

Linear Algebra Notes — Lecture 10/1

Markov Chains

Lecture pages 146–158 (draft) · Lay §5.9 (6th ed.), §4.9 (older editions)

Contents

Overview	1
1 Toy example: two libraries (pp. 146–147)	2
2 Switching to proportions (p. 148)	2
3 Probability vectors, stochastic matrices, Markov chains (p. 149)	3
4 Stochastic matrices keep probability vectors (pp. 150–152)	3
5 Steady states (p. 153)	5
6 Examples: describe all steady states (p. 154)	5
7 Every stochastic matrix has a steady state (p. 155)	6
8 Regular stochastic matrices (p. 156)	7
9 Sinks (pp. 157–158)	8
Corrections to the handwritten notes	9

Overview

A Markov chain is the recursion $X_{k+1} = PX_k$ from the 9/29 lecture, applied to probabilities. The vector X_k lists how things are spread out at time k , and a fixed matrix P says how they move in one step. The big questions are:

- Is there a *steady state*, a distribution that doesn't change?
- Does every starting distribution eventually settle into it?

Every blank and question from the draft notes is answered below (blue boxes), and typos are listed at the end.

1 Toy example: two libraries (pp. 146–147)

The story. A small town has two libraries, A and B . Each month:

- of the books checked out from A , 80% are returned to A and 20% to B ;
- of the books checked out from B , 30% are returned to A and 70% to B .

Right now A has 500 books and B has 500. How many books will each have after k months? (The model assumes every book is checked out and returned once a month.)

Model. Let a_k and b_k be the numbers of books in A and B after k months. Library A next month gets 80% of A 's books plus 30% of B 's:

$$a_{k+1} = 0.8 a_k + 0.3 b_k, \quad b_{k+1} = 0.2 a_k + 0.7 b_k.$$

Fill-in: the matrix P (p. 147)

With $X_k = \begin{bmatrix} a_k \\ b_k \end{bmatrix}$, the two equations are exactly $X_{k+1} = P X_k$ for

$$P = \begin{bmatrix} 0.8 & 0.3 \\ 0.2 & 0.7 \end{bmatrix}.$$

How to read P : column 1 says where A 's books go (0.8 to A , 0.2 to B); column 2 says where B 's books go (0.3 to A , 0.7 to B). Row i , column j = fraction moving from j to i .

Fill-in: X_1 and X_2 , starting from $X_0 = (500, 500)$

$$X_1 = P X_0 = \begin{bmatrix} 0.8(500) + 0.3(500) \\ 0.2(500) + 0.7(500) \end{bmatrix} = \begin{bmatrix} 400 + 150 \\ 100 + 350 \end{bmatrix} = \begin{bmatrix} 550 \\ 450 \end{bmatrix},$$

$$X_2 = P X_1 = \begin{bmatrix} 0.8(550) + 0.3(450) \\ 0.2(550) + 0.7(450) \end{bmatrix} = \begin{bmatrix} 440 + 135 \\ 110 + 315 \end{bmatrix} = \begin{bmatrix} 575 \\ 425 \end{bmatrix}.$$

The total stays 1000 every month. Books move between libraries, but none are created or lost.

2 Switching to proportions (p. 148)

It is better to replace $\begin{bmatrix} a_k \\ b_k \end{bmatrix}$ by $\frac{1}{1000} \begin{bmatrix} a_k \\ b_k \end{bmatrix}$. Now

$$\frac{a_k}{1000} = \text{proportion of books in } A \text{ after } k \text{ months}, \quad \frac{b_k}{1000} = \text{proportion in } B.$$

Why this is better.

- The answer no longer depends on how many books there are: 1000 books or 10 million, the proportions behave the same way.

- The entries are ≥ 0 and add to 1, so they can be read as *probabilities*: the chance that a randomly chosen book is in A or in B .
- The recursion is unchanged. P is linear, so $P(\frac{1}{1000}X_k) = \frac{1}{1000}PX_k$.

