MATH 2010 Chapter 10

10.1 Taylor Series Expansion

Recall

Taylor expansion for 1-variable function $g(t)$ at $t = 0$ up to order k.

$$\boxed{g(t) = g(0) + g'(0)t + \frac{1}{2!}g''(0)t^2 + \cdots + \frac{1}{k!}g^{(k)}(0)t^k + \text{remainder} \quad \circledast}$$

We want a similar formula for a multi-variable function $f(x)$ defined near a, where $x = (x_1, \cdots, x_n), \quad a = (a_1, \cdots, a_n)$.
Let $g(t) = f(a + t(x - a))$
If $\|x - a\|$ is small, then for $|t| \leq 1$,

$$\|t(x-a)\| = |t|\|x-a\| \leq \|x-a\| \text{ is small}$$

and $g(t)$ is defined.

By $\circledast$,

$$f(a + t(x-a)) = g(0) + g'(0)t + \frac{1}{2!}g''(0)t^2 + \cdots + \frac{1}{k!}g^{(k)}(0)t^k + \text{remainder}$$

Put $t = 1$,

$$f(x) = g(0) + g'(0) + \frac{1}{2!}g''(0) + \cdots + \frac{1}{k!}g^{(k)}(0) + \text{remainder}$$

Next, express $g^{(k)}(0)$ in terms of f :

$$g(0) = f(a + t(x-a)) = f(a)$$

$$\begin{aligned} g'(t) &= \nabla f(a + t(x-a)) \cdot \frac{d}{dt}(a + t(x-a)) \\ &= \nabla f(a + t(x-a)) \cdot (x-a) \\ &= \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a + t(x-a))(x_i - a_i) \end{aligned}$$

$$\Rightarrow g'(0) = \nabla f(a) \cdot (x-a)$$
$$= \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a)(x_i - a_i)$$

$$\begin{aligned} g''(t) &= \frac{d}{dt} g'(t) \\ &= \sum_{i=1}^{n} \frac{d}{dt}[\frac{\partial f}{\partial x_i}(a + t(x-a))(x_i - a_i)] \\ &= \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_j \partial x_i}(a + t(x-a))(x_j - a_j)(x_i - a_i) \\ g''(0) &= \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_j \partial x_i}(a)(x_j - a_j)(x_i - a_i) \end{aligned}$$

Hence, Taylor Expansion at a up to order 2 is

$$\boxed{f(x) = f(a) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a)(x_i - a_i) + \frac{1}{2!} \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(a)(x_i - a_i)(x_j - a_j) + \text{remainder}}$$

Similarly, the general term is

$$g^{(k)}(0) = \sum_{i_1, \cdots, i_k = 1}^{n} \frac{\partial^k f}{\partial x_{i_1} \cdots \partial x_{i_k}}(a)(x_{i_1} - a_{i_1}) \cdots (x_{i_k} - a_{i_k})$$

Example 10.1. If $n = 2$, i.e. $f = f(x, y), a = (x_0, y_0)$ f is C^2 (so $f_{xy} = f_{yx}$), then

$$\begin{aligned} f(x,y) =& f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0) \\ &+ \frac{1}{2}[f_{xx}(x_0, y_0)(x - x_0)^2 + 2f_{xy}(x_0, y_0)(x - x_0)(y - y_0) + f_{yy}(x_0, y_0)(y - y_0)^2] \\ &+ \text{remainder} \end{aligned}$$

Theorem 10.2 (Taylor's Theorem). *Let $\Omega \subseteq \mathbb{R}^n$ be open, $f : \Omega \to \mathbb{R}$ be C^k.*

Then for any $x, a \in \Omega$,

$$\begin{aligned} f(x) =& f(a) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a)(x_i - a_i) + \frac{1}{2!} \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(a)(x_i - a_i)(x_j - a_j) + \cdots \\ &+ \frac{1}{k!} \sum_{i_1, \cdots, i_k = 1}^{n} \frac{\partial^k f}{\partial x_{i_1} \cdots \partial x_{i_k}}(a)(x_{i_1} - a_{i_1}) \cdots (x_{i_k} - a_{i_k}) + \varepsilon_k(x, a) \end{aligned},$$

with:

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

Definition 10.3.

$$p_k(x)=f(a)+\sum_{i=1}^{n}\frac{\partial f}{\partial x_i}(a)(x_i-a_i)+\cdots$$
$$+\frac{1}{k!}\sum_{i_1,\cdots,i_k=1}^{n}\frac{\partial^k f}{\partial x_{i_1}\cdots\partial x_{i_k}}(a)(x_{i_1}-a_{i_1})\cdots(x_{i_k}-a_{i_k})$$

is called the **k-th order Taylor polynomial** of f at a.

