Plan

Review Lecture 4

- eigenvalues / eigen-vectors
- diagonalization

1 Markov chains

Characteristics

a) n states

$$\underline{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} \quad \text{state vector} \quad 0 \le v_i \le 1, \quad v_1 + v_2 + \dots + v_n = 1$$

"shares of the population in the different states"

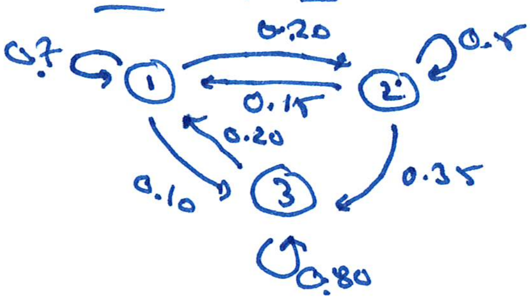
b) transition between states

$\underline{v_t}$: state vector at time t

$\underline{v_{t+1}} = A \cdot \underline{v_t}$ $\quad A = (a_{ij})$ transition matrix

$$\begin{cases} 0 \le a_{ij} \le 1 \\ \text{col. sums in } A \text{ are } 1 \end{cases}$$

Ex: $n = 3$



$$A = \begin{array}{c} \\ \begin{pmatrix} 0.7 & 0.15 & 0.2 \\ 0.2 & 0.5 & 0 \\ 0.1 & 0.35 & 0.8 \end{pmatrix} \\ \begin{array}{ccc} 1 & 2 & 3 \end{array} \end{array} \begin{array}{c} 1 \\ 2 \\ 3 \end{array}$$

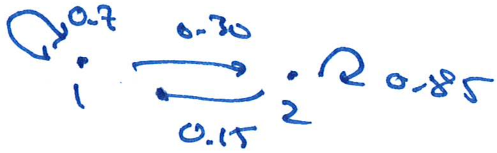
regular $\quad A^2 > 0$
$(t = 2)$

After t time periods: $\underline{v_t} = A^t \cdot \underline{v_0}$

$$\lim_{t \to \infty} \underline{v_t} = \left(\lim_{t \to \infty} A^t \right) \cdot \underline{v_0}$$

Defn: A Markov chain is called regular if there is a $t \ge 1$ such that you can from any state to any other state in t time periods, i.e. $\underline{A^t > 0}$

Ex. 1 $A = \begin{pmatrix} 0.70 & 0.15 \\ 0.30 & 0.85 \end{pmatrix}$



$A > 0$ means all entries $a_{ij} > 0$ $\Rightarrow$ regular

regular $t=1$

Thm (Perron – Frobenius)

If A is the transition matrix of a regular Markov chain, then we have:

i) $\lambda = 1$ is an eigenvalue of A (and all other eigenvalues have $|\lambda| < 1$)

ii) There is a unique eigenvector $\underline{v}$ in E_1 that is a state vector. Moreover, $\lim_{t\to\infty} A^t \cdot \underline{v_0} = \underline{v}$ for any state vector $\underline{v_0}$.

($\underline{v}$ is the eq. state / steady state, and is independent of $\underline{v_0}$)

Ex: $A = \begin{pmatrix} 0.7 & 0.15 & 0.2 \\ 0.2 & 0.5 & 0 \\ 0.1 & 0.35 & 0.8 \end{pmatrix}$

(same as before, regular)

$\lambda = 1$ eigenvalue:

E_1: $\begin{pmatrix} -0.3 & 0.15 & 0.2 \\ 0.2 & -0.5 & 0 \\ 0.1 & 0.35 & -0.2 \end{pmatrix}$

$\rightarrow \begin{pmatrix} -0.1 & -0.35 & 0.2 \\ 0.2 & -0.5 & 0 \\ 0.1 & 0.35 & -0.2 \end{pmatrix}$

$\rightarrow \begin{pmatrix} -0.1 & -0.35 & 0.2 \\ 0 & -1.2 & 0.4 \\ 0 & 0 & 0 \end{pmatrix}$

z free

$-1.2y + 0.4z = 0 \quad y = \frac{-0.4z}{-1.2} = \frac{1}{3}z$

$-0.1x - 0.35\left(\frac{1}{3}z\right) + 0.2z = 0 \quad \cdot 100$

$-10x = \frac{35}{3}z - 20z = -\frac{25}{3}z$

$x = \frac{25}{30}z = \frac{5}{6}z$

$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 5z/6 \\ z/3 \\ z \end{pmatrix} = \frac{z}{6}\underbrace{\begin{pmatrix} 5 \\ 2 \\ 6 \end{pmatrix}}_{= \underline{v_1}}$:

