LECTURE 5 (F)

GRA 6035 MATHEMATICS

EIVIND ERIKSEN

SEP 22, 2014

Plan:

1. Markov chains
2. Quadratic forms
3. Definiteness of quadratic forms

Reading:

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

Review:

A $n \times n$ matrix

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

Method:

Eigenvalues $\lambda_1, \dots, \lambda_r$ of A $(r \leq n)$

If $r = n$:

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

If $r < n$: A is not diagonalizable

Linearly independent eigenvectors $\underline{v}_1, \underline{v}_2, \dots, \underline{v}_k$

If $k = n$:

$$P = \left(\underline{v}_1 \mid \underline{v}_2 \mid \cdots \mid \underline{v}_n \right)$$

If $k < n$: A is not diagonalizable

Conclusion:

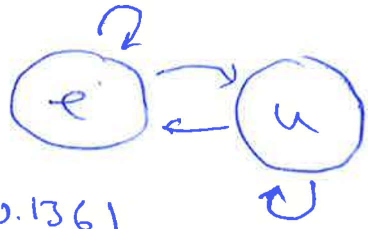
If $r = n$ and $k = n$, then

$$P^{-1}AP = D$$

and A is diagonalizable

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

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

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

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

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

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

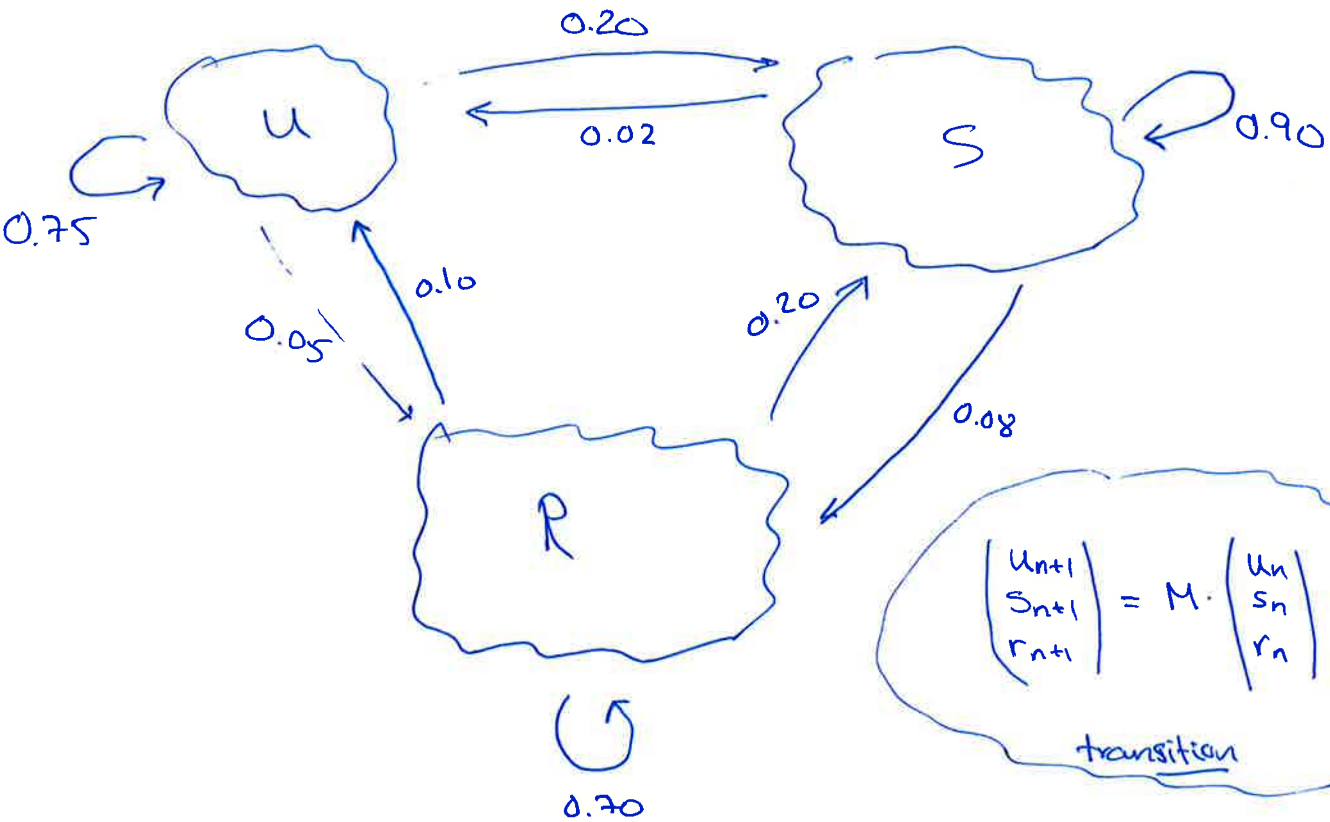
$$\underline{V}_n = \begin{pmatrix} u_n \\ s_n \\ r_n \end{pmatrix} \quad \begin{cases} u_n \geq 0 \\ s_n \geq 0 \\ r_n \geq 0 \end{cases}, \quad u_n + s_n + r_n = 1$$

