- Finding extrema
  - EVT applicable
  - EVT not applicable

- Taylor series expansion.

Thm (Taylor's theorem)

$\underset{\text{open}}{\Omega \subseteq} \mathbb{R}^n$, $f : \Omega \to \mathbb{R}$. C^k-function.

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

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

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

$$+ \mathcal{E}_k(x, a)$$

with $\lim_{x \to a} \frac{\mathcal{E}_k(x,a)}{\|x - a\|^k} = 0$

Def $P_k(x)$

k-th order Taylor polynomial of f at a

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

2nd order Taylor polynomial of f at $(0,0)$?

(sol) $f_x = e^x \cos y$, $f_y = -e^x \sin y$ $\quad f_{yy} = -e^x \cos y$

$f_{xx} = e^x \cos y$, $f_{xy} = -e^x \sin y$, $f_{yx} = -e^x \sin y$.

$f(0) = 1$

$\Rightarrow f_x(0,0) = 1, \quad f_y(0,0) = 0$

$f_{xx}(0,0) = 1, \quad f_{xy}(0,0) = f_{yx}(0,0) = 0, \quad f_{yy}(0,0) = -1$

$$\therefore P_2(x) = 1 + 1\cdot(x-0) + 0\cdot(y-0) + \frac{1}{2!}\Big(1\cdot(x-0)^2 + 0\cdot(x-0)(y-0) + 0\cdot(x-0)(y-0) - 1\cdot(y-0)^2\Big)$$

$$= 1 + x + \frac{1}{2}x^2 - \frac{1}{2}y^2$$

$P_3(x,y)$? $\quad P_3(x,y) = P_2(x,y) + \underline{\text{degree 3 terms}},$

$f_{xxx} = e^x\cos y \quad \Rightarrow f_{xxx}(0,0) = 1$

$f_{xxy} = f_{xyx} = f_{yxx} = -e^x\cos y \Rightarrow f_{xxy}(0,0) = -1$

$f_{xyy} = f_{yxy} = f_{yyx} = -e^x\sin y \Rightarrow f_{xyy}(0,0) = 0$

$f_{yyy} = e^x\sin y \quad \Rightarrow f_{yyy}(0,0) = 0.$

$$\therefore P_3(x,y) = P_2(x,y) + \frac{1}{3!}\Big(f_{xxx}(0,0)x^3 + 3f_{xxy}(0,0)x^2y + 3f_{xyy}(0,0)xy^2 + f_{yyy}(0,0)y^3\Big)$$

$$= 1 + x + \frac{1}{2}x^2 - \frac{1}{2}y^2 + \frac{1}{6}(x^3 - 3x^2 y)$$

$$= 1 + x + \frac{1}{2}x^2 - \frac{1}{2}y^2 + \frac{1}{6}x^3 - \frac{1}{2}x^2 y$$

Matrix form for 2nd order Taylor polynomial.

Def $f : \Omega (\subseteq \mathbb{R}^n) \to \mathbb{R}$, C^2-function.

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

$$Hf(a) = \begin{pmatrix} f_{x_1 x_1}(a) & \cdots & f_{x_1 x_n}(a) \\ \vdots & & \vdots \\ f_{x_n x_1}(a) & \cdots & f_{x_n x_n}(a) \end{pmatrix}$$

Rmk ① $Hf(a)$ is a symmetric $n \times n$ matrix.
($\because$ mixed derivatives theorem)

② In some textbooks, Hessian of f is defined to be the determinant of Hessian matrix.

2nd order Taylor polynomial of f at a can be written as

$$\underset{1\times 1}{P_2(x)} = \underset{1\times 1}{f(a)} + \underbrace{\underset{1\times n}{\nabla f(a)} \cdot \underset{n\times 1}{(x-a)}}_{1\times 1} + \frac{1}{2}\underbrace{\underset{1\times n}{(x-a)^T}\, \underset{n\times n}{Hf(a)}\, \underset{n\times 1}{(x-a)}}_{1\times 1}$$

$x, a \in \mathbb{R}^n$ as column vectors

$(x-a)^T$: the transpose of $x - a = (x_1 - a_1, \cdots, x_n - a_n)$

$$(x-a)^T Hf(a)(x-a)$$

$$= (x_1-a_1, \cdots, x_n-a_n) \begin{pmatrix} f_{x_1x_1}(a) & \cdots & f_{x_1x_n}(a) \\ \vdots & & \vdots \\ f_{x_nx_1}(a) & \cdots & f_{x_nx_n}(a) \end{pmatrix} \begin{pmatrix} x_1-a_1 \\ \vdots \\ x_n-a_n \end{pmatrix}$$

