Chap 6-Week15

2023/05/16 16:00

Chap6.6-6.7 Definite Matrices

1. Definitions of different types of definite matrices
2. Some important theorems
3. Applications: Hessian matrix and identifying local extrema/saddle point

Recall:

Thm A is a real symmetric matrix

$\Leftrightarrow$ A is orthogonally diagonalizable.

Thm If A is a real symmetric matrix, and $\lambda \neq \mu$ are two distinct eigenvalues of A. $\vec{v}$ and $\vec{w}$ are corresponding eigenvectors.

Then $\vec{v} \perp \vec{w}$.

E.g. $A = \begin{bmatrix} 5 & 2 \\ 2 & 5 \end{bmatrix}$ $\quad |A_1| = 5 > 0$, $\quad |A| = 25 - 4 = 21 > 0$

check:

$\forall$ nonzero $\vec{x} \in \mathbb{R}^2$,

$\vec{x}^T A \vec{x} = 5x_1^2 + 4x_1x_2 + 5x_2^2$

$= 5(x_1^2 + \frac{4}{5}x_1x_2 + \frac{4}{25}x_2^2) + \frac{21}{5}x_2^2$ $\quad 2 \cdot x_1 \cdot \frac{2}{5}x_2$

$= 5(x_1 + \frac{2}{5}x_2)^2 + \frac{21}{5}x_2^2 > 0$.

$\Rightarrow$ A is positive definite matrix.

Exercise: $A = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$ is negative definite.

△ **Definition of Definite Matrices**

A real symmetric matrix $A_{n\times n}$ is said to be

(i) positive definite if $\vec{x}^T A \vec{x} > 0$ for all nonzero $\vec{x} \in \mathbb{R}^n$.

(ii) negative definite if $\vec{x}^T A \vec{x} < 0$ for all nonzero $\vec{x} \in \mathbb{R}^n$.

(iii) positive semidefinite if $\vec{x}^T A \vec{x} \geq 0$ for all nonzero $\vec{x} \in \mathbb{R}^n$.

(iv) negative semidefinite if $\vec{x}^T A \vec{x} \leq 0$ for all nonzero $\vec{x} \in \mathbb{R}^n$.

(v) indefinite if $\vec{x}^T A \vec{x}$ takes on values that differ in sign.

Thm Let A be a real symmetric $n\times n$ matrix. Then A is positive definite if and only if all its eigenvalues are positive.

Thm If A is a real symmetric positive (negative) definite matrix, then A is nonsingular.

$\det(A) = \prod_{i=1}^{n} \lambda_i > 0$ since A is positive definite and $\lambda_i > 0 \ \forall i$.

(complete the proof for negative definite matrix).

E.g. $A = \begin{bmatrix} 3 & 2 \\ 2 & 0 \end{bmatrix}$ $\quad \lambda_1 = -1$, $\lambda_2 = 4$

$\Rightarrow$ indefinite matrix

△ **Eigenvalue Test**

(i) A is positive definite $\Leftrightarrow$ All eigenvalues of A are positive.

(ii) A is negative definite $\Leftrightarrow$ All eigenvalues of A are negative.

(iii) A is indefinite $\Leftrightarrow$ Some eigenvalues of $A > 0$ and some < 0.

(iv) A is positive semidefinite $\Leftrightarrow$ All eigenvalues of A are non-negative.

(v) A is negative semidefinite $\Leftrightarrow$ All eigenvalues of A are non-positive.

