21-241 Matrices and Linear Transformations — **Summer 2020**

Lecture 14 — **Monday, June 8**

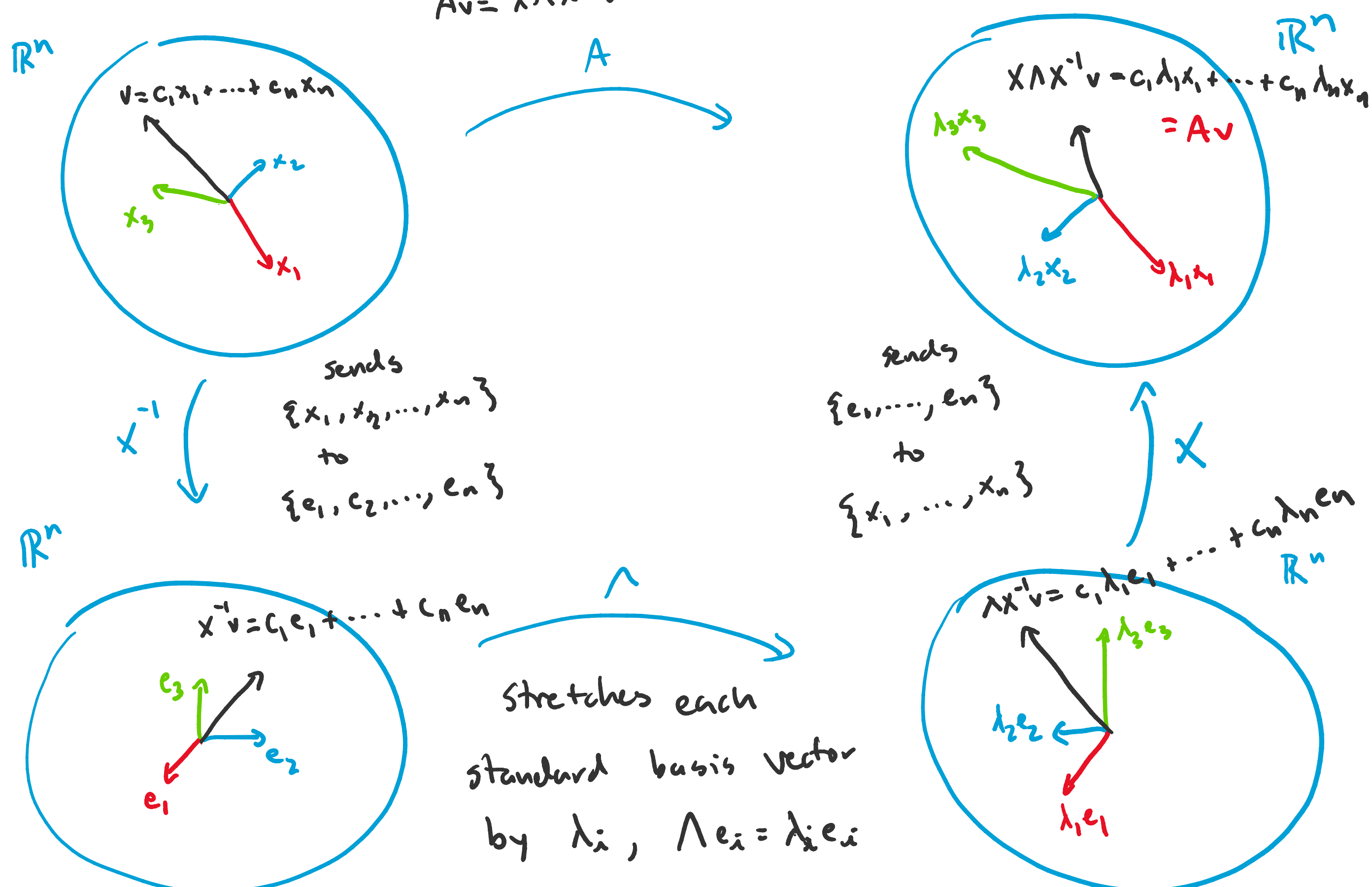
1 Diagonalization as change of basis (Ch. 6.2)

Recall:

Proposition: If an $n \times n$ matrix A has n independent eigenvectors $\{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$, with eigenvalues $\{\lambda_1, \ldots, \lambda_n\}$, then

$A = X \Lambda X^{-1}$, where $X = (x_1 | x_2 | \cdots | x_n)$, $\Lambda = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$

$Av = X\Lambda X^{-1} v$



(Geometric intuition for fact $\det A = \lambda_1\lambda_2\cdots\lambda_n$, HW 7)

Same transformation w.r.t. different bases.