In proportions, $X_0 = (0.5, 0.5)$, $X_1 = (0.55, 0.45)$, $X_2 = (0.575, 0.425)$.

3 Probability vectors, stochastic matrices, Markov chains (p. 149)

Definition (probability vector)

A **probability vector** is $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ with every $x_i \geq 0$ and $x_1 + \dots + x_n = 1$. Its coordinates are the probabilities of the n possible outcomes.

Fill-in: coin toss

Two outcomes, heads and tails. For a fair coin the probability vector is $\begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix}$. (A biased coin landing heads 70% of the time would be $(0.7, 0.3)$.)

Definition (stochastic matrix)

A **stochastic matrix** is an $n \times n$ matrix whose *columns* are probability vectors: all entries are ≥ 0 and each column sums to 1.

Meaning: column j is the probability distribution of where you go next if you are in state j now. The library matrix is stochastic: columns $(0.8, 0.2)$ and $(0.3, 0.7)$.

Definition (Markov chain)

A (discrete) **Markov chain** is a sequence of probability vectors $X_0, X_1, X_2, \dots$ where each vector is determined by the previous one. In this course: $X_{k+1} = PX_k$ for a fixed stochastic matrix P .

4 Stochastic matrices keep probability vectors (pp. 150–152)

Theorem (main example, p. 150)

If X_0 is a probability vector in $\mathbb{R}^n$ and P is an $n \times n$ stochastic matrix, then for every k the vector $X_k = P^k X_0$ is a probability vector. So (X_k) is a Markov chain.

Proof by induction on k .

- **Base case** $k = 0$: X_0 is a probability vector by assumption.
- **Induction step:** if X_k is a probability vector, then $X_{k+1} = PX_k$ is one too. This follows from the lemma below with $x = X_k$.

Lemma (p. 151)

If x is a probability vector and P is stochastic, then Px is a probability vector.

Write $P = [a_{ij}]$ and $x = (x_1, \dots, x_n)$, so that

$$Px = \begin{bmatrix} a_{11}x_1 + \dots + a_{1n}x_n \\ \vdots \\ a_{n1}x_1 + \dots + a_{nn}x_n \end{bmatrix}.$$

Why are the entries of Px all ≥ 0 ?

Each entry is a sum of products $a_{ij}x_j$. Every $a_{ij} \geq 0$ (entries of a stochastic matrix) and every $x_j \geq 0$ (probability vector). Products and sums of nonnegative numbers are nonnegative.

Why do they add up to 1?

Add all the entries of Px and group by x_j instead of by row:

$$(a_{11} + a_{21} + \dots + a_{n1})x_1 + \dots + (a_{1n} + a_{2n} + \dots + a_{nn})x_n.$$

The bracket next to x_j is the sum of column j of P , which is 1 because P is stochastic. So the total is

$$1 \cdot x_1 + \dots + 1 \cdot x_n = x_1 + \dots + x_n = 1,$$

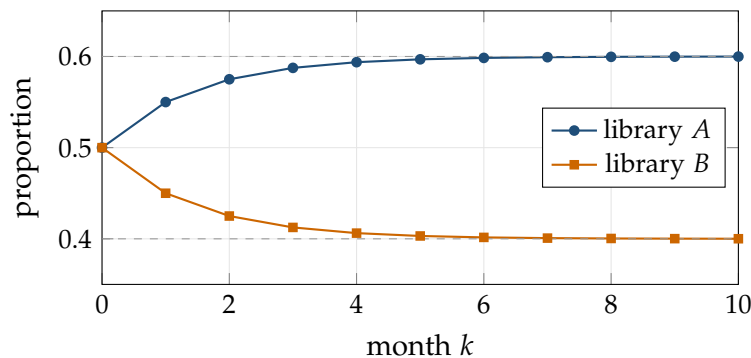
since x is a probability vector.