E.g. $A = \begin{bmatrix} 6 & 4 & -2 \\ 4 & 5 & 3 \\ -2 & 3 & 6 \end{bmatrix}$ $\quad A_1 = 6$, $A_2 = \begin{pmatrix} 6 & 4 \\ 4 & 5 \end{pmatrix}$, $A_3 = A$

By eigenvalue test,

$\det(A - \lambda I) = \begin{vmatrix} 6-\lambda & 4 & -2 \\ 4 & 5-\lambda & 3 \\ -2 & 3 & 6-\lambda \end{vmatrix} = (6-\lambda)[(5-\lambda)(6-\lambda)-9] - 4[4(6-\lambda)+6] - 2[12 + 2(5-\lambda)]$

$= (6-\lambda)(\lambda^2 - 11\lambda + 30 - 25) - 4(5-\lambda) - 48$

$= -\lambda^3 + 17\lambda^2 + 67\lambda - 38$

$\lambda_1 = -0.5015$, $\lambda_2 = 7.8569$, $\lambda_3 = 9.6446$

$\Rightarrow$ A is indefinite.

Def (Leading Principal Submatrix)

For an $n\times n$ matrix A, the leading principal submatrix A_r of A of order r is the matrix formed by deleting the last $n-r$ rows and columns of A.

$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ \vdots & & & & \vdots \\ a_{n1} & \cdots & & \cdots & a_{nn} \end{bmatrix}_{n\times n}$

$A_1 = a_{11}$

$A_2 = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$

$\vdots$

$A_r = \begin{bmatrix} a_{11} & \cdots & a_{1r} \\ \vdots & \ddots & \vdots \\ a_{r1} & \cdots & a_{rr} \end{bmatrix}$

$|A_1| = 6 > 0$
$|A_2| = 30 - 4^2 = 14 > 0$
$|A| = \cdots < 0$, determinant Test fail

By determinant test,

det (A_1) = , det (A_2) =
det (A) =

↳

△ Determinant Test

<u>Thm</u> (i) A is positive definite $\Leftrightarrow$

① the leading principal submatrices A_k are all positive definite

② that is, $\det(A_k) > 0$ for $1 \le k \le n$

(optional)

(ii) A is negative definite $\Leftrightarrow$

① the leading principal submatrices are all negative definite

② that is, $(-1)^k \det(A_k) > 0$ for $1 \le k \le n$. $\Leftrightarrow |A_1| < 0,\ |A_2| > 0,\ |A_3| < 0, \cdots$

(iii) If neither of above fits, the determinant test fails.

Exercise:

$A = \begin{bmatrix} \boxed{2} & 0 & 0 \\ 0 & 5 & 3 \\ 0 & 3 & 5 \end{bmatrix}$ $|A_1| = 2 > 0$, $|A_2| = 10 > 0$, $|A| = 32 > 0$

"+"-tive definite

$B = \begin{bmatrix} -2 & 0 & 1 \\ 0 & -1 & 0 \\ 1 & 0 & -2 \end{bmatrix}$ $|B_1| = -2 < 0$, $|B_2| = 2 > 0$, $|B| = (-2)\times 2 + |\times|$ $= -3 < 0$

"−"-tive definite

$C = \begin{bmatrix} 3 & 2 \\ 2 & 0 \end{bmatrix}$ $\det(C) = \lambda_1 \cdot \lambda_2 < 0$

determinant Test fails
eigenvalue test suggests indefinite matrix

△ Applications: Using Definite Matrices to determine local extrema / saddle point

Recall 1-D case:

$y = f(x)$ has a critical point x_0
if $f'(x_0) = 0$.

↳ Second derivative test:

$$\begin{cases} f''(x_0) > 0, & \text{concave up} \Rightarrow \min \\ f''(x_0) < 0, & \text{concave down} \Rightarrow \max \end{cases}$$

$$f(x_0 + h) = f(x_0) + hf'(x_0) + \tfrac{1}{2}h^2 f''(x_0) + \cdots + \tfrac{1}{n!}h^n f^{(n)}(x_0) + \cdots$$
$$= f(x_0) + \tfrac{1}{2}h^2 f''(x_0) + R$$

the remainder $R = \frac{1}{3!}h^3 f^{(3)}(x^*)$, $x^* \in (x_0 - \epsilon,\ x_0 + \epsilon)$

