alpha_1, \ldots, \alpha_n)$$

The **order** of α is the sum of its components

$$|\alpha| = \alpha_1 + \alpha_2 + \cdots + \alpha_n$$

We define the multi-index **factorial** to be

$$\alpha! = \alpha_1! \alpha_2! \ldots \alpha_n!$$

If $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ then the multi-index **exponential** is

$$x^\alpha = x_1^{\alpha_1} x_2^{\alpha_2} \ldots x_n^{\alpha_n}$$

and if $f : \mathbb{R}^n \to \mathbb{R}$ we right

$$\partial^\alpha = \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2} \ldots \partial x_n^{\alpha_n}}$$

2 Taylor Series

2.1 review

Derivatives can be a tool for linearly approximating a function

$$f(x) \approx f(a) + f'(a)(x - a)$$

We can go beyong just linear approximationg and introduce quadratic, cubic, quartic approximations.

$$p_{n,a}(x) = c_n x^n + c_{n-1} x^{n-1} + \cdots + c_1 x + c_0 \text{ , where } c_k = \frac{f^{(k)}(a)}{k!}$$

$$p_{n,a}(x) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!} (x - a)^k$$

Definition 2.1. Single variable Taylor's Theorem Let $f : [a, b] \subseteq \mathbb{R} \to \mathbb{R}$. Let $n > 0$ $n \in \mathbb{Z}$. Suppose $f^{(n)}$ is continuous on $[a, b]$ and $f^{(n+1)}(t)$ exists on (a, b). Let $\alpha, \beta \in [a, b]$. Then Taylor polynomial of degree $n - 1$ of function f at point t, is denoted as

$$p(t) = p_{n,\alpha} = \sum_{k=0}^{n} C_k (t - \alpha)^k, \text{ , where } C_k = \frac{f^{(k)}(\alpha)}{k!} \in \mathbb{R}$$

Remark. Here $p(t)$ and $f(t)$ have derivatives at α that agree up to order n; that is

$$\forall k \in \{1, \ldots, n\} : p^{(k)}(\alpha) = f^{(k)}(\alpha)$$

Also note that

$$f(t) = p_{n,\alpha}(t) + r_{n,\alpha}(t)$$

If f is defined above, then for each β there eists a point x between α, β such that

$$f(\beta) = p(\beta) + \frac{f^{(n)}(x)}{(n+1)!}(\beta - \alpha)^{n+1}$$

Theorem 2.1. ***Rolle's Theorem*** *If a real-valued function f is continuous on a proper closed interval $[a, b]$, differentiable on the open interval (a, b), and $f(a) = f(b)$, then there exists at least one c in the open interval (a, b) such that*

$$f'(c) = 0$$

Proof. Since $[a, b]$ closed and bounded, intermediate value theorem applies here; that is, $f(x)$ achieves its maximum and minimum over $[a, b]$. Let $c \in [a, b]$. If $c \in (a, b)$, since f is differentiable on (a, b), $f'(c) = 0$ because it is an extremum. If $c \in \{a, b\}$ or maximum and minimum occurs at endpoints. Because $f(a) = f(b)$, then it means that $f(x)$ cannot be greater or smaller than $f(a) = f(b)$, then $f(x)$ is a constant function and $f'(x)$ is therefore 0 over (a, b) □

Theorem 2.2. ***Higher Order Rolle's Theorem*** *Assume that $f : \mathbb{R} \to \mathbb{R}$ is continuous on $[a, b]$ and $n + 1$ times differentiable on $[a, b]$. If $f(a) = f(b)$ and $f^{(k)}(a) = 0$ for all $k \in \{1, \ldots, n\}$ then there exists a $c \in (a, b)$ such that $f^{(n+1)}(c) = 0$*

Theorem 2.3. ***Taylor's Theorem with Lagrange Reminder*** *Suppose that f is $n + 1$ times differentiable on an interval I with $a \in I$. For each $x \in I$ there is a point c between a and x such that*

$$r_{n,a}(x) = \frac{f^{(n+1)}(c)}{(n+1)!}(x - a)^{n+1}$$

so if f is k times differentiable at the point a, then

$$f(x) = p_{n,a}(x) + r_{n,a}(x) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x - a)^k + \frac{f^{(n+1)}(c)}{(n+1)!}(x - a)^{n+1}$$

Corollary 2.3.1. ***Taylor reminder is a good approximation*** *If f is of type C^{n+1} on an open interval I with $a \in I$, then*

$$\lim_{x \to a} \frac{r_{n,a}(x)}{|x-a|^n} = 0$$

Moreover, we could bound $r_{n,a}(x)$ as

$$|r_{n,a}(x)| \leq \frac{M}{(n+1)!}|x-a|^{n+1} \text{ , for some } M > 0$$

Remark. This corollary just shows that the Taylor reminder is a good approximation, since error vanishes faster than order n. Note how $M > 0$ here is arbitrary because $f^{(n+1)}(c)$ is a continuous function on a compact set and therefore achieves its maximum/minimum by extreme value theorem. Also note that we can use this to determine error bounds on Taylor series

Theorem 2.4. ***Multi-variable Taylor's Theorem*** *Let $f : \mathbb{R}^n \to \mathbb{R}$ where $f \in C^{k+1}(S, \mathbb{R})$ where $S \subseteq \mathbb{R}^n$ be an open and convex set. Let $a = (a^1, \ldots, a^n) \in S$ and $x = (x^1, \ldots, x^n) \in S$. Then multivariate Taylor polynomial is given by*

$$f(x) = \sum_{|\alpha| \leq n} \frac{(\partial^\alpha f)(a)}{\alpha!}(x-a)^\alpha + r_{n,a}(x)$$

Or consider $h = x - a$, then

$$f(a+h) = \sum_{|\alpha| \leq n} \frac{(\partial^\alpha f)(a)}{\alpha!}h^\alpha + r_{n,a}(h) \text{ , where } r_{n,a}(h) = \sum_{|\alpha| = k+1} \frac{\partial^\alpha f(u)}{\alpha!}h^\alpha$$

for some u on the line joining a to x

2.2 The Hessian Matrix

Definition 2.2. Hessian Matrix If $f : \mathbb{R}^n \to \mathbb{R}$ is of class C^2 then the Hessian matrix of f at $a \in \mathbb{R}^n$ is the symmetric (i.e. $H = H^T$) $n \times n$ matrix of second order partial derivatives

$$H(a) = \begin{bmatrix} \partial_{11} f(a) & \partial_{12} f(a) & \ldots & \partial_{1n} f(a) \\ \partial_{21} f(a) & \partial_{22} f(a) & \ldots & \partial_{2n} f(a) \\ \vdots & \vdots & \ddots & \vdots \\ \partial_{n1} f(a) & \partial_{n2} f(a) & \ldots & \partial_{nn} f(a) \end{bmatrix}$$

Remark. We can use notion of Hessian matrix to simplify Taylor series formula. For second order Taylor polynomial where $x, a \in \mathbb{R}^n$

$$f(x) = f(a) + \nabla f(a)(x-a) + \frac{1}{2}(x-a)^T H(a)(x-a) + r_{2,a}(x)$$

Now we can compute simple Taylor polynomial not only from formula given but also from gradient and Hessian matrix.

Theorem 2.5. ***Spectral Theorem*** *If $A : \mathbb{R}^n \to \mathbb{R}^n$ is a symmetric matrix then there exists an ortthonormal basis consisting of eigenvectors of A.*