Remark.

- $p_1(x)=L(x)=$ Linearization of f at a
- p_k and f have equal partial derivatives up to order k at a.

IFRAME
open in new window

Example 10.4. $f(x,y)=e^x\cos y$ Find the 2^{nd} order Taylor polynomial at $a=(0,0)$

Solution.

$$\begin{aligned} f_x &= e^x\cos y & f_y &= -e^x\sin y\\ f_{xx} &= e^x\cos y & f_{yx} &= -e^x\sin y\\ f_{xy} &= -e^x\sin y & f_{yy} &= -e^x\cos y \end{aligned}$$

$\Rightarrow \quad f(0,0)=1$,

$$\begin{aligned} f_x(0,0) &= 1 & f_y(0,0) &= 0\\ f_{xx}(0,0) &= 1 & f_{yy}(0,0) &= -1\\ f_{xy}(0,0) &= f_{yx}(0,0)=0 \end{aligned}$$

$$\begin{aligned} p_2(x,y) &= f(0,0)+f_x(0,0)x+f_y(0,0)y+\frac{1}{2!}(f_{xx}(0,0)x^2+2f_{xy}(0,0)xy+f_{yy}(0,0)y^2)\\ &= 1+x+\frac{1}{2}x^2-\frac{1}{2}y^2 \end{aligned}$$

How about $p_3(x,y)$ at $(0,0)$?

$$p_3(x,y) = p_2(x,y) + \underbrace{\frac{1}{3!}g^{(3)}(0)}_{3^{rd}\text{ order terms}}$$

$$\begin{aligned} f_{xxx} &= e^x \cos y & f_{xxy} &= f_{xyx} = f_{yxx} = -e^x \sin y \\ f_{yyy} &= e^x \sin y & f_{xyy} &= f_{yxy} = f_{yyx} = -e^x \cos y \end{aligned}$$

$$\begin{aligned} \Rightarrow f_{xxx}(0,0) &= 1 & f_{xxy}(0,0) &= 0 \\ f_{xyy}(0,0) &= -1 & f_{yyy}(0,0) &= 0 \end{aligned}$$

$$\begin{aligned} g^{(3)}(0) &= f_{xxx}(0,0)x^3 + 3f_{xxy}(0,0)x^2y + 3f_{xyy}(0,0)xy^2 + f_{yyy}(0,0)y^3 \\ &= x^3 - 3xy^2 \end{aligned}$$

$$\begin{aligned} p_3(x,y) &= p_2(x,y) + \frac{1}{3!}(x^3 - 3xy^2) \\ &= 1 + x + \frac{1}{2}x^2 - \frac{1}{2}y^2 + \frac{1}{6}x^3 - \frac{1}{2}xy^2 \end{aligned}$$

Question If $f = f(x,y,z)$ is C^6 , then coefficient of xy^2z^3 in $p_6(x,y,z)$ at $(0,0,0)$ is $\alpha f_{xyyzzz}(0,0,0), \alpha =?$

10.1.1 Matrix form for 2nd order Taylor Polynomial

Definition 10.5. Let $\Omega \subseteq \mathbb{R}^n$ be open, $f : \Omega \to \mathbb{R}$ be C^2.

Then the **Hessian matrix** of f at $a \in \Omega$ is:

$$Hf(a) = \begin{bmatrix} f_{x_1x_1}(a) & \cdots & f_{x_1x_n}(a) \\ \vdots & \ddots & \vdots \\ f_{x_nx_1}(a) & \cdots & f_{x_nx_n}(a) \end{bmatrix}$$

Remark.

- $Hf(a)$ is a symmetric $n \times n$ matrix by the mixed derivatives theorem.
- In Thomas' Calculus, Hessian of f is defined to be the determinant of our Hessian matrix.