2D case: (x_0, y_0) as a critical point:

$$f(x_0 + h,\ y_0 + k) \approx f(x_0, y_0) + \frac{1}{2}\underbrace{\left[h^2 f_{xx} + 2hk f_{xy} + k^2 f_{yy}\right]\Big|_{(x_0,y_0)}}_{\text{always } >0\ ?\ <0\ ?\ \forall\, h,k}$$

$$\frac{1}{2}\begin{bmatrix} h & k \end{bmatrix}\begin{bmatrix} f_{xx}(x_0,y_0) & f_{xy}(x_0,y_0) \\ f_{xy}(x_0,y_0) & f_{yy}(x_0,y_0) \end{bmatrix}\begin{bmatrix} h \\ k \end{bmatrix}$$

Where we define the **Hessian matrix** of $f(x,y)$ as

$$H(x,y) = \begin{bmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{bmatrix}$$

E.g. Let $f(x,y) = 3x^2 - xy + y^2$, find out the local extrema or saddle points.

Solution: ① Detect critical points: $\frac{\partial f}{\partial x} = 6x - y = 0$, $y_0 = 6x_0$; $\frac{\partial f}{\partial y} = -x + 2y$. $x_0 = 2y_0$

$\Rightarrow (x_0, y_0) = (0, 0)$

② Hessian matrix of f:

$$H(x,y) = \begin{bmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{bmatrix} = \begin{bmatrix} 6 & -1 \\ -1 & 2 \end{bmatrix}$$

Hessian matrix at the critical points $(0,0)$:

$$H(0,0) = \begin{bmatrix} 6 & -1 \\ -1 & 2 \end{bmatrix}$$ ↙ positive definite

$\Rightarrow H(0,0)$ is positive definite.

$(0,0)$ is a local minimum.

By eigenvalue test. $\lambda_1 = 1.7639$, $\lambda_2 = 6.2361$

By determinant test, $|H_1| = 6 > 0$, $|H| = 13 > 0$

<u>Thm</u> Suppose an everywhere infinitely differentiable function $F(\vec{x})$ in n variables has a critical point $\vec{x}_0$

Then (i) $\vec{x}_0$ is a local minimum of F if $H(\vec{x}_0)$ is positive definite.

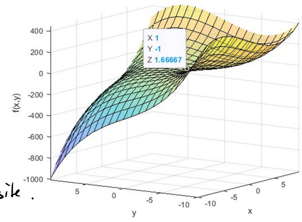
(ii) $\vec{x}_0$ is a local maximum of F if $H(\vec{x}_0)$ is negative definite.

(i) $\vec{x}_0$ is a local minimum of F if $H(\vec{x}_0)$ is positive definite.
(ii) $\vec{x}_0$ is a local maximum of F if $H(\vec{x}_0)$ is negative definite.
(iii) $\vec{x}_0$ is a saddle point of F if $H(\vec{x}_0)$ is indefinite.

Exercise: $f(x,y) = \frac{1}{3}x^3 - \frac{1}{3}y^3 + 3xy + 2x - 2y$ at $(1,-1)$ $\longrightarrow$ critical point $\to$ saddle

Hessian matrix $= \begin{bmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{bmatrix} = \begin{bmatrix} 2x & 3 \\ 3 & -2y \end{bmatrix}$

```
>> X=[-10:1:10]; Y=X;      % generate two uniformly-spaced vectors x and y
>> [x,y]=meshgrid(X,Y);  %  generate a grid with x and y
>> f=1/3*x.^3-1/3*Y.^3+3*x.*y+2*x-2*y; % compute the function value at different grid points
>> surf(x,y,f)                  % plot a 3-D graph
```



Thus $H(1,-1) = \begin{bmatrix} 2 & 3 \\ 3 & 2 \end{bmatrix}$ is

$\det(H) = 4 - 9 = -5 < 0$
$= \lambda_1 \cdot \lambda_2$

So the signs of λ_1 and λ_2 are opposite.

