Lecture 6B

Eivind Eriksen — Sep 18, 2014 — GRA 6035 Mathematics

Plan:

1. Markov chains
2. Quadratic forms
3. Definiteness of symmetric matrices

Reading:

[ME] 6.2 (Ex 3), 23.1 (Ex 23.4), 23.6, 13.1–13.5, 16.1–16.4, 23.8

Review: Diagonalization

A $n\times n$-matrix

A diagonalizable if $P^{-1}AP = D$ is diagonal for some invertible matrix P.

Method:

Eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_r$ $(r \leq n)$

If $r = n$:

$$D = \begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix}$$

If $r < n$: A is not diagonalizable

Eigenvectors: $\underline{v_1}, \underline{v_2}, \ldots, \underline{v_k}$ (linearly independent)

If $k = n$:

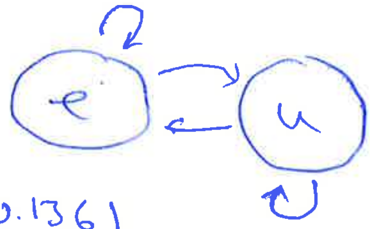
$$P = \begin{pmatrix} \underline{v_1} & \underline{v_2} & \cdots & \underline{v_n} \end{pmatrix}$$

If $k < n$: A is not diagonalizable

$$\boxed{\begin{array}{l} \text{Conclusion: If } r = n \text{ and } k = n, \text{ then} \\ \qquad P^{-1}AP = D \\ \text{and } A \text{ is diagonalizable.} \end{array}}$$

① Markov chains

Ex: employment - unemployment



Transition matrix: $A = \begin{pmatrix} 0.98 & 0.136 \\ 0.02 & 0.864 \end{pmatrix}$

Starting state: $\begin{pmatrix} e_0 \\ u_0 \end{pmatrix} = \begin{pmatrix} 0.9 \\ 0.1 \end{pmatrix} = \underline{x}_0$

After n time periods: $\boxed{A^n \cdot \underline{x}_0 = \underline{x}_n}$

$P^{-1}AP = D \quad | P \cdot$

$AP = PD \quad | \cdot P^{-1}$

$\underline{A = PDP^{-1}} \quad \Rightarrow A^n = (PD\cancel{P^{-1}}) \cdot (\cancel{P}D\cancel{P^{-1}}) \ldots (\cancel{P}DP^{-1})$

$$\boxed{A^n = P \cdot D^n \cdot P^{-1}}$$

In the Ex: $\lambda_1 = 1 \quad \lambda_2 = 0.844$

$D = \begin{pmatrix} 1 & 0 \\ 0 & 0.844 \end{pmatrix} \quad D^n = \begin{pmatrix} 1^n & 0 \\ 0 & 0.844^n \end{pmatrix}$

It is much easier to compute D^n than A^n.

Markov process

Ex: Families are classified as U (urban), S (suburban) and R (rural). At time $t=n$ (after n years), the share of families in these groups can be described by the state vector