$E_1 = \text{span}(\underline{v_1})$

$5+2+6 = 13 \Rightarrow \underline{v} = \frac{1}{13}\begin{pmatrix} 5 \\ 2 \\ 6 \end{pmatrix} = \begin{pmatrix} 5/13 \\ 2/13 \\ 6/13 \end{pmatrix}$ eq. state

② Definiteness of quadratic forms

Defn A function $f(x_1, x_2, \ldots, x_n)$ is called a quadratic form if f is polynomial where each term has degree 2.

Ex: $f(x,y,z) = x^2 + 4xy - 6xz + 5y^2 - 2yz + z^2$

Result: Any quadratic form in n variables can be expressed uniquely as

$$\underline{x}^T A \underline{x}, \text{ where } \underline{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

and A is a symmetric $n \times n$ matrix.

$f = \ldots + c_i x_i^2 + \ldots + c_{ij} x_i x_j$

$\Downarrow$

$A = (a_{ij})$ $\quad a_{ii} = c_i$ $\quad a_{ij} + a_{ji} = c_{ij}$

Ex:

$$A = \begin{array}{c} \begin{pmatrix} 1 & 2 & -3 \\ 2 & 5 & -1 \\ -3 & -1 & 1 \end{pmatrix} \begin{array}{cc} 1 & x \\ 2 & y \\ 3 & z \end{array} \\ \begin{array}{ccc} 1 & 2 & 3 \\ x & y & z \end{array} \end{array}$$

Symmetric matrix of $f(x,y,z) = x^2 + 4xy - 6xz + 5y^2 - 2yz + z^2$

$2xy + 2yx = 4xy$

Definiteness:

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

Ex: $f(x,y,z) = x^2 + 2y^2 + z^2$ positive defn.

since $f(x,y,z) > 0$ for $(x,y,z) \neq (0,0,0)$.

$f(x,y,z) = x^2 - y^2 + 3z^2$ indefinite

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

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

Defn.: $f(\underline{x}) = f(x_1, \dots, x_n)$ quadr. form.

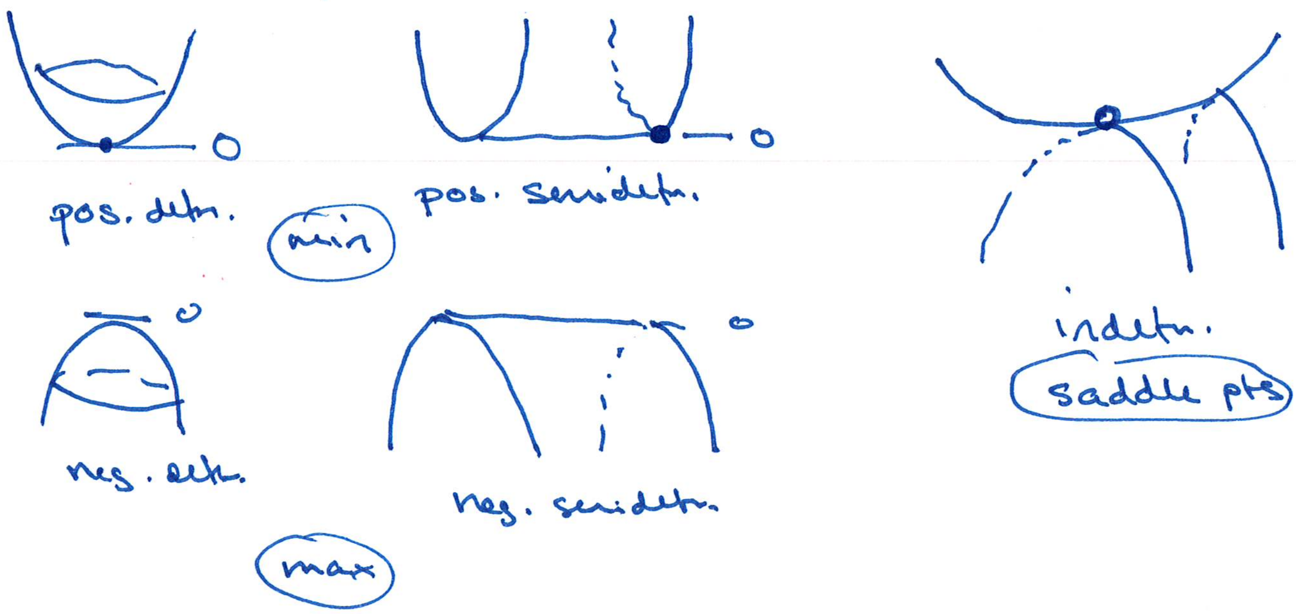
f is called positive semidefn. if $f(\underline{x}) \geq 0$ for all $\underline{x}$