With the Hessian matrix, the 2nd order Taylor polynomial of f at a can be written as:

$$\boxed{\underset{1\times 1}{p_2(x)} = \underset{1\times 1}{f(a)} + \underset{1\times n}{\nabla f(a)}\underset{n\times 1}{(x-a)} + \frac{1}{2}\underset{1\times n}{(x-a)^\top}\underset{n\times n}{Hf(a)}\underset{n\times 1}{(x-a)}}$$

where $x, a \in \mathbb{R}^n$ are written as column vectors:

$$\begin{aligned}(x-a)^\top &= \text{Transpose of } x-a \\ &= [x_1 - a_1, \cdots, x_n - a_n]\end{aligned}$$

Remark.

$$\begin{aligned}
&(x-a)^\top Hf(a)(x-a) \\
=&[x_1 - a_1, \cdots, x_n - a_n]\begin{bmatrix} f_{x_1x_1}(a) & \cdots & f_{x_1x_n}(a) \\ \vdots & \ddots & \vdots \\ f_{x_nx_1}(a) & \cdots & f_{x_nx_n}(a)\end{bmatrix}\begin{bmatrix} x_1 - a_1 \\ \vdots \\ x_n - a_n\end{bmatrix} \\
=&[x_1 - a_1, \cdots, x_n - a_n]\begin{bmatrix} f_{x_1x_1}(a)(x_1-a_1) + \cdots + f_{x_1x_n}(a)(x_n-a_n) \\ \vdots \\ f_{x_nx_1}(a)(x_1-a_1) + \cdots + f_{x_nx_n}(a)(x_n-a_n)\end{bmatrix} \\
=&f_{x_1x_1}(a)(x_1-a_1)(x_1-a_1) + \cdots + f_{x_1x_n}(a)(x_1-a_1)(x_n-a_n) \\
&+ \cdots \\
&\ \vdots \\
&+ f_{x_nx_1}(a)(x_1-a_1)(x_n-a_n) + \cdots + f_{x_nx_n}(a)(x_n-a_n)(x_n-a_n) \\
=&\sum_{i,j=1}^{n} f_{x_ix_j}(a)(x_i-a_i)(x_j-a_j) \\
=&g^{(2)}(0)
\end{aligned}$$

Example 10.6.

$$f(x,y) = e^x \cos y$$

Find $p_2(x,y)$ at $a = (0,0)$ using matrix form.

Solution.

$$f(0,0) = 1$$

$$\nabla f(0,0) = (1,0)$$

$$Hf(0,0) = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

$$\begin{aligned} p_2(x,y) &= f(0,0) + \nabla f(0,0) \begin{bmatrix} x-0 \\ y-0 \end{bmatrix} + \frac{1}{2}[x-0 \quad y-0] Hf(0,0) \begin{bmatrix} x-0 \\ y-0 \end{bmatrix} \\ &= 1 + [1 \quad 0] \begin{bmatrix} x \\ y \end{bmatrix} + \frac{1}{2}[x \quad y] \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \\ &= 1 + x + \frac{1}{2}x^2 - \frac{1}{2}y^2 \end{aligned}$$

Example 10.7.

$$g(x,y) = \frac{\ln x}{1-y}$$

Find $p_2(x,y)$ at $(1,0)$.

Solution.

$$g(1,0) = 0$$

$$\nabla g = [g_x, g_y] = \left[\frac{1}{x(1-y)}, \frac{\ln x}{(1-y)^2} \right]$$

$$Hg = \begin{bmatrix} g_{xx} & g_{xy} \\ g_{yx} & g_{yy} \end{bmatrix} = \begin{bmatrix} -\frac{1}{x^2(1-y)} & \frac{1}{x(1-y)^2} \\ \frac{1}{x(1-y)^2} & \frac{2\ln x}{(1-y)^3} \end{bmatrix}$$

$$\nabla g(1,0) = [1 \quad 0] \quad Hg(1,0) = \begin{bmatrix} -1 & 1 \\ 1 & 0 \end{bmatrix}$$

$$\begin{aligned} p_2(x,y) &= g(0,0) + \nabla g(0,0) \begin{bmatrix} x-1 \\ y \end{bmatrix} + \frac{1}{2}[x-1 \quad y] Hg(1,0) \begin{bmatrix} x-1 \\ y \end{bmatrix} \\ &= 0 + [1 \quad 0] \begin{bmatrix} x-1 \\ y \end{bmatrix} + \frac{1}{2}[x-1 \quad y] \begin{bmatrix} -1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x-1 \\ y \end{bmatrix} \\ &= (x-1) - \frac{1}{2}(x-1)^2 + (x-1)y \end{aligned}$$

10.1.2 Application to local maximum / minimum

Suppose $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ is C^2, and a is a critical point of f.