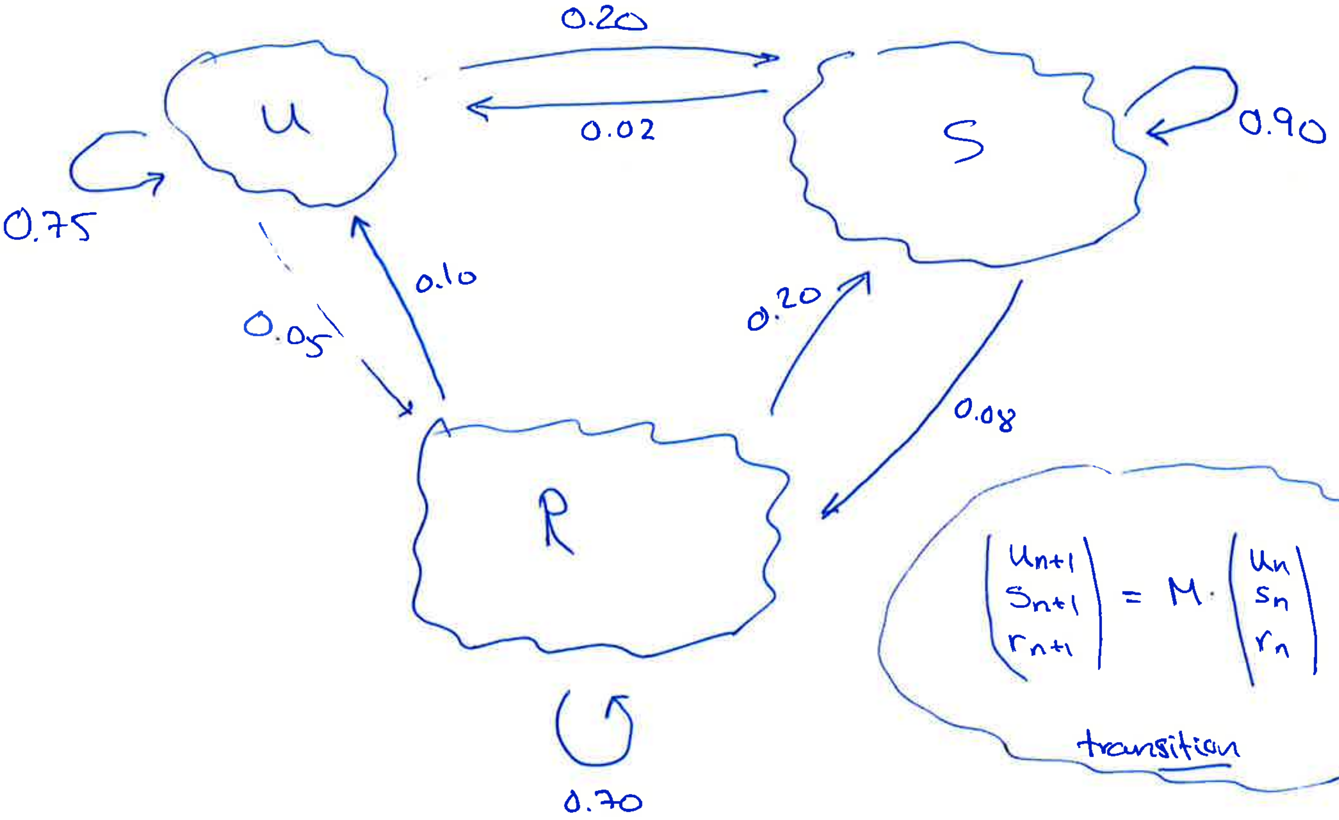
$$\underline{V}_n = \begin{pmatrix} u_n \\ s_n \\ r_n \end{pmatrix} \quad \begin{cases} u_n \geq 0 \\ s_n \geq 0 \\ r_n \geq 0 \end{cases}, \quad u_n + s_n + r_n = 1$$

Ex: $\underline{V} = \begin{pmatrix} 0.8 \\ 0.1 \\ 0.1 \end{pmatrix}$

From year n to year n+1, the change in the shares are given by a transition matrix or Markov matrix

$$M = \begin{pmatrix} 0.75 & 0.02 & 0.10 \\ 0.20 & 0.90 & 0.20 \\ 0.05 & 0.08 & 0.70 \end{pmatrix} \quad \begin{cases} m_{ij} \geq 0 \\ \text{each column has sum } 1 \end{cases}$$

It can be described graphically as follows:



$$\begin{pmatrix} u_{n+1} \\ s_{n+1} \\ r_{n+1} \end{pmatrix} = M \cdot \begin{pmatrix} u_n \\ s_n \\ r_n \end{pmatrix}$$

transition

Markov process:

$$\underbrace{\underline{V}_0 = \begin{pmatrix} u_0 \\ s_0 \\ r_0 \end{pmatrix}}_{\text{start}} \dashrightarrow \underline{V}_1 = M \cdot \underline{V}_0 \dashrightarrow \underline{V}_2 = M\underline{V}_1 = M^2 \underline{V}_0 \dashrightarrow \cdots \longrightarrow \underline{V}_n = M^n \cdot \underline{V}_0$$