Conclusion (p. 152). The recursion $X_{k+1} = PX_k$ encodes how outcome probabilities change from time k to time $k + 1$.

Back to the library. $P = \begin{bmatrix} 0.8 & 0.3 \\ 0.2 & 0.7 \end{bmatrix}$ is stochastic, $X_0 = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}$ is a probability vector, and

$$X_1 = \begin{bmatrix} 0.4 + 0.15 \\ 0.1 + 0.35 \end{bmatrix} = \begin{bmatrix} 0.55 \\ 0.45 \end{bmatrix}.$$

Continuing, the proportion in A creeps up toward 0.6:



The next section explains the limit $(0.6, 0.4)$.

5 Steady states (p. 153)

Definition

A **steady state** for $X_{k+1} = PX_k$ is a probability vector q with $Pq = q$.

Meaning: if the chain is ever at q , it stays there forever. In 9/29 language, q is an eigenvector of P with eigenvalue 1 that happens to be a probability vector.

How to find one. Use $Pq = q \iff Pq - q = 0 \iff (P - I)q = 0$. So find $\text{Nul}(P - I)$ and pick the vector in it whose entries are ≥ 0 and add to 1.

Fill-in: steady state for the library

$$P - I = \begin{bmatrix} 0.8 - 1 & 0.3 \\ 0.2 & 0.7 - 1 \end{bmatrix} = \begin{bmatrix} -0.2 & 0.3 \\ 0.2 & -0.3 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 0 \end{bmatrix} = (P - I) \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -0.2x + 0.3y \\ 0.2x - 0.3y \end{bmatrix}.$$

Both rows say the same thing: $0.2x = 0.3y$, i.e. $x = 1.5y$. A probability vector also needs $x + y = 1$, so $1.5y + y = 1$, giving $y = 0.4$ and $x = 0.6$:

$$q = \begin{bmatrix} 0.6 \\ 0.4 \end{bmatrix}.$$

Check: $Pq = (0.48 + 0.12, 0.12 + 0.28) = (0.6, 0.4) = q$. ✓

In the long run 60% of the books, i.e. 600, are in A and 400 are in B , matching the graph.

Why the library converges (using 9/29). The vector $(-1, 1)$ satisfies $P(-1, 1)^T = (-0.5, 0.5)^T = \frac{1}{2}(-1, 1)^T$, so it is an eigenvector with eigenvalue $\frac{1}{2}$. Any probability vector is $(x, 1 - x) = q + (0.6 - x)(-1, 1)$. So by the 9/29 theorem,

$$X_k = q + (0.6 - x_0) \left(\frac{1}{2}\right)^k \begin{bmatrix} -1 \\ 1 \end{bmatrix} \longrightarrow q.$$

Check with $x_0 = 0.5$: $X_2 = (0.6, 0.4) + 0.1 \cdot \frac{1}{4}(-1, 1) = (0.575, 0.425)$. ✓ The gap to the steady state halves every month.

6 Examples: describe all steady states (p. 154)

In each case write $q = (x, y)$ with $x, y \geq 0$ and $x + y = 1$, and solve $Pq = q$.

$$P = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I: \text{every probability vector is a steady state}$$

$Iq = q$ for every q , so the steady states are all $(x, 1 - x)$ with $0 \leq x \leq 1$. There are infinitely many. In words: nothing ever moves, so whatever you start with stays.

$$P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}: \text{exactly one steady state, } (\frac{1}{2}, \frac{1}{2})$$

$Pq = (y, x)$, so $Pq = q$ means $y = x$. With $x + y = 1$: $q = (\frac{1}{2}, \frac{1}{2})$. In words: every step *swaps* the two states. The chain started at $(1, 0)$ goes $(1, 0), (0, 1), (1, 0), \dots$ forever; it only sits still if it starts half and half.

$$P = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}: \text{exactly one steady state, } (1, 0)$$

$Pq = (x + y, 0)$, so $Pq = q$ means $y = 0$ and $x = x + y$. With $x + y = 1$: $q = (1, 0)$. In words: everything in state 2 moves to state 1 and state 1 keeps everything, so after one step everything is in state 1: $P(x, y)^T = (x + y, 0)^T = (1, 0)^T$ for every probability vector.

