9.5 - Taylor's Formula

9.5 a - Notation: Let \(a, b \in \mathbb{R}^d\). Denote the line segment connecting \(a\) to \(b\) by \([a, b]\).

\[ [a, b] = \{(1-t)a + tb : 0 \leq t \leq 1 \} \]

Note that \((1-t)a + tb = a + t(b-a)\).

9.5 b - Definition: A set \(S \subset \mathbb{R}^d\) is said to be convex if \(\forall a, b \in S\), \([a, b] \subset S\).

9.5 c - Note: One can show that \(B(0, r)\) is convex for all \(r > 0\).

9.5 d - Prop [MVT]: Let \(U \subset \mathbb{R}^d\) be open and let \(f : U \to \mathbb{R}\) be \(C^1\) on \(U\). Let \(a, b \in U\) s.t. \([a, b] \subset U\). Then \(\exists c \in [a, b] \) s.t. \(f(b) - f(a) = \frac{d}{dt}f(c)(b-a)\).

PF: Let \(\gamma : [0, 1] \to \mathbb{R}^d\) by \(\gamma(t) = (1-t)a + tb\) for \(t \in [0, 1]\) and \(g : [0, 1] \to \mathbb{R}\) by \(g(t) = f(\gamma(t))\) for \(t \in [0, 1]\). Then since \(\gamma\) is cts on \([0, 1]\) and diff. on \((0, 1)\) and \(f\) is diff. on \(U\), \(g\) is cts on \([0, 1]\) and diff. on \((0, 1)\). Hence by the MVT, \(\exists e \in (0, 1)\) s.t. \(g(1) - g(0) = g'(\xi) = \frac{d}{dt}f(c)(b-a)\). Setting \(c = \xi\) yields the desired result.

9.5 e - Cor: Let \(U \subset \mathbb{R}^d\) be convex and \(f : U \to \mathbb{R}\) be \(C^1\) on \(U\). Assume that \(\exists K > 0\) s.t. \(||d(f(a))|| \leq K\) for all \(a \in U\). Then \(f\) is cont on \(U\) and \(f\) is cont on \(U\). Hence \(f(x) - f(y) = ||d(f(c))|| ||x - y|| \leq K ||x - y||\).

Given \(\epsilon > 0\), \(\exists \delta > 0\) s.t. \(||x - y|| < \delta\) then \(|f(x) - f(y)| < |f(0)| - K < K\). Hence \(f\) is cont.

9.5 f - Cor: Let \(U \subset \mathbb{R}^d\) be convex and \(f : U \to \mathbb{R}\) be \(C^1\) on \(U\). If \(df(x) = 0\) for all \(x \in U\), then \(f\) is a const. function.

PF: This follows immediately from Prop 9.5 d.

9.5 g - Notation: Let \(U \subset \mathbb{R}^d\) be open and \(f : U \to \mathbb{R}\) be \(C^r\) on \(U\) where \(r \geq 2\).

We denote the dxd matrix \(\frac{\partial^2 f(a)}{\partial x_i \partial x_j}\) by \(\frac{d^2 f(a)}{\partial x_i \partial x_j}\). Note that \(\frac{d^2 f(a)}{\partial x_i \partial x_j}\) is a symmetric dxd matrix. Given \(h \in \mathbb{R}^d\), write \(\frac{d^2 f(a)}{\partial x_i \partial x_j}h_i h_j = \frac{\partial^2 f(a)}{\partial x_i \partial x_j}h_i h_j\).

More generally, for \(r \geq 1\), set \(d^r f(a) = \frac{\partial^r f(a)}{\partial x_i \cdots \partial x_i}h_i \cdots h_r\) and for \(h \in \mathbb{R}^d\)

\[ d^r f(a) = \sum_{i_1, i_2, \cdots, i_r} \frac{\partial^r f(a)}{\partial x_{i_1} \cdots \partial x_{i_r}} h_{i_1} \cdots h_{i_r} \]

9.5 h - Prop: Let \(f : U \to \mathbb{R}\) and \(a \in U\) be as above. Let \(h \in \mathbb{R}^d\) s.t. \(f(a + th)\) is \(C^r\) for all \(t \in \mathbb{R}\) and \(f(a + th) = \sum_{i_1, i_2, \cdots, i_r} \frac{\partial^r f(a)}{\partial x_{i_1} \cdots \partial x_{i_r}} h_{i_1} \cdots h_{i_r} \) and for \(t \in \mathbb{R}\).