— " — negative — " — $f(\underline{x}) \leq 0$ — " —

— " — indefinite otherwise

f is called positive defn. if $f(\underline{x}) > 0$ for $\underline{x} \neq \underline{0}$

— " — negative defn. $f(\underline{x}) < 0$ — " —



Methods to determine the definiteness:

$f(\underline{x}) = f(x_1, \ldots, x_n)$ quadr. form

A $n\times n$ symmetric matrix with eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$.

Thm:

(1) $\lambda_1, \lambda_2, \ldots, \lambda_n \geq 0 \iff$ pos. semidefin.

(2) $\lambda_1, \lambda_2, \ldots, \lambda_n \leq 0 \iff$ neg —"—

(3) Both pos. and neg. eigenvalues $\iff$ indefin.

$\lambda_1, \ldots, \lambda_n > 0 \iff$ pos. defin.

$\lambda_1, \ldots, \lambda_n < 0 \iff$ neg. —"—

Ex: $f(x,y,z) = x^2 + 4xy - 6xz + 5y^2 - \underset{2yz}{\cancel{}} + z^2$

$$A = \begin{pmatrix} 1 & 2 & -3 \\ 2 & 5 & -1 \\ -3 & -1 & 1 \end{pmatrix}$$

$$-\lambda^3 + c_1\lambda^2 - c_2\lambda + c_3 = 0$$

$c_1 = tr(A) = 7$

$c_3 = \det(A) = 1\cdot 4 - 2(-1) - 3\cdot 13 = -33$

$\lambda_1 + \lambda_2 + \lambda_3 = tr(A) = 7$

$\lambda_1 \cdot \lambda_2 \cdot \lambda_3 = \det(A) = -33$

pos. and neg. eigenvalues $\Rightarrow$ f is indefin

Principal minors of A

Defn An r-minor of A is principal Δ_r if we choose the same rows and cols, and leading principal D_r if we choose the first r row/cols.

Ex: $A = \begin{pmatrix} 1 & 2 & -3 \\ 2 & 5 & -1 \\ -3 & -1 & 1 \end{pmatrix}$

$D_1 = 1$
$D_2 = 1$
$D_3 = |A| = -33$

leading principal minor

$M_{12,12}$ $M_{23,23}$ $M_{13,13}$

$\Delta_1 = 1, 5, 1$
$\Delta_2 = 1, 4, -8$
$\Delta_3 = -33$

principal minors

Ex: $f = -x^2 - y^2 - 2z^2$ neg. defn

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

$D_1 = -1$
$D_2 = (-1)(-1) = 1$
$D_3 = (-1)(-1)(-2) = -2$

Result 1:

$D_1, D_2, \ldots, D_n > 0 \iff$ pos. defn.

$D_1 < 0, D_2 > 0, \ldots \iff$ neg. defn. $\quad (-1)^i D_i > 0 \quad i = 1, 2, \ldots, n$

Result 2:

$\Delta_1, \Delta_2, \ldots, \Delta_n \geq 0 \iff$ pos. semidefn.

$\Delta_1 \leq 0, \Delta_2 \geq 0, \ldots \iff$ neg. semidefn. $\quad (-1)^i \Delta_i \geq 0 \quad i = 1, 2, \ldots, n$

otherwise $\iff$ indefn.