- $A(\overbrace{c_1\mathbf{x}_1 + \cdots + c_n\mathbf{x}_n}^{\text{lin. comb. of eigenbasis}}) = c_1Ax_1 + \cdots + c_nAx_n = c_1\lambda_1x_1 + \cdots + c_n\lambda_nx_n$
- $\Lambda(\underbrace{c_1\mathbf{e}_1 + \cdots + c_n\mathbf{e}_n}_{\text{lin. comb. of std. basis}}) = c_1\Lambda e_1 + \cdots + c_n\Lambda e_n = c_1\lambda_1e_1 + \cdots + c_n\lambda_ne_n$

2 Powers of a matrix

Suppose $A = X\Lambda X^{-1}$. ← diagonalization

Then $A^n = (X\Lambda X^{-1})^n$

hard to compute

$$= (X\Lambda \underbrace{X^{-1})(X}_{I_n}\Lambda \underbrace{X^{-1})(X}_{I_n}\Lambda X^{-1})\cdots(X\Lambda X^{-1})$$

$$= X\Lambda^n X^{-1}$$

easy to compute Λ^n

E.g. (lec 13)

eigenvalues for A^{200} are $(-1)^{200}, 5^{200}$

$$\underbrace{\begin{pmatrix} 1 & 2 \\ 4 & 3 \end{pmatrix}^{200}}_{A^{200}} = \underbrace{\begin{pmatrix} -1 & 1/2 \\ 1 & 1 \end{pmatrix}}_{X} \underbrace{\begin{pmatrix} (-1)^{200} & 0 \\ 0 & 5^{200} \end{pmatrix}}_{\Lambda^{200}} \underbrace{\begin{pmatrix} -2/3 & 1/3 \\ 2/3 & 2/3 \end{pmatrix}}_{X^{-1}}$$

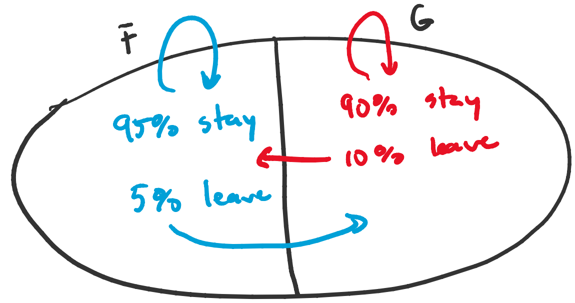
eigenvectors for A^{200} = eigenvectors for A

diagonalization for A^{200}

3 Markov matrices (Ch. 10.3)

Example: Population dynamics

Suppose there are two states F and G.



Let

- $u_F(i) =$ pop. of F @ time i
- $u_G(i) =$ " " G @ " "

Then

- $u_F(1) = .95\, u_F(0) + .1\, u_G(0)$
- $u_G(1) = .05\, u_F(0) + .9\, u_G(0)$

and in general

- $u_F(i+1) = .95\, u_F(i) + .1\, u_G(i)$
- $u_G(i+1) = .05\, u_F(i) + .9\, u_G(i)$

$$\begin{pmatrix} u_F(i+1) \\ u_G(i+1) \end{pmatrix} = \begin{array}{c|c} & \text{today} \\ & \begin{array}{cc} F & G \end{array} \\ \text{tomorrow} \begin{array}{c} F \\ G \end{array} & \underbrace{\begin{pmatrix} .95 & .1 \\ .05 & .9 \end{pmatrix}}_{M} \end{array} \begin{pmatrix} u_F(i) \\ u_G(i) \end{pmatrix}$$

Question: What happens to $\mathbf{u}_i = \begin{pmatrix} u_F(i) \\ u_G(i) \end{pmatrix}$ as i becomes large?

$$\vec{u}_{i+1} = M \vec{u}_i$$

E.g. Suppose $\mathbf{u}_0 = \begin{pmatrix} u_F(0) \\ u_G(0) \end{pmatrix} = \begin{pmatrix} 3 \\ 0 \end{pmatrix}$.

$$\vec{u}_1 = \begin{pmatrix} u_F(1) \\ u_G(1) \end{pmatrix} = M \begin{pmatrix} 3 \\ 0 \end{pmatrix} \simeq \begin{pmatrix} 2.85 \\ .15 \end{pmatrix} \leftarrow \text{sum} = 3$$

$$\vec{u}_2 = \begin{pmatrix} u_F(2) \\ u_G(2) \end{pmatrix} = M \begin{pmatrix} 2.85 \\ .15 \end{pmatrix} = M^2 \begin{pmatrix} 3 \\ 0 \end{pmatrix} = \begin{pmatrix} 2.7225 \\ 0.2775 \end{pmatrix} \leftarrow \text{sum} = 3$$

$$\cdots$$

$$\vec{u}_{20} = M^{20} \begin{pmatrix} 3 \\ 0 \end{pmatrix} = \begin{pmatrix} 2.038 \\ .96 \end{pmatrix}$$

$$\cdots$$

$$\vec{u}_{50} = M^{50} \begin{pmatrix} 3 \\ 0 \end{pmatrix} \approx \begin{pmatrix} 2 \\ 1 \end{pmatrix} \leftarrow \text{equilibrium}$$