Then, $\nabla f(a) = \vec{0}$. For x near a ,

$$
\begin{aligned}
f(x) &\approx p_2(x) \\
&= f(a) + \nabla f(a)(x-a) + \frac{1}{2}(x-a)^\top Hf(a)(x-a) \\
&= \quad f(a) + \underbrace{\frac{1}{2}(x-a)^\top Hf(a)(x-a)}_{\text{This term determines whether } f(x)>f(a) \text{ or } f(x)<f(a)}
\end{aligned}
$$

For $n = 1$, i.e. f is 1-variable.

$$\frac{1}{2}(x-a)^\top Hf(a)(x-a) = \frac{1}{2}f''(a)(x-a)^2$$

Recall: Second Derivative Test

This may be viewed as a consequence of Taylor's Theorem. That is, if $f'(a) = 0$, then near $x = a$, we have:

$$f(x) \approx f(a) + \underbrace{f'(a)(x-a)}_{=0} + \frac{1}{2}f''(a)(x-a)$$

IFRAME

The sign of the second derivative at $x = a$ essentially tells us whether locally the graph of the function looks like an upward or downward parabola.

For $n = 2$, the 2^{nd} order term is:

$$\frac{1}{2}[x-x_0 \quad y-y_0] \underbrace{\begin{bmatrix} f_{xx}(x_0,y_0) & f_{xy}(x_0,y_0) \\ {}_{yx}(x_0,y_0) & f_{yy}(x_0,y_0) \end{bmatrix}}_{f \text{ is } C^2 \Rightarrow \text{Symmetric}} \begin{bmatrix} x-x_0 \\ y-y_0 \end{bmatrix}$$

To understand the nature of critical points, we study **quadratic forms** of 2 variables.

$$
\begin{aligned}
q(x,y) &= [x \quad y] \begin{bmatrix} A & B \\ B & C \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \\
&= Ax^2 + 2Bxy + Cy^2
\end{aligned}
$$

Does $q(x,y)$ have a definite sign (always positive or always negative) for $(x,y) \neq (0,0)$?

We can determine it by completing square.

Example 10.8.

$$q(x,y) = 2xy \left(= [x \;\; y] \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \right)$$

IFRAME

Note $q(x,y) = \frac{1}{2}(x+y)^2 - \frac{1}{2}(x-y)^2$ Along $x+y=0$, i.e. $y=-x$,

$$q(x,-x) = -2x^2 < 0 \text{ for } x \neq 0$$

Along $x-y=0$, i.e. $y=x$

$$q(x,x) = 2x^2 > 0 \text{ for } x \neq 0$$

Hence, q has no definite sign, i.e. indefinite.

Clearly $(0,0)$ is a critical point of $q(x,y)$ but neither local maximum nor minimum.

Such a critical point is called a **saddle point** .

Example 10.9.

$$q(x,y) = 17x^2 - 12xy + 8y^2 \left(= [x \;\; y] \begin{bmatrix} 17 & -6 \\ -6 & 8 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \right)$$

IFRAME

Does $q(x,y)$ have a definite sign?

Solution.

$$\begin{aligned} q(x,y) &= 17[x^2 - \frac{2\cdot 6}{17}xy + (\frac{6}{17})^2y^2] + (8 - \frac{36}{17})y^2 \\ &= 17(x - \frac{6}{17}y)^2 + 10y^2 \quad \circledast \end{aligned}$$

Hence, $q(x,y) > 0 = q(0,0)$ for $(x,y) \neq (0,0)$ Hence, The critical point $(0,0)$ is a local minimum. Also global minimum of $q(x,y)$.

Remark. Expression like $\circledast$ is called diagonalization of quadratic form. It is not unique!

For example $q(x,y) = \underset{\uparrow}{5}(\frac{x+2y}{\sqrt{5}})^2 + \underset{\uparrow}{20}(\frac{2x-y}{\sqrt{5}})^2$ is another diagonalization.