7 Every stochastic matrix has a steady state (p. 155)

Theorem

Every stochastic matrix has a steady state.

Proof (only for 2×2 matrices). A 2×2 stochastic matrix has columns $(r, 1 - r)$ and $(1 - s, s)$:

$$P = \begin{bmatrix} r & 1 - s \\ 1 - r & s \end{bmatrix}, \quad 0 \leq r, s \leq 1, \quad P - I = \begin{bmatrix} r - 1 & 1 - s \\ 1 - r & s - 1 \end{bmatrix}.$$

One equation, not two. The second row of $P - I$ is -1 times the first, because each column of $P - I$ sums to $1 - 1 = 0$. So $(P - I) \begin{bmatrix} x \\ y \end{bmatrix} = 0$ is the single equation $(r - 1)x + (1 - s)y = 0$. Adding the probability condition $x + y = 1$ gives the system

$$\begin{bmatrix} 1 & 1 \\ r - 1 & 1 - s \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

Is it consistent? Yes

Row reduce the augmented matrix, adding $(1 - r)$ times row 1 to row 2:

$$\left[\begin{array}{cc|c} 1 & 1 & 1 \\ r - 1 & 1 - s & 0 \end{array} \right] \longrightarrow \left[\begin{array}{cc|c} 1 & 1 & 1 \\ 0 & 2 - r - s & 1 - r \end{array} \right].$$

Case $2 - r - s \neq 0$. Then there is a unique solution:

$$y = \frac{1-r}{2-r-s}, \quad x = 1 - y = \frac{1-s}{2-r-s}.$$

Case $2 - r - s = 0$. Since $r, s \leq 1$, $r + s = 2$ forces $s = 1 = r$. Then $1 - r = 0$ too, so row 2 reads $0 = 0$: consistent! Here $P = I$ and every $(x, 1 - x)$ works (p. 154).

The step the notes skip: the solution is a probability vector

It isn't enough to solve the equations; a steady state must also have $x, y \geq 0$. In the first case, the numerators $1 - r$ and $1 - s$ are ≥ 0 because $r, s \leq 1$. The denominator $2 - r - s = (1 - r) + (1 - s)$ is ≥ 0 , and nonzero in this case, so it is > 0 . So $x, y \geq 0$, and $x + y = 1$ by construction.

Check with the library ($r = 0.8, s = 0.7$): $y = \frac{0.2}{0.5} = 0.4, x = \frac{0.3}{0.5} = 0.6$. ✓

8 Regular stochastic matrices (p. 156)

Definition

A stochastic matrix P is **regular** if, for some $k \geq 1$, the power P^k has *all* entries strictly positive.

Meaning: after k steps you can get from any state to any state with positive probability.

Is $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ regular? No

$P^2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = I$, so the powers alternate: $P, I, P, I, \dots$. Every one of them has zeros, so no power is all positive. (After an even number of steps you can't be in the other state; after an odd number you can't be in the same one.)

Is $\begin{bmatrix} 1 & 0.5 \\ 0 & 0.5 \end{bmatrix}$ regular? No

State 1 never sends anything to state 2 (column 1 is $(1, 0)$), so once you're in state 1 you stay. Concretely,

$$P^2 = \begin{bmatrix} 1 & 0.75 \\ 0 & 0.25 \end{bmatrix}, \quad P^3 = \begin{bmatrix} 1 & 0.875 \\ 0 & 0.125 \end{bmatrix}, \quad \text{and in general} \quad P^k = \begin{bmatrix} 1 & 1 - 0.5^k \\ 0 & 0.5^k \end{bmatrix}.$$

(Induction: $P^{k+1} = P \cdot P^k$ has bottom row $(0 \cdot 1 + 0.5 \cdot 0, 0 \cdot (1 - 0.5^k) + 0.5 \cdot 0.5^k) = (0, 0.5^{k+1})$.) The bottom-left entry is always 0.