The Markov process is regular if $m_{ij} > 0$ for all i,j. We assume that this the case. The following holds for all regular Markov processes:

Fact: i) $\lambda = 1$ is an eigenvalue of M, and there is a unique eigenvector $\underline{v}$ with eigenvalue $\lambda = 1$ that is a state vector (i.e. $\underline{v} = (v_i)$ with $v_i \geq 0$, $v_1 + \dots + v_k = 1$)

ii) $\lim_{n\to\infty} M^n \cdot \underline{v_0} = \underline{v}$ and $\lim_{n\to\infty} M^n = (\underline{v} | \underline{v} | \dots | \underline{v})$

Ex: $M = \begin{pmatrix} 0.75 & 0.02 & 0.10 \\ 0.20 & 0.90 & 0.20 \\ 0.05 & 0.08 & 0.70 \end{pmatrix}$

$\lambda = 1$: $\begin{pmatrix} -0.25 & 0.02 & 0.10 \\ 0.20 & -0.10 & 0.20 \\ 0.05 & 0.08 & -0.30 \end{pmatrix} \to \begin{pmatrix} 5 & 8 & -30 \\ -25 & 2 & 10 \\ 20 & -10 & 20 \end{pmatrix} \to \begin{pmatrix} 5 & 8 & -30 \\ 0 & 42 & -140 \\ 0 & -42 & 140 \end{pmatrix}$

$\to \begin{pmatrix} 5 & 8 & -30 \\ 0 & 42 & -140 \\ 0 & 0 & 0 \end{pmatrix}$

$5x + 8y - 30z = 0$
$42y - 140z = 0$
z free

$y = \frac{140z}{42} = \frac{10}{3}z$

$5x = 30 \cdot z - 8 \cdot \frac{10}{3}z = \frac{90-80}{3}z$

$x = \frac{2}{3}z$

$\frac{2}{3}z + \frac{10}{3}z + z = 1$
$5z = 1$
$z = 1/5$

$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2/3 \cdot z \\ 10/3 \cdot z \\ z \end{pmatrix} = z/3 \cdot \begin{pmatrix} 2 \\ 10 \\ 3 \end{pmatrix} \Rightarrow \underline{v} = \begin{pmatrix} 2/15 \\ 10/15 \\ 3/15 \end{pmatrix}$ (with $z = 1/5$)

Conclusion: As $n \to \infty$ (in the long run) $u = 2/15 \approx 13.3\%$ of families are urban, $s = 10/15 \approx 66.7\%$ are suburban, and $r = 3/15 = 20\%$ are rural.

Check: Compute M^{10}, M^{50}, M^{100} using Wolfram Alpha or other software.

It is also possible to compute M^n as

$$M^n = P \cdot \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}^n \cdot P^{-1} \approx P \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} P^{-1}$$

since $\lambda_1 = 1$, $\lambda_2, \lambda_3 < 1$

② Quadratic forms

A function $f(x_1, \dots, x_n)$ is called a quadratic form is $f(\underline{x})$ is a sum of terms of order two.

Ex: $f(x) = ax^2 \qquad (n=1)$

$f(x,y) = ax^2 + bxy + cy^2 \qquad (n=2)$

$$f(x_1, x_2, \dots, x_n) = a_{11}x_1^2 + a_{12}x_1x_2 + a_{13}x_1x_3 + \dots + a_{1n}x_1x_n + a_{22}x_2^2 + a_{23}x_2x_3 + \dots$$

Fact: $f(\underline{x}) = \underline{x}^t A \underline{x}$ where $\underline{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{pmatrix}$, A $n \times n$-matrix

Any quadr. form can be written like this with A symmetric (A unique)

$$= (x_1\ x_2 \dots x_n) \cdot \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & & a_{2n} \\ \vdots & & & \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

