MAT 307, Multivariable Calculus with Linear Algebra – Fall 2024

1 Taylor's Theorem

If $f : U \to \mathbb{R}$ is a differentiable function, $P \in U$, one can use the derivative to write down the best linear (tangent plane) approximation to f at P. One might also like to do better using quadratic, or higher degree, approximation.

Recall in one-variable calculus that

Definition 1. $f : I \subset \mathbb{R} \to \mathbb{R}$ is $\mathcal{C}^k$-function, for $x_0 \in I$, the **k-th order Taylor polynomial of f centered at** x_0 is given by

$$T_k f(x, x_0) = f(x_0) + f'(x_0)(x - x_0) + \cdots \frac{f^k(x_0)}{k!}(x - x_0)^k = \sum_{i=0}^{k} \frac{f^i(x_0)}{i!}(x - x_0)^i.$$

The **remainder** is the difference

$$R_k f(x, x_0) = f(x) - T_k f(x, x_0), \qquad \frac{R_k f(x, x_0)}{(x - x_0)^k} \to 0.$$

The Taylor polynomial is chosen so that the first k derivatives of $T_k f$ at x_0 are precisely the same as those of f. So the first k derivatives of the remainder are all zero. The remainder is a measure of how good the Taylor approximation of f.

Generalizing to functions from $\mathbb{R}^n$ to $\mathbb{R}$.

Definition 2. Let $U \subset \mathbb{R}^n$ be an open subset which is convex (i.e. if P and Q are in U, then so does every point on the line between $P = (p_1, \ldots, p_n)$ and $Q = (x_1, \ldots, x_n)$). Suppose $f : U \to \mathbb{R}$ is $\mathcal{C}^k$. Given $P \in U$, the **k-th Taylor polynomial of f centered at** P is

$$T_k f(Q, P) = f(P) + \sum_{1 \le i \le n} \frac{\partial f}{\partial x_i}(P)(x_i - p_i) + \frac{1}{2} \sum_{1 \le i,j \le n} \frac{\partial^2 f}{\partial x_i \partial x_j}(P)(x_i - p_i)(x_j - p_j)$$

$$+ \cdots \frac{1}{k!} \sum_{1 \le i_1, i_2, \cdots, i_k \le n} \frac{\partial^k f}{\partial x_{i_1} \cdots \partial x_{i_k}}(P)(x_{i_1} - p_{i_1}) \cdots (x_{i_k} - p_{i_k}).$$

The **remainder** is the difference

$$R_k f(Q, P) = f(Q) - T_k f(Q, P).$$

If f is $\mathcal{C}^{k+1}$, then

$$\frac{R_k(f(Q, P)}{\|\overrightarrow{PQ}\|^k} \to 0.$$

Writing out the first few terms of the Taylor polynomial of f, we have

$$T_2f(Q,P) = f(P) + \sum_{1\le i\le n} \frac{\partial f}{\partial x_i}(P)(x_i - p_i) + \frac{1}{2}\sum_{1\le i,j\le n} \frac{\partial^2 f}{\partial x_i \partial x_j}(P)(x_i - p_i)(x_j - p_j)$$

The second term is simply

$$\nabla f(P) \cdot \overrightarrow{PQ}.$$

The third term is

$$\frac{1}{2}(\overrightarrow{PQ})^{tr} H(f)(P)\overrightarrow{PQ}, \qquad H(f)(P) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(P)\right),$$

where $H(f)(P)$ is called the **Hessian** of f at P. It is a symmetric matrix because mixed partials are equal.