PF: We prove this by induction. The case \(r = 1\) follows from Cor 9.3 g. Now \(h \in \mathbb{R}^d\) s.t. \(f(a + th)\) is \(C^r\). Since it is a linear combination of all \(r\)th order partial derivatives of \(f\). If we denote this function by \(g^r\) then since \(g^r(a) = \sum_{i_1, i_2, \cdots, i_r} \frac{\partial^r f(a)}{\partial x_{i_1} \cdots \partial x_{i_r}} h_{i_1} \cdots h_{i_r} \) we have by 9.3 g that \(g^r(a) = \frac{\partial^r f(a)}{\partial x_{i_1} \cdots \partial x_{i_r}} h_{i_1} \cdots h_{i_r} \)
9.5 ctd.
Taylor's formula in this setting follows from 9.5 h and the 1-var case.

9.5 i - Theorem: Let \( U \subset R^d \) be open and let \( f: U \to R \) be \( C^n \) on \( U \). Let \( a, b \in U \)
such that \( \{a, b\} \subset U \). Then \( \exists \epsilon \in (a, b) \) s.t.
\[
(f(b) - f(a)) = f'(\epsilon)(b-a)\]

\[
f''(\epsilon)(a-x)^2 + \cdots + \frac{1}{n!}f^{(n)}(\epsilon) (a-x)^n
\]

\( \epsilon = \text{Apply } 1\text{-var Taylor's formula to } g(t) = f(a + t(b-a)) \text{ and use } 9.5 h. \)

With \( f \) as above, the \( k^{th} \) Taylor poly of \( f \) at \( a \) for \( k \leq n+1 \) is given by
\[
T_k(x) = f(a) + df(a)(x-a) + \frac{1}{2}d^2f(a)(x-a)^2 + \cdots + \frac{1}{k!}d^k f(a)(x-a)^k
\]

9.5 j - Ex: Let \( f: R^2 \to R \) be defined by \( f(x,y) = x^2 y + 3y^3 - 12y - x^2 \). Find \( T_2 \) at \((1,2)\).
\[
\begin{align*}
\frac{df}(x,y) &= (2xy - 2x, x^2 + 3y^2 - 12) \\
\frac{d^2 f}(x,y) &= (2y - 2, 2x \quad 6y)
\end{align*}
\]

So \( \frac{df}(1,2) = (2, 2) \) and \( \frac{d^2 f}(1,2) = (0, 12) \). Since \( f(1,2) = -15 \),
\[T_2(x,y) = -15 + 2(x-1) + (y-2) + (x-1)^2 + 2(x-1)(y-2) + 6(y-2)^2.\]

9.5 l - Def: Let \( U \subset R^d \) be open, let \( a \in U \) and let \( f: U \to R \) be a function. We say that \( f \) has a \textit{local max} at \( a \) if \( \forall \epsilon > 0 \) s.t. \( x \in U \) \( f(x) \leq f(a) \) for all \( x \in U \) \( \text{s.t.} \ |x-a| < \epsilon \). Sim. say that \( f \) has a \textit{local min} at \( a \) if \( \forall \epsilon > 0 \) s.t. \( f(x) \geq f(a) \) for all \( x \in U \) \( \text{s.t.} \ |x-a| < \epsilon \). A \textit{local extremum} is a local max or a local min.

9.5 m - Prop: Let \( U \subset R^d \) be open, let \( a \in U \) and let \( f: U \to R \) be a function. If \( f \) is diff. at \( a \) and it has a local extremum at \( a \), then \( df(a) = 0 \).

pf: Suppose \( f \) is diff. at \( a \) and has a local ext at \( a \). Then for every \( v \in R^d \), \( g(v) = f(a + tv) \) has a local ext at \( 0 \). Hence \( 0 = g'(0) = df(a)(v) \). Thus \( df(a) = 0 \).

9.5 n - Cor: Let \( K \subset R^d \) be compact and let \( f: K \to R \) be cts. Then both the abs max and min occur on \( K \). If an extremum occurs at \( x_0 \), then \( x_0 \in K \) and either \( df(x_0) = 0 \) or \( f \) is not diff at \( x_0 \).

9.5 o - Def: Let \( A \) be a \( d \times d \) matrix. We say \( A \) is \textit{positive definite} if \( x^TAx > 0 \) for all \( x \neq 0 \) and negative definite if \( x^TAx < 0 \) for all \( x \neq 0 \).

9.5 p - Prop: Let \( A \) be a symmetric \( d \times d \) matrix. Then all the eigenvalues of \( A \) are real. If they are all positive, \( A \) is positive definite. If they are all negative, \( A \) is negative definite. (Recall that \( df(a) \) is sym. if \( f \) is \( C^2 \) on \( U \).