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:

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

start

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: $M = \begin{pmatrix} 0.75 & 0.02 & 0.10 \\ 0.20 & 0.90 & 0.20 \\ 0.05 & 0.08 & 0.70 \end{pmatrix}$ $\quad D = \begin{pmatrix} 1 & & \\ & \lambda_2 & \\ & & \lambda_3 \end{pmatrix}$

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

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

$5x + 8y - 30z = 0$

$42y - 140z = 0$

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) $u = 2/15 \simeq 13.3\%$ of families are urban, $s = 10/15 \simeq 66.7\%$ 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

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

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

2 Quadratic forms

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

Ex:

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

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

$n=3$: $f(x,y,z) = ax^2 + bxy + cxz + dy^2 + eyz + fz^2$

Quadratic forms in matrix notation:

Ex:
$$(x \; y) \cdot \begin{pmatrix} 1 & 2 \\ 3 & 7 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} \qquad \underline{x} = \begin{pmatrix} x \\ y \end{pmatrix}$$
$$= (x+3y \quad 2x+7y) \cdot \begin{pmatrix} x \\ y \end{pmatrix}$$
$$= (x+3y)x + (2x+7y) \cdot y \qquad \underline{x}^t \cdot A \cdot \underline{x}$$
$$= x^2 + 3yx + 2xy + 7y^2$$
$$= \underline{x^2 + 5xy + 7y^2} \qquad A = \begin{pmatrix} 1 & 5/2 \\ 5/2 & 7 \end{pmatrix}$$

Fact: Any quadratic form can be written as $\underline{x}^t A \underline{x}$. It is possible to choose A to be symmetric, and then A is unique.

Quadratic forms in n var's $\rightleftarrows$ symmetric $n \times n$-matrices

$$f(\underline{x}) \longmapsto \underline{x}^t A \cdot \underline{x}$$

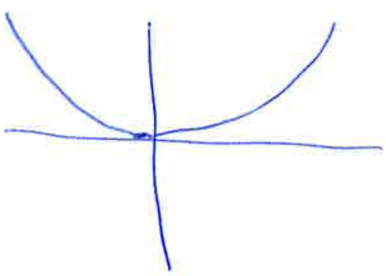
$\uparrow$ the symmetric matrix of the quadratic form

Ex: $f(x,y,z) = x^2 + 2xy + 4y^2 + 3yz - z^2$

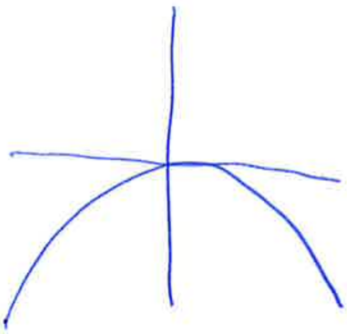
$$A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 4 & 3/2 \\ 0 & 3/2 & -1 \end{pmatrix}$$ is the symmetric matrix of the quadratic form.

$a_{ij} + a_{ji} =$ coeff. in front of $x_i x_j$

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

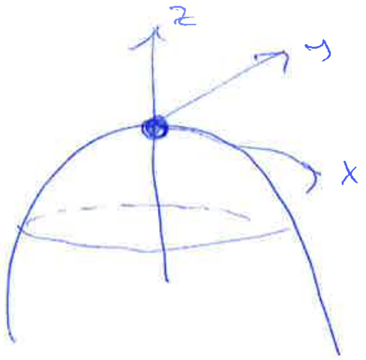


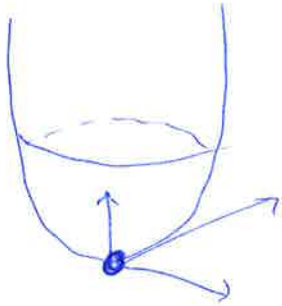
$a>0$



$a<0$

Ex: $n=2$





$f(x,y) = -x^2 - y^2$ negative semidefn.

$f(x,y) = x^2 + y^2$ positive semidefn.