Ex: $n=2$

$$(x_1\ x_2) \cdot \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + a_{21}x_2 & a_{12}x_1 + a_{22}x_2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

$$= (a_{11}x_1 + a_{21}x_2)x_1 + (a_{12}x_1 + a_{22}x_2)x_2$$

$$= a_{11}x_1^2 + a_{21} \cdot x_1x_2 + a_{12}x_1x_2 + a_{22}x_2^2$$

$$= a_{11}x_1^2 + (a_{21} + a_{12})x_1x_2 + a_{22}x_2^2$$

Ex: $4x_1^2 + 6x_1x_2 - x_2^2 = (x_1 \; x_2) \cdot \begin{pmatrix} 4 & 3 \\ 3 & -1 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$

$= \underline{x}^t A \underline{x}$ with A symmetric

Ex: $x_1^2 + x_2^2 - 4x_2x_3 + x_3^2 - 6x_1x_3$

$$= (x_1 \; x_2 \; x_3) \cdot \begin{pmatrix} 1 & 0 & -3 \\ 0 & 1 & -2 \\ -3 & -2 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \underline{x}^t A \cdot \underline{x}$$

A is called <u>the symmetric matrix</u> of the quadratic form

---

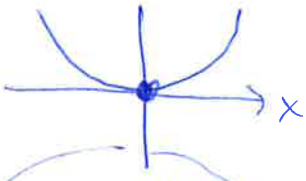
Definition:

Let $f(x_1, \dots, x_n)$ be a quadratic form, and let A be its symmetric matrix. Then both f and A are called

- positive definite $\iff f(\underline{x}) > 0$ for all $\underline{x} \neq \underline{0}$
- positive semidefinite $\iff f(\underline{x}) \geq 0$ for all $\underline{x}$
- negative definite $\iff f(\underline{x}) < 0$ for all $\underline{x} \neq 0$
- negative semidefinite $\iff f(\underline{x}) \leq 0$ for all $\underline{x}$
- indefinite $\iff$ neither positive or negative semidefinite, or in other words $f(\underline{x})$ take both positive and negative values

Ex: $n=1$ $\quad f(x) = ax^2$ $\quad A = (a)$

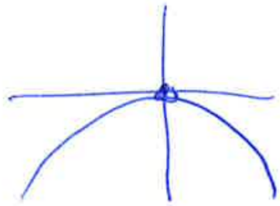
$a>0$:



positive definite

$x=0$ is a global minimum

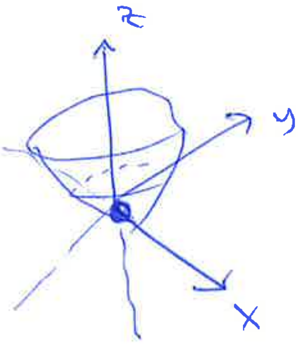
$a<0$:



~~pos~~ negative definite

$x=0$ is a global max.

Ex: $f(x,y) = x^2+y^2$



f is positive definite

$f(0,0)=0$

$f(x,y)>0$ if $(x,y) \neq (0,0)$

$f(x,y) = x^2-y^2$ is indefinite

$f(1,0)=1$

$f(0,1)=-1$

$f(x,y) = -x^2-2y^2$ is negative definite

When a quadratic form has only squares (no crossterms), i.e. A is diagonal:

$$f(x_1,\dots,x_n) = c_1 \cdot x_1^2 + c_2 x_2^2 + \dots + c_n \cdot x_n^2 \qquad A = \begin{pmatrix} c_1 & & 0 \\ & c_2 & \\ 0 & & \ddots \ c_n \end{pmatrix}$$

Then:

f positive definite: $c_1, c_2, \dots, c_n > 0$

f negative definite: $c_1, c_2, \dots, c_n < 0$

f positive semidefinite: $c_1, c_2, \dots, c_n \geq 0$

f negative semidefinite: $c_1, c_2, \dots, c_n \leq 0$

f is indefinite

both positive and negative c_i's

Fact: Classification of quadratic forms

If $f(x_1, \dots, x_n)$ is a quadratic form with symmetric matrix A, with eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$; then