9.5 q - Theorem: Let \( U \subset R^d \) be open, let \( f: U \to R^2 \in C^2 \) and \( \forall a \in U \) \( df(a) = 0 \).

Then
\[
\begin{enumerate}
\item If \( df(a) \) is positive definite, \( f \) has a local min at \( a \).
\item If \( df(a) \) is negative definite, \( f \) has a local max at \( a \).
\end{enumerate}

The proof of (i) depends on the following lemma (the proof of (ii) is similar).

9.5 r - Lemma: Let \( A \) be a pos. def. \( d \times d \) matrix. Then \( \exists \epsilon > 0 \) s.t. if \( B \) is a \( d \times d \) matrix s.t. \( ||A-B||_M \leq \epsilon \), then \( u^TBu > \epsilon \) for all unit vectors \( u \in R^d \) and \( B \) is pos. def.

pf: Note that \( K = \{ x \in R^d : ||x||_M = 1 \} \) is compact \& \( u \mapsto u^TAu \) is cts. Thus \( E = \inf \{ u^TAu : u \in K \} > 0 \).

Let \( B \) be a \( d \times d \) matrix s.t. \( ||A-B||_M < \epsilon \). Let \( u \in K \).

Then \( u^T(A-B)u \leq ||u||_M^2 ||A-B||_M \leq ||A-B||_M \leq \epsilon \), hence \( u \mapsto u^T(A-B)u \geq E - \epsilon \) s.t. \( u \mapsto u^TBu \geq E \), thus \( B \) is pos. def.
9.5 - ctd.

pf. of Theorem 9.5p(i): Note that $d^2f(a)$ is pos. def. Let $E > 0$ be as in 9.5q with $A = d^2f(a)$. Since $f$ is $C^2$, $x \mapsto d^2f(x)$ is cts. Hence, $\exists \delta > 0$ s.t. if $||x-a|| < \delta$, then $||d^2f(x) - d^2f(a)|| < E$ and so $d^2f(x)$ is pos. def. Now let $x \in U - \{a\}$ s.t. $||x-a|| < \delta$ and $[a,x] \subset U$. Then by 9.5i (Taylor’s formula) with $n = 1$ and $b = x$,

$\exists \epsilon \subset [a,x]$ s.t. $f(x) = f(a) + d^2f(a)(x-a) + \frac{1}{2} d^2f(c)(x-a)^2 = f(a) + \frac{1}{2} d^2f(a)(x-a)^2$

since $||c-a|| < ||x-a|| < \delta$, $d^2f(c)$ is pos. def. So if $x \neq a$, $d^2f(c)(x-a)^2 = (x-a) \cdot d^2f(c)(x-a) > 0$, hence, $f(x) > f(a)$ and $f$ has a local min at $a$. $\square$

9.5r - Def: Let $f: U \rightarrow \mathbb{R}$ and $a \in U$ be as above and s.t. $d^2f(a) = 0$. If $\exists u, v \in \mathbb{R}$ s.t. $d^2f(a)u^2 > 0$ and $d^2f(a)v^2 < 0$, we say that $f$ has a saddle pt. at $a$.

Note that this happens iff $d^2f(a)$ has at least one pos. and one neg. eval. In this case, there is no local extremum at $a$.

9.5s - Cor: Let $U \subset \mathbb{R}^2$, $f: U \rightarrow \mathbb{R}$ be $C^2$ and let $a \in U$. S.t. that $d^2f(a) = 0$ and $\Delta = \text{det} d^2f(a) \neq 0$. Then

i) If $\Delta > 0$ and $\frac{\partial^2 f}{\partial x^2}(a) > 0$, $f$ has a loc. min at $a$.

ii) If $\Delta > 0$ and $\frac{\partial^2 f}{\partial x^2}(a) < 0$, $f$ has a loc. max at $a$.

iii) If $\Delta < 0$, $f$ has a saddle pt. at $a$.

9.5t - Ex: Let $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ be given by $f(x,y) = x^2y + y^3 - 12y - x^2$ for $(x,y) \in \mathbb{R}^2$. Find all loc. extrema and saddle points of $f$.

First show that $d^2f(a) = 0$ iif $a = (0,2)$, or $a = (\pm 3,1)$. We have $d^2f(x,y) = \begin{pmatrix} 2y & 2x \\ 2x & 6y \end{pmatrix}$ and so $\Delta = \text{det} d^2f(x,y) = 12y(y-1) - 4x^2$. Thus, there are saddle pts at $(\pm 3,1)$, a local min at $(0,2)$ with $f(0,2) = -16$ and a local max at $(0,-2)$ with $f(0,-2) = 16$. 