$$\qquad\qquad\qquad\qquad\qquad\qquad n\times n \qquad\qquad\qquad n\times 1$$

$$= [x_1-a_1, \cdots, x_n-a_n] \begin{pmatrix} f_{x_1x_1}(a)(x_1-a_1) + \cdots + f_{x_1x_n}(a)(x_n-a_n) \\ \vdots \\ f_{x_nx_1}(a)(x_1-a_1) + \cdots + f_{x_nx_n}(a)(x_n-a_n) \end{pmatrix}$$

$$1\times n \qquad\qquad\qquad\qquad n\times 1$$

$$= f_{x_1x_1}(a)(x_1-a_1)(x_1-a_1) + \cdots + f_{x_1x_n}(a)(x_n-a_n)(x_1-a_1)$$
$$+ \cdots$$
$$+ f_{x_nx_1}(a)(x_1-a_1)(x_n-a_n) + \cdots + f_{x_nx_n}(a)(x_n-a_n)(x_n-a_n)$$

$$= \sum_{i,j=1}^{n} f_{x_i x_j}(a)\,(x_i-a_i)(x_j-a_j)$$

eg $f(x,y) = e^x \cos y$ $\quad p_2(x,y)$ at $(0,0)$

$f(0,0) = 1$, $\quad Df = (e^x\cos y, -e^x \sin y)$

$Df(0,0) = (1, 0)$

$$Hf = \begin{pmatrix} e^x \cos y & -e^x \sin y \\ -e^x \sin y & -e^x \cos y \end{pmatrix} \quad Hf(0,0) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

$$P_2(x,y) = f(0,0) + Df(0,0)\begin{pmatrix} x-0 \\ y-0 \end{pmatrix} + \frac{1}{2}(x-0 \,.\, y-0)\, Hf(0,0)\begin{pmatrix} x-0 \\ y-0 \end{pmatrix}$$

$$= 1 + (1 \; 0)\begin{pmatrix} x \\ y \end{pmatrix} + \frac{1}{2}(x \; y)\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$$

$$= 1 + x + \frac{1}{2}x^2 - \frac{1}{2}y^2$$

eg $\quad g(x,y) = \dfrac{\ln x}{1-y}$. $\quad P_2(x,y)$ of g at $(1,0)$?

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

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

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

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

$$Hg(1,0) = \begin{pmatrix} -1 & 1 \\ 1 & 0 \end{pmatrix}$$

$$\therefore P_2(x,y) = g(1,0) + \nabla g(1,0)\begin{pmatrix} x-1 \\ y-0 \end{pmatrix} + \frac{1}{2}(x-1\ \ y-0)\, Hg(1,0)\begin{pmatrix} x-1 \\ y-0 \end{pmatrix}$$

$$= 0 + (1,0)\begin{pmatrix} x-1 \\ y-0 \end{pmatrix} + \frac{1}{2}(x-1\ \ y-0)\begin{pmatrix} -1 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} x-1 \\ y \end{pmatrix}$$

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

Application to local max/min.

Suppose f is C^2, a is a critical point of f.

$\nabla f(a) = 0$.

$$p_2(x) = f(a) + \nabla f(a)\cdot(x-a) + \frac{1}{2}(x-a)^T Hf(a)(x-a)$$

$$= f(a) + 0 + \underbrace{\frac{1}{2}(x-a)^T Hf(a)(x-a)}$$

This term determines whether $f(x) > f(a)$ or $f(x) < f(a)$.

Recall: $n=1$, one-variable calculus

$Hf(a) = f''(a)$

$$\frac{1}{2}(x-a)^T Hf(a)(x-a) = \frac{1}{2}f''(a)|x-a|^2$$

If a is a critical point, ($f'(a)=0$)

$$\begin{cases} f''(a) > 0 \Rightarrow \bigcup \text{ local min at } a \\ f''(a) < 0 \Rightarrow \text{ local max at } a \end{cases} \quad \left(\begin{array}{c}\text{2nd} \\ \text{derivative} \\ \text{test}\end{array}\right)$$

For $n=2$, the 2nd order term is

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

To understand nature of critical points, we study <u>quadratic forms</u> of 2 variables.

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

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

(in one variable, ax^2, $-ax^2$ $(a>0)$)

$ax^2 \to$ positive $\to$ local min $x=0$

$-ax^2 \to$ negative $\to$ local max $x=0$

eg1 $g(x,y) = 2xy = (x\ \ y)\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$

Note that $g(x,y) = \frac{1}{2}\left((x+y)^2 - (x-y)^2\right)$

Along $x+y=0$, $g(x,y) = g(x,-x) = -2x^2 < 0$ $x \neq 0$

" $x-y=0$, $g(x,y) = g(x,x) = 2x^2 > 0$ $x \neq 0$.

$\therefore g$ has no definite sign.