Example 1. Find the second order Taylor polynomial approximation of $f(x,y,z) = x^2y - 4z$ near the point $P = (1,2,4)$.
Let $Q = (x,y,z)$, then

$$\begin{aligned}
& T_2f(Q,P) \\
= \;& f(P) + \nabla f(P)\cdot \overrightarrow{PQ} + \frac{1}{2}\overrightarrow{PQ}H(f)(P)\cdot\overrightarrow{PQ} \\
= \;& -14 + (4,1,-4)\cdot(x-1,y-2,z-4) + \frac{1}{2}(x-1,y-2,z-4)\begin{pmatrix} 2y & 2x & 0 \\ 2x & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}_{|P} \begin{pmatrix} x-1 \\ y-2 \\ z-4 \end{pmatrix} \\
= \;& -14 + (4,1,-4)\cdot(x-1,y-2,z-4) + \frac{1}{2}(x-1,y-2,z-4)\begin{pmatrix} 4 & 2 & 0 \\ 2 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x-1 \\ y-2 \\ z-4 \end{pmatrix} \\
= \;& -14 + 4(x-1) + (y-2) - 4(z-4) + 2(x-1)^2 + 2(x-1)(y-2)
\end{aligned}$$

Example 2.

$$\begin{bmatrix} x_1 & x_2 \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = ax_1y_1 + bx_1y_2 + cx_2y_1 + dy_1y_2$$

$$\begin{bmatrix} x & y \end{bmatrix}\begin{bmatrix} a & b \\ b & c \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = ax^2 + 2bxy + cy^2$$

Example 3. Find the second order Taylor polynomial approximation of $f(x,y,z) = x^2y - 4z$ near the point $P = (1,2,4)$.

Let $Q = (x, y, z)$, then

$$\begin{aligned}
& T_2 f(Q, P) \\
= \;& f(P) + \nabla f(P) \cdot \overrightarrow{PQ} + \frac{1}{2}\overrightarrow{PQ} H(f)(P) \cdot \overrightarrow{PQ} \\
= \;& -14 + (4, 1, -4) \cdot (x-1, y-2, z-4) + \frac{1}{2}(x-1, y-2, z-4) \begin{pmatrix} 2y & 2x & 0 \\ 2x & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}_{|P} \begin{pmatrix} x-1 \\ y-2 \\ z-4 \end{pmatrix} \\
= \;& -14 + (4, 1, -4) \cdot (x-1, y-2, z-4) + \frac{1}{2}(x-1, y-2, z-4) \begin{pmatrix} 4 & 2 & 0 \\ 2 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x-1 \\ y-2 \\ z-4 \end{pmatrix} \\
= \;& -14 + 4(x-1) + (y-2) - 4(z-4) + 2(x-1)^2 + 2(x-1)(y-2)
\end{aligned}$$

Note that $H(f)(P)$ is symmetric.

What about Taylor approximation at $P = (0, 0, 0)$? $f(x, y, z)$ is already a polynomial!

$$T_1 f(Q, (0,0,0)) = T_2 f(Q, (0,0,0)) = -4z, \qquad T_3(f(Q, (0,0,0)) = x^2 y - 4z.$$

Recall that $K \subset \mathbb{R}^n$ is **closed** if the complement K^c is open, and this is equivalent to saying that K contains all of its boundary points.

Definition 3. $K \subset \mathbb{R}^n$ is said to be **bounded** if there is a real number M such that $\|\vec{x}\| \leq M$ for all $\vec{x} \in K$. K is **compact** if K is closed and bounded.

Example 4.

1. $[a, b]$ is compact.
2. $(a, b]$ is bounded but not closed, so not compact.
3. $[a, \infty)$ is closed but not bounded, so not compact.
4. $K = \{x \in \mathbb{R}^n \mid \|x\| \leq M\}$ is compact.
5. $K = \{x \in \mathbb{R}^n \mid \|x\| < M\}$ is bounded but not closed.

We quote a key theorem without proof:

Theorem 1. *Suppose $f : K \to \mathbb{R}$ is continuous, if K is compact, then there are two points P_{min} and P_{max} in K such that for all $P \in K$, we have*

$$f(P_{min}) \leq f(P) \leq f(P_{max}).$$

To find the maxima and minima of $f : K \to \mathbb{R}$, we break the problem into two parts:

I. investigate the interior points, using the derivative test.

II. investigate the boundary points, use Lagrange multipliers (next lecture).

Definition 4. Let $f : K \to \mathbb{R}$ be a function, $P \in K$ an interior point. We say f has a **local minimum (resp. maximum) at** P if there is an open ball $U = B_\delta(P)$ centered at P contained in K such that $f(P) \leq f(Q)$ (resp. $f(P) \geq f(Q)$) for all $Q \in U$. An interior point $P \in K$ is called a **critical point** if f is not differentiable at P or if it is, then $Df(P) = \nabla f(P) = \vec{0}$.