E.g. Suppose $\mathbf{u}_0 = \begin{pmatrix} u_F(0) \\ u_G(0) \end{pmatrix} = \begin{pmatrix} 0 \\ 3 \end{pmatrix}$.

$$M^k \begin{pmatrix} 0 \\ 3 \end{pmatrix} \longrightarrow \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

E.g. Suppose $\mathbf{u}_0 = \begin{pmatrix} u_F(0) \\ u_G(0) \end{pmatrix} = \begin{pmatrix} 1.5 \\ 1.5 \end{pmatrix}$.

$$M^k \begin{pmatrix} 1.5 \\ 1.5 \end{pmatrix} \longrightarrow \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

What's special about $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$?

It's a "steady state" (fixed by M) → eigenvector of M w/ eigenvalue $\lambda = 1$

$$M \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} .95 & .1 \\ .05 & .9 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

Note: $M^k \longrightarrow \begin{pmatrix} 2/3 & 2/3 \\ 1/3 & 1/3 \end{pmatrix}$ (column space spanned by $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$)

each column is e-vector w/ $\lambda = 1$

Eigenvectors for M:

basis for $\mathbb{R}^2$ $\left\{ \begin{array}{ll} x_1 = \binom{2}{1}, & \lambda = 1 \leftarrow \text{bigger eigenvalue} \\ x_2 = \binom{-1}{1}, & \lambda = .85 \end{array} \right.$

Any $\mathbf{v} \in \mathbb{R}^2$ can be written in terms of the basis of eigenvectors. $v = c_1 x_1 + c_2 x_2$

$c_1 \binom{2}{1} + c_2 \binom{-1}{1}$

Then $M\vec{v} = M\left(c_1 \binom{2}{1} + c_2 \binom{-1}{1}\right)$

$= c_1 M \binom{2}{1} + c_2 M \binom{-1}{1}$

$= c_1 \, 1 \binom{2}{1} + c_2 (.85) \binom{-1}{1}$

$M^2 \vec{v} = c_1 1^2 \binom{2}{1} + c_2 (.85)^2 \binom{-1}{1}$

...

$M^k \vec{v} = c_1 \underbrace{1^k}_{\to 1} \binom{2}{1} + c_2 \underbrace{(.85)^k}_{\to 0} \binom{-1}{1}$

$\to c_1 \binom{2}{1}$

E.g.

- $\binom{3}{0} = \underset{c_1}{1}\binom{2}{1} - 1\binom{-1}{1}$, so $M^k \binom{3}{0} \longrightarrow 1 \binom{2}{1}$
- $\binom{1.5}{1.5} = \underset{c_1}{1}\binom{2}{1} + \frac{1}{2}\binom{-1}{1}$, so $M^k \binom{1.5}{1.5} \to 1 \binom{2}{1}$

An $n \times n$ matrix M is a **Markov matrix** if

"for all"

- $M_{ij} \geq 0 \quad \forall i,j$
- $\sum_{i=1}^{n} M_{ij} = 1 \quad \forall j$ (columns sum to 1)

"positive" if $M_{ij} > 0$ for all i,j

Notation: Let $\mathbf{1} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$

Note that M is a Markov matrix if and only if $M^T \mathbb{1} = \mathbb{1}$ & $M_{ij} \geq 0 \ \forall ij$

E.g. $M^T \mathbb{1} = \underbrace{\begin{pmatrix} .95 & .05 \\ .1 & .9 \end{pmatrix}}_{M^T} \underbrace{\begin{pmatrix} 1 \\ 1 \end{pmatrix}}_{\mathbb{1}} = \underbrace{\begin{pmatrix} 1 \\ 1 \end{pmatrix}}_{\mathbb{1}}$

A **Markov chain** has states and transition probabilities between states.

↳ only depend on current state

↳ ≥ 0, sum to 1

(prior) states

(subsequent) states

$$\begin{array}{cc} & i \\ j & \begin{pmatrix} & \\ & M_{ij} \end{pmatrix} \end{array}$$

Markov Matrix M

M_{ij} ↳ Pr(transition from state i to state j)

4 Properties of Markov matrices

Proposition: If M is a diagonalizable Markov matrix with

- eigenvector $\mathbf{x}_1$ and eigenvalue $\lambda_1 = 1$, and
- all other eigenvalues have $|\lambda| < 1$, then

then $\vec{x}_1$ is an attracting steady state (i.e. $M^k v \to x_1$)