$\lambda_1, \lambda_2, \dots, \lambda_n > 0$: f positive definite
$\lambda_1, \lambda_2, \dots, \lambda_n \geq 0$: f positive semidefinite
$\Big\}$ $\underline{x=0}$ is global $\underline{\text{min}}$

$\lambda_1, \lambda_2, \dots, \lambda_n < 0$: f negative definite
$\lambda_1, \lambda_2, \dots, \lambda_n \leq 0$: f negative semidefinite
$\Big\}$ $\underline{x=0}$ is global $\underline{\text{max}}$

Both positive and negative λ_i's : f indefinite $\Big\}$ $\underline{x=0}$ is saddle point

Why:

In general, you can rewrite

$$f(x_1, \dots, x_n) = \lambda_1 \cdot u_1^2 + \lambda_2 \cdot u_2^2 + \dots + \lambda_n \cdot u_n^2$$

where $u_1, u_2, \dots, u_n$ are linear combinations of the x_i's.

Ex: $$4xy = (\overset{u_1}{\overset{\shortparallel}{x+y}})^2 - (\overset{u_2}{\overset{\shortparallel}{x-y}})^2$$
$$= (\cancel{x^2} + 2xy + \cancel{y^2}) - (\cancel{x^2} - 2xy + \cancel{y^2})$$

Ex: $f(x,y,z) = x^2 - 6xz + 2y^2 + 2z^2$

$$A = \begin{pmatrix} 1 & 0 & -3 \\ 0 & 2 & 0 \\ -3 & 0 & 2 \end{pmatrix}$$

$$\begin{vmatrix} 1-\lambda & 0 & -3 \\ 0 & 2-\lambda & 0 \\ -3 & 0 & 2-\lambda \end{vmatrix} = (2-\lambda) \cdot \begin{vmatrix} 1-\lambda & -3 \\ -3 & 2-\lambda \end{vmatrix}$$

$$= (2-\lambda) \cdot (\lambda^2 - 3\lambda - 7) = 0$$

$\underline{\lambda = 2}$, $\lambda = \dfrac{3 \pm \sqrt{3^2 - 4 \cdot (-7)}}{2}$

$$\lambda = \frac{3 \pm \sqrt{37}}{2}$$

f is indefinite $\Longleftarrow$ $\lambda_1 = 2$ (>0) $\lambda_2 = \dfrac{3+\sqrt{37}}{2}$ (>0) $\lambda_3 = \dfrac{3-\sqrt{37}}{2}$ (<0)

Ex: $f(\underline{x}) = 2x_1^2 - 4x_1x_3 + 7x_2^2 - 14x_1x_4 + x_4^2$

$$A = \begin{pmatrix} 2 & 0 & -2 & -7 \\ 0 & 7 & 0 & 0 \\ -2 & 0 & 0 & 0 \\ -7 & 0 & 0 & 1 \end{pmatrix}$$

<u>Eigenvalues:</u>

$$\begin{vmatrix} 2-\lambda & 0 & -2 & -7 \\ 0 & 7-\lambda & 0 & 0 \\ -2 & 0 & -\lambda & 0 \\ -7 & 0 & 0 & 1-\lambda \end{vmatrix} = 0$$

$$(7-\lambda) \cdot \begin{vmatrix} 2-\lambda & -2 & -7 \\ -2 & -\lambda & 0 \\ -7 & 0 & 1-\lambda \end{vmatrix} = 0$$

$\underline{\lambda_1 = 7}$,

$$2 \cdot \big(-2(1-\lambda)\big) - \lambda \cdot (\lambda^2 - 3\lambda - 47) = 0$$

$$-4(1-\lambda) - \lambda(\lambda^2 - 3\lambda - 47) = 0$$

$$-\lambda^3 + \dots + (-4)$$

$$(7-\lambda) \cdot (-\lambda^3 + \dots - 4) = \lambda^4 + \dots - 28$$

$$|A| = -28 = \underset{\substack{\parallel \\ 7}}{\lambda_1} \cdot \lambda_2 \cdot \lambda_3 \cdot \lambda_4$$

} At least one of $\lambda_2, \lambda_3, \lambda_4$ is negative

$\Downarrow$

<u>f is indefinit</u>

Method using principal minors

A $n\times n$ matrix, symmetric

A leading principal minor of A of order i is the minor obtained by selecting the first i rows and i colums. We call it D_i.