Ex. 1:

$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 5 & 7 \\ 3 & 7 & 10 \end{pmatrix}$ $\quad D_1 = 1 \quad \Delta_1 = 1, 5, 10$; $\quad D_2 = 1 \quad \Delta_2 = 1, 1, 1$; $\quad D_3 = 0 \quad \Delta_3 = 0$

pos. semidefn

Ex. 2

$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 3 & 6 & 8 \end{pmatrix}$ $\quad D_1 = 1 \quad \Delta_1 = 1, 4, 8$; $\quad D_2 = 0 \quad \Delta_2 = 0, -4, -1$; $\quad D_3 = 0 \quad \Delta_3 = 0$

indefinite

Result 3 Reduced rank criterion

If $\operatorname{rk} A = r < n$ and $D_1, D_2, \ldots D_r > 0$ then A pos. semidefn.

— " — $< n$ and $D_1 < 0, D_2 > 0, \ldots$ (up to r) then A neg. semidefn.

Ex:

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

$\operatorname{rk} A = 2$

$$\left.\begin{array}{l} D_1 = 1 \\ D_2 = 2 \\ D_3 = 0 \\ D_4 = 0 \end{array}\right\} \underset{\Longrightarrow}{\text{RRC}} \text{ pos. semidefn.}$$

(3) Orthogonal diagonalization

Defn An invertible matrix P is called orthogonal if $P^{-1} = P^T$. An orthogonal diagonalization of A is a diag. $P^{-1}AP = D$ such that P is orthogonal.

$(P^T A P = D)$

Facts: (1) $\underset{n \times n}{A}$ is orth. diagonalizable $\Longleftrightarrow$ A symmetric.

(2) $P = (\underline{v_1} | \underline{v_2} | \ldots | \underline{v_n})$ is orthogonal

$\Updownarrow$

(a) $\|v_i\| = 1$ and (b) $\underline{v_i} \cdot \underline{v_j} = 0$ for $i \neq j$

$P^{-1} = P^T$

$\Updownarrow$

$P^T \cdot P = I$

$\Updownarrow$

$$\begin{pmatrix} \underline{v_1} \\ \vdots \\ \underline{v_n} \end{pmatrix} \cdot (\underline{v_1} | \ldots | \underline{v_n}) = I$$

$$\begin{pmatrix} \underline{v_1} \cdot \underline{v_1} & \underline{v_1} \cdot \underline{v_2} & \cdots \\ \vdots & \vdots & \end{pmatrix}$$

(3) $\underline{v_i}, \underline{v_j}$ are eigenvectors of A (symm) with different eigenval. then $\underline{v_i} \cdot \underline{v_j} = 0$

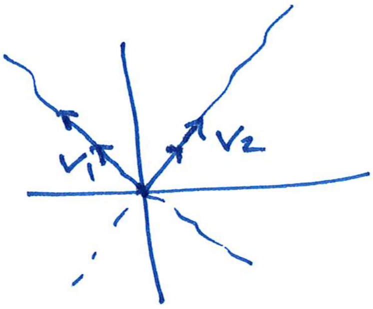
Eivind Eriksen (eivind.eriksen@bi.no), Office B3y-085
https://www.dr-eriksen.no/teaching/GRA6035/
https://www.dr-eriksen.no/teaching/ELE3781/

Ex 1: Easy case

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \quad \lambda_1 = -1 \quad \underline{v_1} = \begin{pmatrix} -1 \\ 1 \end{pmatrix} \qquad D = \begin{pmatrix} -1 & 0 \\ 0 & 3 \end{pmatrix} \quad P = \begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix}$$

$$\lambda_2 = 3 \quad \underline{v_2} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

ord. diag.



$$\underline{v_1} \cdot \underline{v_2} = 0$$

normalization

$$\|\underline{v_1}\| = \sqrt{2} \;\rightarrow\; \frac{1}{\sqrt{2}}\begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

$$\|\underline{v_2}\| = \sqrt{2} \;\rightarrow\; \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$D = \begin{pmatrix} -1 & 0 \\ 0 & 3 \end{pmatrix} \quad P = \begin{pmatrix} -1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix}$$

orth. diag.