Proof: Let $\{x_1, \ldots, x_n\}$ be a basis of e-vectors for M,

w/ eigenvalues $\{\underset{=1}{\lambda_1}, \underbrace{\lambda_2, \ldots, \lambda_n}_{|\lambda_i| < 1}\}$

Let $v \in \mathbb{R}^n$. Then $\vec{v} = c_1 \vec{x}_1 + \cdots + c_n \vec{x}_n$,

& by linearity,

$$M^k \vec{v} = c_1 \underbrace{\lambda_1^k}_{\to 1} x_1 + c_2 \underbrace{\lambda_2^k}_{\to 0} x_2 + \cdots + c_n \underbrace{\lambda_n^k}_{\to 0} x_n$$

$$\approx c_1 x_1 \quad \text{for } k \text{ large.} \qquad \square$$

(Largest eigenvalues control long-term behavior of A^k)

Key properties of positive Markov matrices (no 0 entries)

- Largest eigenvalue $\lambda_1 = 1$
- Other eigenvalues $|\lambda_i| < 1$
- Attracting steady state $\vec{x}_1$
- $M^k \to$ $\begin{pmatrix} x_1 & \cdots & x_1 \end{pmatrix}$ for large k.

Proposition: For any $n \times n$ matrix A, A and A^T have the same eigenvalues.

Proof: $\underbrace{\det(A - \lambda I)}_{\text{char. poly for } A} = \det(A - \lambda I)^T$ $\det A = \det A^T$ (Prop a)

$= \det(A^T - (\lambda I)^T)$ $(A-B)^T = A^T - B^T$

$= \underbrace{\det(A^T - \lambda I)}_{\text{char. poly for } A^T}$ λI diagonal

A & A^T have same char. poly $\Rightarrow$ same eigenvalues. $\square$

Proposition: Every Markov matrix has eigenvalue $\lambda = 1$.

Proof: M Markov $\Rightarrow M^T \mathbb{1} = \mathbb{1}$

$\Rightarrow \mathbb{1}$ is an eigenvector for M^T w/ eigenvalue $\lambda = 1$.

$\Rightarrow \lambda = 1$ is an eigenvalue for M. $\square$

Proposition: If M is a Markov matrix and $\mathbf{x}$ is a vector, then the sum of the components of $M\mathbf{x}$ is the same as the sum of the components of $\mathbf{x}$.

E.g. $M\begin{pmatrix}3\\0\end{pmatrix} = \begin{pmatrix}2.85\\.15\end{pmatrix}$, $3+0=3$, $2.85+.15=3$

Proof: component sum of $M\vec{x}$ = component sum of $x_1 M_{*1} + \cdots + x_n M_{*n}$

$= x_1 \underbrace{(M_{11} + \cdots + M_{n1})}_{=1} + \cdots + x_n \underbrace{(M_{1n} + \cdots + M_{nn})}_{=1}$

$= x_1 + \cdots + x_n$

$=$ component sum of $\vec{x}$ $\square$

Proposition: If A and B are $n \times n$ Markov matrices, then AB is Markov.

Proof: • $(AB)_{ij} = A_{i*} B_{*j} = \underset{\geq 0}{a_{i1}} \underset{\geq 0}{b_{1j}} + \cdots + \underset{\geq 0}{a_{in}} \underset{\geq 0}{b_{nj}} \geq 0$.

• $AB = \left(A(B_{*1}) \middle| \cdots \middle| A(B_{*n}) \right)$

Note: component sum of $B_{*j} = 1$

$\Rightarrow$ component sum of $A(B_{*j}) = 1$

So cols of AB sum to 1. $\square$

Proposition: If λ is an eigenvalue for an $n \times n$ Markov matrix M, then $|\lambda| \leq 1$.

Proof: Let $\vec{x} = (x_1, \ldots, x_n)$ be a unit eigenvector for M, w/ eigenvalue λ.

i.e. $x_1^2 + \cdots + x_n^2 = 1 \Rightarrow |x_i| \leq 1$

M^k Markov $\Rightarrow 0 \leq (M^k)_{ij} \leq 1$

So $|i^{th}$ entry of vector $M^k \vec{x}| = \underbrace{(\text{row } i \text{ of } M^k) \cdot \vec{x}}_{\text{dot product of \#'s} \leq 1}$

$\boxed{\leq n}$

But $M^k \vec{x} = \lambda^k \vec{x}$

If $|\lambda| > 1$, then $\lambda^k \vec{x}$ has entries $\boxed{> n}$ for k sufficiently large.

Contradiction $\Rightarrow |\lambda| \leq 1$. $\square$

5 Applications of matrix powers

Markov matrices

- Population dynamics
- Epidemiology S-I-R models
- Random walks
- Page rank (Google)
- Games of skill and chance

Non-Markov matrices

- Recurrence relations
- Economics (Ch. 10)