At critical points where derivatives are defined, all directional derivatives are 0.

Proposition 2. *Let $K \subset \mathbb{R}^n$ be compact, and $f : K \to \mathbb{R}$ a differentiable function. If an interior point $P \in K$ is a local maximum (resp. minimum), then P is a critical point.*

Proof. Note that since P is a local maximum, for $h > 0$, we have

$$\frac{\partial f}{\partial x_i}(P) = \lim_{h\to 0^+} \frac{f(P + h\hat{e}_i) - f(P)}{h} \leq 0, \qquad \frac{\partial f}{\partial x_i}(P) = \lim_{h\to 0^+} \frac{f(P - h\hat{e}_i) - f(P)}{-h} \geq 0$$

for all i. All the partial derivatives exist, and they must vanish. □

Example 5.
- $f(x, y) = x^4 + y^4 - 4xy$, $\nabla f(x, y) = (4(x^3 - y), 4(y^3 - x))$, f has critical points at $(0, 0)$, $(1, 1)$ and $(-1, -1)$.
- Volcano: $f(x, y) = (x^2 + y^2)e^{-(x^2+y^2)}$, $\nabla f(x, y) = (2x(1 - (x^2 + y^2)), 2y(1 - (x^2 + y^2)))e^{-(x^2+y^2)}$. The unit circle and $(0, 0)$ are critical points.

Recall the one variable second derivative test that follows from Taylor's Theorem applied to the second Taylor polynomial.

1. If $f'(a) = 0$ and $f''(a) < 0$, then a is a local maximum of f.
2. If $f'(a) = 0$ and $f''(a) > 0$, then a is a local minimum of f.
3. If $f'(a) = 0$ and $f''(a) = 0$, then inconclusive.

In higher dimensions, the second Taylor polynomial centered at P is

$$T_2 f(Q, P) = f(P) + \nabla f(P) \cdot \overrightarrow{PQ} + \frac{1}{2}\overrightarrow{PQ} H(f)(P) \overrightarrow{PQ}^{tr}.$$

where

$$H(f)(P) = \left[\frac{\partial^2 f}{\partial x_i \partial x_j}(P)\right].$$

Note that $H(f)(P)$ is a symmetric matrix when f is $\mathcal{C}^2$. At a critical point, the second term vanishes. The important term is the quadratic Hessian term. The difference between $f(Q)$ and $T_2 f(Q, P)$ is small compared to the quadratic term if f is $\mathcal{C}^3$.

2 Optimization

In higher dimensions, the second Taylor polynomial centered at P is

$$T_2 f(Q,P) = f(P) + \nabla f(P)\cdot \overrightarrow{PQ} + \frac{1}{2}\overrightarrow{PQ}H(f)(P)\overrightarrow{PQ}^{tr}.$$

where

$$H(f)(P) = \left[\frac{\partial^2 f}{\partial x_i \partial x_j}(P)\right].$$

At a critical point, the second term vanishes. The important term is the quadratic Hessian term. The difference between $f(Q)$ and $T_x f(Q,P)$ is small compared to the quadratic term if f is $\mathcal{C}^3$.

Definition 5. If A is a symmetric $n \times n$ matrix, then the function

$$Q_A : \mathbb{R}^n \to \mathbb{R}, \qquad Q_A(\vec{x}) = \vec{x}^{tr} A \vec{x}$$

is called a **symmetric quadratic form**. Q_A is **positive definite (resp. negative)** if $Q_A(\vec{x}) > 0$ (resp. $Q_A(\vec{x}) < 0$) for all $\vec{x} \neq 0$; Q_A is called **positive semidefinite (resp. negative semidefinite)** if $Q_A(\vec{x}) \geq 0$ (resp. $Q_A(\vec{x}) \leq 0$) for all $\vec{x} \neq 0$.