Defn: $\boxed{\text{Remember: } f(x_1,\dots,x_n) = 0 \text{ !}}$

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

1. positive semidefinite $\iff$ $f(x_1,\dots,x_n) \geq 0$ for all $\underline{x}$
2. negative semidefinite $\iff$ $f(\underline{x}) \leq 0$ for all $\underline{x}$
3. indefinite $\iff$ neither pos./neg. semidefinite = f takes both pos. and neg. values

positive definite $\iff$ $f(\underline{x}) > 0$ for all $\underline{x} \neq \underline{0}$

negative definite $\iff$ $f(\underline{x}) < 0$ for all $\underline{x} \neq \underline{0}$

f pos. semidefinite : $\underline{x} = \underline{0}$ is global minimum

f neg. semidefinite : $\underline{x} = \underline{0}$ is global maximum

f indefinite : $\underline{x} = \underline{0}$ is saddle point

③ Definiteness of symmetric matrices

Ex: A is diagonal

$$A = \begin{pmatrix} a_{11} & & & 0 \\ & a_{22} & & \\ & & \ddots & \\ 0 & & & a_{nn} \end{pmatrix}$$

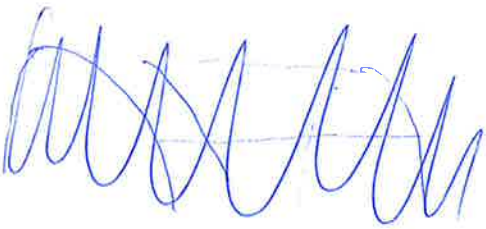
$f(x_1, \dots, x_n)$
$= a_{11}x_1^2 + a_{22}x_2^2 + \dots + a_{nn}x_n^2$

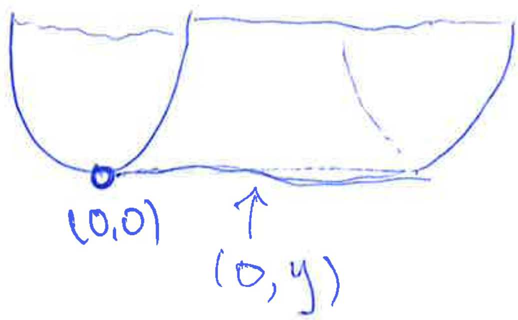
$a_{11}, a_{22}, \dots, a_{nn} \geq 0$: f positive semidefinite
$a_{11}, a_{22}, \dots, a_{nn} > 0$: f positive definite

$a_{11}, a_{22}, \dots, a_{nn} \leq 0$: f negative semidefinite
$a_{11}, a_{22}, \dots, a_{nn} < 0$: f negative definite

otherwise = both positive and negative numbers among $a_{11}, \dots, a_{nn}$: f indefinite

Ex: $f(x,y) = 2x^2 + 0 \cdot y^2$ $\begin{cases} \text{pos. semidefn.} \\ \underline{\text{not}} \text{ pos. defn.} \end{cases}$





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

$$= \underbrace{\tfrac{1}{2}x^2 + xy + \tfrac{1}{2}y^2}_{\substack{\frac{1}{2}(x^2+2xy+y^2) \\ =\frac{1}{2}(x+y)^2}} + \tfrac{1}{2}x^2 + \tfrac{1}{2}y^2$$

$\frac{3}{2}u_1^2 + \frac{1}{2}u_2^2$

pos. semidefn. $\longrightarrow$ $= \frac{1}{2}(x+y)^2 + \frac{1}{2}x^2 + \frac{1}{2}y^2$

Fact:

Let $f(x_1,\dots,x_n)$ be a quadr. form with symm. matrix A, and let $\lambda_1, \lambda_2, \dots, \lambda_n$ be the eigenvalues of A.

A pos. defn. $\iff \lambda_1,\dots,\lambda_n > 0$

A pos. semidefn. $\iff \lambda_1,\dots,\lambda_n \geq 0$

A neg. defn. $\iff \lambda_1,\dots,\lambda_n < 0$

A neg. semidefn. $\iff \lambda_1,\dots,\lambda_n \leq 0$

A indefinite $\iff$ there are both pos. and neg. eigenvalues

Ex: $f(x,y) = x^2 + xy + y^2 \rightsquigarrow A = \begin{pmatrix} 1 & 1/2 \\ 1/2 & 1 \end{pmatrix}$

$\lambda^2 - 2\lambda + 3/4 = 0$

$\lambda = \dfrac{2 \pm \sqrt{2^2 - 4 \cdot 3/4}}{2}$

$= 1 \pm 1/2$

$= 3/2,\ 1/2 > 0$

$$\boxed{\begin{gathered}\lambda_1, \lambda_2 > 0 \\ \Downarrow \\ A \text{ pos. defn.}\end{gathered}}$$

Why does this work?