Exercise: $F(x,y,z) = x^2 + xz - 3\cos y + z^2$

Find the local extremum or saddle points and plot the graph using MATLAB!

Δ More about positive definite matrices

properties: ① If $A_{n\times n}$ is a symmetric positive definite matrix, then A is nonsingular.

A is positive definite $\Rightarrow$ all eigenvalues of A >0
$\Downarrow$
A is nonsingular $\Leftarrow$ $\det(A) = \lambda_1 \lambda_2 \cdots \lambda_n > 0$

② If A is a symmetric positive definite matrix, then $\det(A) > 0$.

③ If A is symmetric positive definite, then the leading principal submatrices of A are all positive definite.

$\vec{x}_r^T A_r \vec{x}_r > 0$ for $r = 1, 2, \cdots, n$

$$\underbrace{[\vec{x}_r^T \;\; \vdots\; 0 \cdots 0]}_{\vec{x}^T} \underbrace{\left[\begin{array}{ccc|ccc} a_{11} & a_{12} \cdots & a_{1r} & \cdots & & a_{1n} \\ \vdots & \ddots & \vdots & & & \vdots \\ a_{r1} & \cdots & a_{rr} & & & \\ \hline \vdots & & & \ddots & & \\ a_{n1} & & \cdots & & & a_{nn} \end{array}\right]_{n\times n}}_{A} \underbrace{\begin{bmatrix} \vec{x}_r \\ \hdashline 0 \\ \vdots \\ 0 \end{bmatrix}_{n\times 1}}_{\vec{x}} = \vec{x}_r^T A_r \vec{x}_r > 0$$

A_r (upper-left block); $\vec{x}_r \neq \vec{0}_{r\times 1}$

Since A is positive definite

Similarly, we can prove each A_r is positive definite for $r = 1, 2, \cdots, n$.

By property ③, then ①, ②: If A is symmetric & positive definite, then $\det(A_r) > 0$ for all r.

(optional) ④ If A is symmetric positive definite, then A can be reduced to upper triangular form using only row operation $\underline{\text{III}}$.
And the pivots will all be positive.

(optional) ⑤ $A = LL^T$ where L is lower triangular with positive diagonals.

⑥ $A = B^T B$ for some nonsingular matrix B.

Idea of proof in your textbook:

A positive definite $\Rightarrow$ ③ $\Rightarrow$ ④ $\Rightarrow$ ⑤ $\Rightarrow$ ⑥ $\Rightarrow$ ①

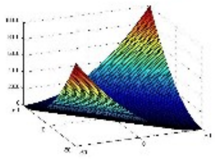
Idea of proof in your textbook:

A positive definite $\Rightarrow$ ③ $\Rightarrow$ ④ $\Rightarrow$ ⑤ $\Rightarrow$ ⑥ $\Rightarrow$ ①

Matlab code for plotting $z=x^2+2xy+y^2$ on domain [-50,50]X[-50,50]

```
>> x=[-50:2:50]; y=x;       % generate two uniformly-spaced vectors x and y
>> [X,Y]=meshgrid(x,y);  %  generate a grid with x and y
>> Z=X.^2+2*X.*Y+Y.^2; % compute the function value at different grid points
>> surf(X,Y,Z)               % plot a 3-D graph
```

Try to plot the other examples in the notes!



Chap6.6-6.7

1. Definitions of positive/negative definite/semidefinite, indefinite matrix
2. Different ways to identify the matrices
   a. Eigenvalues <===> positive, negative, indefinite
   b. Every leading principal submatrix is positive definite <=> det$|A_i|$>0 for all i <=>A positive definite
3. Applications in classification of stationary points
   a. Hessian matrix
      i. Positive definite <==> local min
      ii. Negative definite <==> local max
      iii. Indefinite <==> saddle