Example 6. Let $A = I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, then $Q_A(x,y) = (x,y)A\begin{pmatrix} x \\ y \end{pmatrix} = x^2 + y^2$ is positive definite. If $A = -I_2$, then $Q_A(x,y) = -x^2 - y^2$ is negative definite. If $A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, then $Q_A(x,y) = x^2 - y^2$ is neither positive nor negative definite.

Using Taylor's Theorem, we have

Proposition 3. *$f : K \to \mathbb{R}$ is $\mathcal{C}^3$, if $P \in K \subset \mathbb{R}^n$ is an interior critical point, then*

(a) If $H(f)(P)$ is positive definite, then P is a local minimum.

(b) If $H(f)(P)$ is negative definite, then P is a local maximum.

(c) If $\vec{x}^{tr} H(f)(P)\vec{x} \neq \vec{0}$ for all $\vec{v} \neq 0$, and $H(f)(P)$ and is neither positive nor negative definite, then P is a saddle point.

Examples of critical points behavior at $(0,0)$:

- $x^2 + y^2$: local min
- $-x^2 - y^2$: local max
- $x^2 - y^2$: saddle

- $x^4 \pm y^4, -x^4 - y^4, x^2 \pm y^4$: none of the above

Given the Hessian matrix, we have a determinant test for positive definiteness:

Theorem 4 (Determinant Test)**.** *If A is an $n \times n$ matrix, let d_i be the determinant of the upper left $i \times i$ submatrix. Let $Q_A(\vec{x}) = \vec{x}^{tr} A \vec{x}$.*

(a) If $d_i > 0$ for all i, then Q_A is positive definite.

(b) If $d_i > 0$ for i even and $d_i < 0$ for i odd, then Q_A is negative definite.

In all other cases, the test is inconclusive.

We do not prove this. In the case of a 2×2 matrix $A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$, we have

$$Q(x, y) = ax^2 + 2bxy + cy^2.$$

Assume $d_1 = a > 0$, then

$$Q(x, y) = (\sqrt{a}x + \frac{b}{\sqrt{a}}y)^2 + (c - \frac{b^2}{a})y^2 = (\sqrt{a}x + \frac{b}{\sqrt{a}}y)^2 + \frac{ac - b^2}{a}y^2.$$

$d_1 = a > 0$ and $d_2 = ac - b^2$.

Hence if $a > 0$ and $ac - b^2 > 0$, then we have a local min; if $a < 0$ and $ac - b^2 > 0$, we have a local max; if $ac - b^2 < 0$, we have a saddle; if $ac - b^2 = 0$, then we have an indeterminate critical point that has to be analyzed by other methods.

Example 7. $f(x, y) = x^3/3 - x - (y^3/3 - y)$, $\nabla f(x, y) = (x^2 - 1, 1 - y^2)$. The critical points are at $(1, 1)$, $(1, -1)$, $(-1, 1)$, and $(-1, -1)$. The Hessian $H(f)(x, y) = \begin{pmatrix} 2x & 0 \\ 0 & -2y \end{pmatrix}$. $(1, 1)$ is a saddle, $(-1, 1)$ is a local max, $(1, -1)$ is a local min, $(-1, -1)$ is a saddle.

Example 8.

$$f(x, y) = x^2 - y^4, \qquad g(x, y) = x^2 + y^4$$

$$H(f)(0, 0) = H(g)(0, 0) = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}$$

but $(0, 0)$ is a saddle for f and a local min for g.

To find global min and max in a domain D, need to check all of the following:

- behavior as P goes to infinity if D is unbounded

- points along the boundary of D

- critical points in interior of D

Example 9. Find global min and max of $f(x,y) = x+y+\frac{8}{xy}$ in the first quadrant where $x > 0$ and $y > 0$.