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

where $u_1, u_2, \dots, u_n$ are linear comb. of $x_1, \dots, x_n$.

Ex: $x^2 - 6xz + 2y^2 + 2z^2 \rightsquigarrow A = \begin{pmatrix} 1 & 0 & -3 \\ 0 & 2 & 0 \\ -3 & 0 & 2 \end{pmatrix}$

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

$2-\lambda = 0$ or $\lambda^2 - 3\lambda - 7 = 0$

$\lambda_1 = 2 > 0$ $\quad \lambda = \dfrac{3 \pm \sqrt{9+28}}{2}$

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

$\lambda_2 = \dfrac{3+\sqrt{37}}{2} > 0$

$\lambda_3 = \dfrac{3-\sqrt{37}}{2} < 0$

indefinite

$(\lambda_1 > 0, \lambda_2 > 0, \lambda_3 < 0)$

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

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

Eigenvalues:

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

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

$\lambda_1 = 7$,

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

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

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

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

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

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

$\Downarrow$

f is indefinit

Method using principal minors

A: $n\times n$ symmetric matrix

The <u>leading principal minor</u> D_i of order i is obtained by keeping rows $1,2,\dots,i$ and columns $1,2,\dots,i$.

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

$D_1 = 1$

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

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

$= 2\cdot(-7) = \underline{-14}$

<u>Fact</u>:

If $D_1, D_2, \dots, D_n > 0$ then A is positive definite

If $D_1, D_2, \dots, D_n \geq 0$ then A <u>maybe</u> positive semidefinite (but we must check principal minors)

If $\underbrace{D_1<0,\ D_2>0,\ D_3<0,\ \dots,\ (-1)^n\cdot D_n>0}_{(-1)^i\cdot D_i>0 \text{ for } i=1,2,\dots,n}$, then A is <u>negative defn.</u>

If $D_1\leq 0,\ D_2\geq 0,\ D_3\leq 0,\ \dots,\ (-1)^n D_n\geq 0$, then A <u>may be</u> negative semidefn. (but we must check all principal minors)

All other cases, A is indefinite.

Ex:

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

$\lambda_1 = -1$

$\lambda_2 = -2$

$\lambda_3 = -3$

$D_1 = \lambda_1 = -1$

$D_2 = \lambda_1 \lambda_2 = 2$

$D_3 = \lambda_1 \lambda_2 \lambda_3 = -6$

negative definite

Ex:

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

$\lambda_1 = -1$

$\lambda_2 = 0$

$\lambda_3 = a$

$D_1 = -1$

$D_2 = 0$

$D_3 = 0$

$a \leq 0$ means neg. semidefn.

If you cannot decide the definiteness of A using leading principal minors (because at least one is zero), then you compute all principal minors:

A principal minor Δ_i of order i is obtained by keeping i rows and the same i columns.

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

$\Delta_1 = -1, \; 0, \; a$

$\Delta_2 = 0, \quad \begin{vmatrix} -1 & 0 \\ 0 & a \end{vmatrix} = -a, \quad \begin{vmatrix} 0 & 0 \\ 0 & a \end{vmatrix} = 0$

$\Delta_3 = 0$

$D_1 = -1$

$D_2 = 0$

$D_3 = 0$

A neg. semidefn. $\Updownarrow$ $a \leq 0$

Result:

If $\Delta_i \geq 0$ for all principal minors Δ_i of order i for $i = 1, 2, \dots, n$, then A is positive semidefinite.

If $\Delta_i \cdot (-1)^i \geq 0$ for all principal minors Δ_i of order i for $i = 1, 2, \dots, n$ (i.e. all $\Delta_1 \leq 0$, all $\Delta_2 \geq 0, \dots$) then A is negative semidefinite.

Ex: $x^2 + 2xy + y^2 + z^2 \rightsquigarrow A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

$D_1 = 1$

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

$D_3 = 1 \cdot D_2 = 0$

must compute Δ_i to find out if it is pos. semidefn. or indefinite

$$\left.\begin{aligned} \Delta_1 &= 1, 1, 1 \\ \Delta_2 &= 0, 1, 1 \\ \Delta_3 &= 0 \end{aligned}\right\} \text{all } \Delta_i \geq 0 \Downarrow \text{positive semidefn.}$$