Ex: $A = \begin{pmatrix} 1 & 0 & -3 \\ 0 & 2 & 0 \\ -3 & 0 & 2 \end{pmatrix}$

$D_1 = 1 > 0$

$D_2 = \begin{vmatrix} 1 & 0 \\ 0 & 2 \end{vmatrix} = 2 > 0$

$D_3 = |A| = \begin{vmatrix} 1 & 0 & -3 \\ 0 & 2 & 0 \\ -3 & 0 & 2 \end{vmatrix}$

$= 2 \cdot (1 \cdot 2 - (-3)^2) = -14 < 0$

Fact:

i) If $D_1, D_2, \ldots, D_n > 0$, then A is positive definite

ii) If $D_1 < 0$, $D_2 > 0$, $D_3 < 0$, ... then A is negative definite

iii) If $D_1, D_2, \ldots, D_n$ fails both patterns above, and the reason for the failure is wrong sign (not zero), then A is indefinite.

In the example, A is indefinite.

Ex:

$$A = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -3 \end{pmatrix}$$

$$f = -x^2 - 2y^2 - 3z^2$$

$\lambda_1 = -1,\ \lambda_2 = -2,\ \lambda_3 = -3$

negative definite

$D_1 = -1 < 0$

$D_2 = (-1)\cdot(-2) = 2 > 0$

$D_3 = (-1)(-2)\cdot(-3) = -6 < 0$

If at least one of ~~the~~ the leading priciple minors are zero, you have to look at all principle minors.

Ex:

$$A = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -3 \end{pmatrix}$$

$D_1 = -1$

$D_2 = 0$

$D_3 = 0$

$\lambda_1 = -1$

$\lambda_2 = 0$

$\lambda_3 = -3$

A principle minor of order i is a minor obtained by keeping i rows and the same i columns. We call the Δ_i.

Ex:

$$A = \begin{pmatrix} 1 & 0 & -3 \\ 0 & 2 & 0 \\ -3 & 0 & 2 \end{pmatrix} = \begin{pmatrix} \boxed{1} & 0 & -3 \\ 0 & \boxed{2} & 0 \\ -3 & 0 & \boxed{2} \end{pmatrix}$$

$\Delta_1 = 1,\ 2,\ 2$

$\Delta_2 = 2,\ \begin{vmatrix} 1 & -3 \\ -3 & 2 \end{vmatrix} = -7,\ \begin{vmatrix} 2 & 0 \\ 0 & 2 \end{vmatrix} = 4$

$\Delta_3 = -14$

$D_1 = 1$

$D_2 = 2$

$D_3 = |A| = -14$

Fact:

If all principal minors $\Delta_1, \Delta_2, \dots, \Delta_n \geq 0$, then A is positive semidefinite

If $\Delta_1 \leq 0$, $\Delta_2 \geq 0$, $\Delta_3 \leq 0, \dots$ for all principal minors Δ_i, then A is negative semidefinite

Ex: $f(x,y,z) = x^2 + 2xy + y^2 + z^2$

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$D_1 = 1$

$D_2 = \begin{vmatrix} 1 & 1 \\ 1 & 1 \end{vmatrix} = 0$

$D_3 = \begin{vmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = 1 \cdot \begin{vmatrix} 1 & 1 \\ 1 & 1 \end{vmatrix} = 0$

need all principal minors

Principal minors:

$$\left.\begin{aligned} \Delta_1 &= 1, 1, 1 \\ \Delta_2 &= 0, 1, 1 \\ \Delta_3 &= 0 \end{aligned}\right\}$$

$\Delta_1, \Delta_2, \Delta_3 \geq 0$ for all principal minors

$\Rightarrow$ positive semidefinite