(i.e. indefinite)

Clearly $(0,0)$ is a critical point of $g(x,y)$.

but neither local max or local min.

Such a critical point is called a saddle point.

eg2 $g(x,y) = 17x^2 - 12xy + 8y^2$

Definite sign?

Yes, $g(x,y) = 17\left(x^2 - \frac{12}{17}xy\right) + 8y^2$

$$= 17\left(x - \frac{6}{17}y\right)^2 - \frac{36}{17}y^2 + 8y^2$$

$$= 17\left(x - \frac{6}{17}y\right)^2 + \frac{100}{17}y^2 \quad (*)$$

$\therefore g(x,y) > 0 = g(0,0)$ for $(x,y) \neq (0,0)$

$\therefore$ The critical point $(0,0)$ is a local min of $g(x,y)$

(global minimum in this case).

Rmk Expression like (*) is called diagonalization of quadratic form. It is not unique.

For example, $q(x,y) = 5\cdot\left(\frac{x+2y}{\sqrt{5}}\right)^2 + 20\left(\frac{2x-y}{\sqrt{5}}\right)^2$

is another diagonalization.

Higher dimensional example.

eg $q(x,y,z) = xy + yz + zx$. Definite sign for $(x,y,z) \neq (0,0,0)$?

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

Let $u = \frac{x+y}{2}$, $v = \frac{x-y}{2}$

$$= u^2 - v^2 + 2uz$$
$$= (u^2 + 2uz + z^2) - z^2 - v^2$$
$$= (u+z)^2 - v^2 - z^2$$
$$= 1\cdot\left(\frac{x+y}{2} + z\right)^2 - 1\cdot\left(\frac{x-y}{2}\right)^2 - z^2$$

positive

negative

$\Rightarrow q(x,y,z)$ indefinite.

$\therefore (0,0,0)$ is a saddle point.

For general theory (all n variables) need linear algebra

: diagonalization of quadratic forms, eigenvalues.

Def Let A be a $n \times n$ symmetric matrix. Then A is said to be

① positive definite if $x^T A x > 0$ for all column vectors $x \in \mathbb{R}^n - \{0\}$.

② negative definite if $x^T A x < 0$ for all columns vectors $x \in \mathbb{R}^n - \{0\}$.

③ indefinite if $\exists$ column vectors $x, y \in \mathbb{R}^n - \{0\}$ s.t. $x^T A x > 0$ $y^T A y < 0$.

Eg ① $(x \ y)\begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = x^2 + 4y^2 > 0$ $\begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 - \{0\}$

$\therefore \begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix}$ is positive definite

② $(x \ y)\begin{pmatrix} -1 & 0 \\ 0 & -4 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = -x^2 - 4y^2 < 0$ $\begin{pmatrix} x \\ y \end{pmatrix} \neq 0$

$\therefore \begin{pmatrix} -1 & 0 \\ 0 & -4 \end{pmatrix}$ negative definite

③ $(x \ y)\begin{pmatrix} -1 & 0 \\ 0 & 4 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = -x^2 + 4y^2$ indefinite.

If $\begin{pmatrix}x\\y\end{pmatrix}=\begin{pmatrix}1\\0\end{pmatrix} \Rightarrow -x^2+4y^2=-1<0$

$\begin{pmatrix}x\\y\end{pmatrix}=\begin{pmatrix}0\\1\end{pmatrix} \Rightarrow \sim = 4 > 0$

Rmk It is possible that symmetric matrix is neither positive definite, negative definite, indefinite.

① $(x\ \ y)\begin{pmatrix}1&0\\0&0\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix} = x^2 \geq 0$

but if $\begin{pmatrix}x\\y\end{pmatrix}=\begin{pmatrix}0\\1\end{pmatrix} \Rightarrow (0\ \ 1)\begin{pmatrix}1&0\\0&0\end{pmatrix}\begin{pmatrix}0\\1\end{pmatrix}=0$

$\Rightarrow$ not positive definite.

$\begin{pmatrix}1&0\\0&0\end{pmatrix}$ is not positive definite
negative definite
indefinite

eg $(x\ \ y)\begin{pmatrix}1&2\\2&5\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}$

$= x^2+4xy+5y^2$

$= (x^2+4xy+4y^2)+y^2$

$= (x+2y)^2+y^2 > 0$ for $\forall \begin{pmatrix}x\\y\end{pmatrix} \in \mathbb{R}^2 - \{0\}$

$\therefore \begin{pmatrix}1&2\\2&5\end{pmatrix}$ is positive definite.

Second derivative test

Thm $\Omega \subseteq \mathbb{R}^n$ open, $f: \Omega \to \mathbb{R}$ is C^2.

$a \in \Omega$ is a critical point (i.e. $\nabla f(a) = 0$).