Is $P = \begin{bmatrix} 0.5 & 0.3 & 0.4 \\ 0.5 & 0.3 & 0.1 \\ 0 & 0.4 & 0.5 \end{bmatrix}$ regular? Yes

It is stochastic: the columns sum to $0.5 + 0.5 + 0 = 1$, $0.3 + 0.3 + 0.4 = 1$, $0.4 + 0.1 + 0.5 = 1$. P itself has a 0, so check P^2 (row of P times column of P):

$$P^2 = \begin{bmatrix} 0.40 & 0.40 & 0.43 \\ 0.40 & 0.28 & 0.28 \\ 0.20 & 0.32 & 0.29 \end{bmatrix}.$$

For example, the bottom-left entry is $0(0.5) + 0.4(0.5) + 0.5(0) = 0.2$. Every entry of P^2 is positive, so P is regular with $k = 2$.

9 Sinks (pp. 157–158)

Definition

A steady state q of a stochastic matrix P is a **sink** if for *every* probability vector X_0 , the sequence $X_{k+1} = PX_k$ converges to q . (Lay doesn't use this term.)

If P has a sink q , then q is the only steady state

Let q' be any steady state. Start the chain at $X_0 = q'$. Since $Pq' = q'$, the chain never moves: $X_k = q'$ for all k , so it converges to q' . But q is a sink, so the same sequence converges to q . A sequence has only one limit, so $q' = q$.

Does $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ have a sink? No

Its only steady state is $(\frac{1}{2}, \frac{1}{2})$ (p. 154), but starting from $X_0 = (1, 0)$ the chain goes $(1, 0), (0, 1), (1, 0), (0, 1), \dots$ and never converges. So no steady state attracts every starting vector.

Theorem (p. 158)

Every regular stochastic matrix P has a sink q . That is, for every probability vector X_0 ,

$$\lim_{k \rightarrow \infty} P^k X_0 = q.$$

Proof later (for $n = 2$).

Examples.

- The library matrix has all positive entries, so it is regular with $k = 1$. Its sink is $q = (0.6, 0.4)$, and the box in Section 5 shows the convergence directly.
- The converse is false: *regular* is enough for a sink, but not required. $\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$ is not regular (its

bottom row is always 0), yet it sends every probability vector to $(1, 0)$ in one step, so $(1, 0)$ is a sink. Likewise $\begin{bmatrix} 1 & 0.5 \\ 0 & 0.5 \end{bmatrix}$ is not regular, but $P^k(x, y)^T = (x + (1 - 0.5^k)y, 0.5^k y)^T \rightarrow (1, 0)$.

- I has no sink: it has infinitely many steady states, and by the box above a sink would have to be the only one.

P		steady states	regular?	sink?
$\begin{bmatrix} 0.8 & 0.3 \\ 0.2 & 0.7 \end{bmatrix}$	(library)	$(0.6, 0.4)$ only	yes	yes
$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$		all $(x, 1 - x)$	no	no
$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$		$(\frac{1}{2}, \frac{1}{2})$ only	no	no (oscillates)
$\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$		$(1, 0)$ only	no	yes

Corrections to the handwritten notes

Page	As written	Should be	Why
151	last term $(a_{n1} + \cdots + a_{nn})x_n$	$(a_{1n} + \cdots + a_{nn})x_n$	x_n multiplies column n , not row n
156	regular if P^k has positive entries	...if <i>some</i> P^k has <i>all</i> entries > 0	one positive entry isn't enough; k is not fixed
157	$x_{k+1} = Px_k$	$X_{k+1} = PX_k$	vectors are capital X