- As $x \to \infty$ or $y \to \infty$, $f(x,y) \to \infty$, so global max does not exist.
- As (x,y) approaches the boundary of D, i.e. $x \to 0$ or $y \to 0$, $f(x,y) \to \infty$, so there must be min in the domain D.
- $f_x = 1 - \frac{8}{x^2y}$ and $f_y = 1 - \frac{8}{xy^2}$ are defined everywhere in the domain.
- $f_x = 0$ and $f_y = 0$ implies $x^2y = xy^2 = 8$, so $(x,y) = (2,2)$. $f(2,2) = 6$ is global min.

Example 10. Let $K = \{(x,y) \mid x^2 + y^2 \leq 2\}$, then K is compact. Define a function $f : K \to \mathbb{R}$ by $f(x,y) = xy$. f is continuous, since K is compact, f must attain its max and min on K. Since $\nabla f(x,y) = (y,x)$, the only critical point is $(0,0)$.

$$H(f)(x,y) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Since $d_1 = 0$ and $d_2 = -1 < 0$, determinant test is inconclusive. As $f(x,x) > 0$ and $f(x,-x) < 0$, $(0,0)$ is a saddle point. Hence the max and min of f must be on the boundary given by

$$C = \{(x,y) \mid x^2 + y^2 = 2\}.$$

Let $g : \mathbb{R}^2 \to \mathbb{R}$ be the function $g(x,y) = x^2 + y^2$, we would like to maximize or minimize f subject to the condition $g(x,y) = 2$.

One solution is to eliminate a variable say y from $g(x,y) = 2$, and substitute it into f to find max and min.

Alternatively, we can parametrize the boundary by

$$\vec{r}(t) = \sqrt{2}(\cos t, \sin t).$$

We need to maximize the composition $h : [0, 2\pi] \to \mathbb{R}$ where $h(t) = 2\cos t \sin t$. Since $[0, 2\pi]$ is compact, h attains its max and min.

$$h'(t) = 2\cos^2 t - 2\sin^2 t = 0, \qquad t = \pi/4, 3\pi/4, 5\pi/4, 7\pi/4.$$

The method of Lagrange multiplier does the opposite. Instead of eliminating a variable, we add one variable called λ.

To find min/max of f subject to the constraint $g = c$, where f and g are both $\mathcal{C}^1$, **Lagrange multiplier method**, whereby $\nabla f = \lambda \nabla g$ where $\nabla g \neq \vec{0}$ at a min/max, works, because ∇g is perpendicular to level curve of $g = c$. For f to attain min/max, ∇f must be perpendicular to level curve of $g = c$ as well, otherwise, the directional derivative of f in one of the two directions along $g = c$ would be positive, so one could increase f by moving in that direction, meaning we weren't really at maximum. Since both ∇f and ∇g are perpendicular to the level curve, ∇f must be a multiple of ∇g at a local max or min.

$$\nabla f(x, y) = \lambda \nabla g(x, y), \qquad g(x, y) = 2$$

$$y = 2\lambda x, \qquad x = 2\lambda y, \qquad x^2 + y^2 = 2$$

$$x = 2\lambda(2\lambda x) = 4\lambda^2 x$$

Either $x = 0$, then $y = 2\lambda x = 0$, but $g(0, 0) = 0 \neq 2$!
Or $x \neq 0$, and $4\lambda^2 = 1$, $\lambda = \pm\frac{1}{2}$.
When $\lambda = \frac{1}{2}$, $y = x$, $g(x, y) = 2x^2 = 2$, $f(x, y) = x^2 = 1$.
When $\lambda = -\frac{1}{2}$, $y = -x$, $g(x, y) = 2x^2 = 2$, $f(x, y) = -x^2 = -1$.
We have four critical points $(\pm 1, \pm 1)$. The maximum of f is 1 at $\pm(1, 1)$ while the minimum of f is -1 at $\pm(1, -1)$.