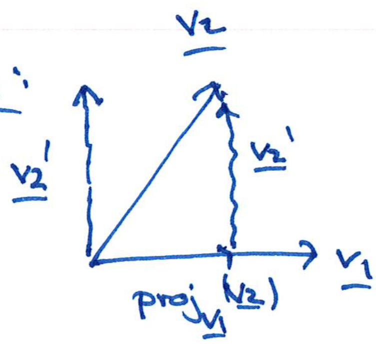
Ex. 2

$$A = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix} \quad \begin{matrix} \lambda_1 = 1 \\ \lambda_2 = 1 \\ \lambda_3 = 4 \end{matrix} \quad v_1 = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \quad \underline{v_2} = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \quad \underline{v_3} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

$$\underline{v_1} \cdot \underline{v_3} = \underline{v_2} \cdot \underline{v_3} = 0$$

$$\underline{v_1} \cdot \underline{v_2} = 1 \neq 0$$

$\underline{E_1}$:



change $\underline{v_2}$ to $\underline{v_2}' = \underline{v_2} - \mathrm{proj}_{\underline{v_1}}(\underline{v_2})$

$$= \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} - \frac{\underline{v_2} \cdot \underline{v_1}}{\underline{v_1} \cdot \underline{v_1}} \cdot \underline{v_1}$$

$$= \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} - \frac{1}{2} \cdot \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}$$

$$= \begin{pmatrix} -1/2 \\ -1/2 \\ 1 \end{pmatrix} = \frac{1}{2} \cdot \begin{pmatrix} -1 \\ -1 \\ 2 \end{pmatrix}$$

Hence $\text{span}(\underline{v_1}, \underline{v_2}) = E_1 = \text{span}(\underline{v_1}, \underline{v_2}')$ $\quad \underline{v_1} = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \quad \underline{v_2}'' = \begin{pmatrix} -1 \\ -1 \\ 2 \end{pmatrix}$

$\left(v_2' = \frac{1}{2} v_2'' \right)$

$\Rightarrow P = \begin{pmatrix} -1 & -1 & 1 \\ 1 & -1 & 1 \\ 0 & 2 & 1 \end{pmatrix}$ satisfies

$$\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \cdot \begin{pmatrix} -1 \\ -1 \\ 2 \end{pmatrix} = 1 - 1 = 0, \quad \left\| \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \right\| = \sqrt{2}$$

$$\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = -1 + 1 = 0, \quad \left\| \begin{pmatrix} -1 \\ -1 \\ 2 \end{pmatrix} \right\| = \sqrt{6}$$

$$\begin{pmatrix} -1 \\ -1 \\ 2 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = -1 - 1 + 2 = 0, \quad \left\| \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\| = \sqrt{3}$$

Therefore:

$$D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{pmatrix} \quad P = \begin{pmatrix} -1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \\ 1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \\ 0 & 2/\sqrt{6} & 1/\sqrt{3} \end{pmatrix}$$

gives an orthogonal diag. of A

Remark:

If $f(\underline{x}) = \underline{x}^T A \underline{x}$ and A has orth. diag $P^T A P = D$

then $x = P\underline{u}$ gives:

$$f(\underline{x}) = (P\underline{u})^T A (P\underline{u}) = \underline{u}^T P^T A P \underline{u} = \underline{u}^T D \underline{u}$$

$$= \underline{\lambda_1 u_1^2 + \lambda_2 u_2^2 + \dots + \lambda_n u_n^2}$$

in the new variables $\underline{u}$ given by $\underline{x} = P\underline{u}$, or $\boxed{\underline{u} = P^T \cdot \underline{x}}$

Ex:

$$\begin{pmatrix} u \\ v \\ w \end{pmatrix} = \begin{pmatrix} -1/\sqrt{2} & 1/\sqrt{2} & 0 \\ -1/\sqrt{6} & -1/\sqrt{6} & 2/\sqrt{6} \\ 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} :$$

$(\underline{u} = P^T \cdot \underline{x})$

$$\left.\begin{aligned} u &= \tfrac{1}{\sqrt{2}}(-x + y) \\ v &= \tfrac{1}{\sqrt{6}}(-x - y + 2z) \\ w &= \tfrac{1}{\sqrt{3}}(x + y + z) \end{aligned}\right\} \Rightarrow \begin{aligned} f(\underline{x}) &= \underline{x}^T A \underline{x} \\ &= \underline{u}^T D \underline{u} \\ &= 1 \cdot u^2 + 1 \cdot v^2 + 4w^2 \\ &= \underline{\underline{u^2 + v^2 + 4w^2}} \end{aligned}$$