"Orthogonal" change of coordinates

10.1.3 Higher dimension example

Example 10.10.

$$q(x,y,z) = xy + yz + zx$$

Definite sign for $(x,y,z) \neq (0,0,0)$?

Solution.

$$q = \frac{1}{4}(x+y)^2 - \frac{1}{4}(x-y)^2 + z(x+y)$$

Let $u = \frac{x+y}{2} \quad v = \frac{x-y}{2}$. Then

$$\begin{aligned} q &= u^2 - v^2 + 2uz \\ &= (u^2 + 2uz + z^2) - v^2 - z^2 \\ &= (u+z)^2 - v^2 - z^2 \\ &= (\frac{x+y}{2} + z)^2 - (\frac{x-y}{2})^2 - z^2 \\ &= \underset{\substack{\uparrow \\ \text{positive}}}{\frac{1}{4}} (x+y+2z)^2 - \frac{1}{4}(x \underset{\uparrow}{-} y)^2 \underset{\uparrow}{-} z^2 \end{aligned}$$

negative

On the plane $x + y + 2z = 0$, i.e. $z = -\frac{x+y}{2}$

$$\begin{aligned} q &= q(x, y, -\frac{x+y}{2}) \\ &= -\frac{1}{4}(x-y)^2 - \frac{1}{4}(x+y)^2 < 0 \text{ for } (x,y,z) \neq (0,0,0) \end{aligned}$$

Along the line $x - y = z = 0$, i.e. $y = x, z = 0$

$$\begin{aligned} q(x,y,z) &= q(x,x,0) \\ &= x^2 > 0 \text{ for } x \neq 0 \end{aligned}$$

Hence, the critical point $(0,0,0)$ is a saddle point. For general theory, need linear algebra:

Diagonalization of quadratic form, eigenvalues $\cdots$

Definition 10.11. Let A be a $n \times n$ symmetric matrix.

Then A is said to be

- **positive definite** if $x^\top Ax > 0$ for all column vectors $x \in \mathbb{R}^n \backslash \{\vec{0}\}$

- **negative definite** if $x^\top Ax < 0$ for all column vectors $x \in \mathbb{R}^n \backslash \{\vec{0}\}$
- **indefinite** if $\exists$ column vectors $x, y \in \mathbb{R}^n \backslash \{\vec{0}\}$ such that $x^\top Ax > 0$ and $y^\top Ay < 0$

Remark. These are not all the possible cases:

There are symmetric matrix which is not positive definite, negative definite nor indefinite.

Example 10.12.

$$[x \quad y] \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = x^2 + 4y^2 > 0 \quad \text{for all } \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2 \backslash \{\vec{0}\}$$

Hence, $\begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix}$ is positive definite.

Example 10.13.

$$[x \quad y] \begin{bmatrix} -1 & 0 \\ 0 & -4 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = -x^2 - 4y^2 < 0 \quad \text{for all } \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2 \backslash \{\vec{0}\}$$

Hence, $\begin{bmatrix} -1 & 0 \\ 0 & -4 \end{bmatrix}$ is negative definite.

Example 10.14.

$$[x \quad y] \begin{bmatrix} -1 & 0 \\ 0 & 4 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = -x^2 + 4y^2$$
$$[1 \quad 0] \begin{bmatrix} -1 & 0 \\ 0 & 4 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = -1 < 0$$
$$[0 \quad 1] \begin{bmatrix} -1 & 0 \\ 0 & 4 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = 4 > 0$$

Hence, $\begin{bmatrix} -1 & 0 \\ 0 & 4 \end{bmatrix}$ is indefinite.

Example 10.15.

$$[x \quad y] \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = x^2 \geq 0 \quad \text{for all } \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2$$

$$[0 \quad 1] \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = 0 \Rightarrow \text{ not positive definite}$$

Hence, $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ is neither positive/negative definite nor indefinite.

Example 10.16.

$$
\begin{aligned}
&[x \quad y] \begin{bmatrix} 1 & 2 \\ 2 & 5 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \\
=&x^2 + 4xy + 5y^2 \\
=&(x^2 + 4xy + 4y^2) + y^2 \\
=&(x + 2y)^2 + y^2 > 0 \quad \text{for all } \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2 \backslash \{\vec{0}\}
\end{aligned}
$$

Hence, $\begin{bmatrix} 1 & 2 \\ 2 & 5 \end{bmatrix}$ is positive definite.