Then if $Hf(a)$ is

positive definite $\Rightarrow$ a is a local min

negative " $\Rightarrow$ " local max

indefinite $\Rightarrow$ saddle point.

(idea of pf) By Taylor's thm,

$$f(x) \approx f(a) + \nabla f(a) \cdot (x-a) + \frac{1}{2}(x-a)^T Hf(a)(x-a)$$

$$= f(a) + \frac{1}{2}(x-a)^T Hf(a)(x-a).$$

$$f(x) - f(a) \approx \frac{1}{2}(x-a)^T Hf(a)(x-a)$$

If $Hf(a)$ is positive definite, RHS > 0

for all $x - a \neq 0$ i.e. $x \neq a$.

$\Rightarrow f(x) - f(a) > 0$ for all $x \neq a$ near a

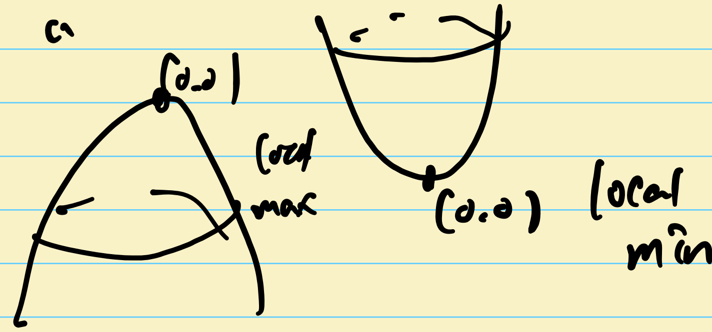
$\Rightarrow$ a is a local min of f.

Similar idea works for other cases $\square$

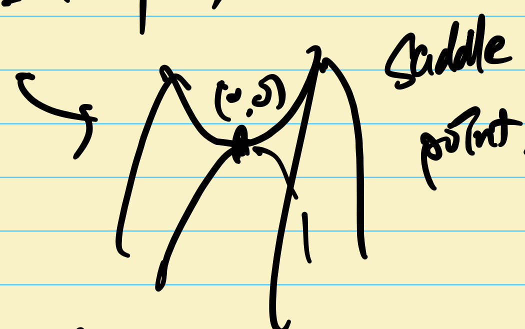
Geometrically,

① $Hf(a)$ is positive definite (e.g. $f = x^2+y^2$ at $(0,0)$)

② " " negative (e.g. $f = -x^2-y^2$)



③ " indefinite (e.g. $f = x^2-y^2$)



Q How can we determine definiteness of $Hf(a)$?

$n=1$; nothing to do

$n=2$; complete the square

Thm Let $M = \begin{pmatrix} A & B \\ B & C \end{pmatrix}$. Then

- M is positive definite $\Leftrightarrow$ $AC-B^2>0$, $A>0$
- " negative " $\Leftrightarrow$ $AC-B^2>0$, $A<0$
- " indefinite " $\Leftrightarrow$ $AC-B^2<0$.

Note that $AC-B^2 = \det M$

(proof) Let $q(x,y) = (x\ y) M \begin{pmatrix} x \\ y \end{pmatrix} = Ax^2 + 2Bxy + Cy^2$

case 1) $(A \neq 0)$

$$Aq(x,y) = A^2x^2 + 2ABxy + ACy^2$$
$$= (Ax+By)^2 + (AC-B^2)y^2$$

Hence $q(x,y) > 0$ $\forall (x,y) \neq (0,0) \iff AC-B^2 > 0,\ A > 0$

" < 0 " $\iff AC-B^2 > 0,\ A < 0$

$q(x,y)$ can have both signs $\iff AC-B^2 < 0$.

case 2) $(A=0)$ $AC-B^2 = -B^2 \leq 0$.

$$q(x,y) = 2Bxy + Cy^2$$
$$= y(2Bx + Cy)$$

q is neither positive definite nor negative definite.

q is indefinite $\iff B \neq 0 \iff AC-B^2 < 0$.

$\square$

Thm (Second derivative test)

$f : \Omega \to \mathbb{R}$ is C^2, $a \in \Omega$. $\nabla f(a) = 0$.
$(\Omega \subset \mathbb{R}^2)$

Then

① $f_{xx} f_{yy} - f_{xy}^2 > 0$, $f_{xx} > 0$ at a

$\Rightarrow$ a is a local min

② $f_{xx} f_{yy} - f_{xy}^2 > 0$, $f_{xx} < 0$ at a

$\Rightarrow$ a local max

③ $f_{xx} f_{yy} - f_{xy}^2 < 0$ at a

$\Rightarrow$ a is a saddle point

④ $f_{xx} f_{yy} - f_{xy}^2 = 0$ at a

$\Rightarrow$ inconclusive.

Rmk ④ ; a can be local max/min, saddle point.