Lecture 18. Markov Chains

YULIA R. GEL

CS/SE/STAT 3341 Probability and Statistics in Computer Science and Software Engineering

March 30, 2014

Matrix Algebra

Since working with Markov chains involves matrices, let us have a brief review of matrix algebra.

$$A = \{A_{ij}, i = 1, \ldots, m; n = 1, \ldots, n\} = \begin{pmatrix} A_{11} & A_{12} & \ldots & A_{1n} \\ A_{21} & A_{22} & \ldots & A_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ A_{m1} & A_{m2} & \ldots & A_{mn} \end{pmatrix}$$

is a $m \times n$-matrix.

Matrices – Special Cases

- A matrix

$$C = \{A_i, i = 1, \ldots, m\} = \begin{pmatrix} C_1 \\ C_2 \\ \ldots \\ C_m \end{pmatrix}$$

is a $m \times 1$-matrix, or **column vector**.

- Similarly, $D = (D_1, D_2, \ldots, D_n)$ is a $1 \times n$-matrix, or **row vector**.

Matrix Addition

Let A and B be two $m \times n$-matrices. Then

$$
\begin{aligned}
A + B &= \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1n} \\ A_{21} & A_{22} & \dots & A_{2n} \\ \dots & \dots & \dots & \dots \\ A_{m1} & A_{m2} & \dots & A_{mn} \end{pmatrix} + \begin{pmatrix} B_{11} & B_{12} & \dots & B_{1n} \\ B_{21} & B_{22} & \dots & B_{2n} \\ \dots & \dots & \dots & \dots \\ B_{m1} & B_{m2} & \dots & B_{mn} \end{pmatrix} \\
&= \begin{pmatrix} A_{11} + B_{11} & A_{12} + B_{12} & \dots & A_{1n} + B_{1n} \\ A_{21} + B_{21} & A_{22} + B_{22} & \dots & V_{2n} + B_{2n} \\ \dots & \dots & \dots & \dots \\ A_{m1} + B_{m1} & A_{m2} + B_{m2} & \dots & A_{mn} + B_{mn} \end{pmatrix}
\end{aligned}
$$

Notice that you can add matrices **only** of the same dimensions!

Matrix Multiplication

Let A be $m \times n$-matrix and C be $n \times p$-matrix. Then

$$
\begin{aligned}
A \times C &= \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1n} \\ A_{21} & A_{22} & \dots & A_{2n} \\ \dots & \dots & \dots & \dots \\ A_{m1} & A_{m2} & \dots & A_{mn} \end{pmatrix} \times \begin{pmatrix} B_{11} & B_{12} & \dots & B_{1p} \\ B_{21} & B_{22} & \dots & B_{2p} \\ \dots & \dots & \dots & \dots \\ B_{n1} & B_{n2} & \dots & B_{np} \end{pmatrix} \\
&= \left\{ (AB)_{ij} \right\}_{i=1,j=1}^{m,p} = \left\{ \sum_{k=1}^{n} A_{ik} B_{kj} \right\}_{i=1,j=1}^{m,p},
\end{aligned}
$$

i.e. we always multiply **row by column**.

Matrix Multiplication – Example

Let $A = \begin{pmatrix} 1 & 0 & -2 \\ 0 & 3 & -1 \end{pmatrix}$ and $B = \begin{pmatrix} 0 & 3 \\ -2 & -1 \\ 0 & 4 \end{pmatrix}$.

Then $AB = \begin{pmatrix} 1 & 0 & -2 \\ 0 & 3 & -1 \end{pmatrix} \begin{pmatrix} 0 & 3 \\ -2 & -1 \\ 0 & 4 \end{pmatrix} = \begin{pmatrix} 0 & -5 \\ -6 & -7 \end{pmatrix}$.

Matrix Multiplication – Example

Let $A = \begin{pmatrix} 1 & 0 & -2 \\ 0 & 3 & -1 \end{pmatrix}$ and $B = \begin{pmatrix} 0 & 3 \\ -2 & -1 \\ 0 & 4 \end{pmatrix}$.

Then $AB = \begin{pmatrix} 1 & 0 & -2 \\ 0 & 3 & -1 \end{pmatrix} \begin{pmatrix} 0 & 3 \\ -2 & -1 \\ 0 & 4 \end{pmatrix} = \begin{pmatrix} 0 & -5 \\ -6 & -7 \end{pmatrix}$.

What is about BA? Is it equal to AB?

Matrix Multiplication – Example

Let $A = \begin{pmatrix} 1 & 0 & -2 \\ 0 & 3 & -1 \end{pmatrix}$ and $B = \begin{pmatrix} 0 & 3 \\ -2 & -1 \\ 0 & 4 \end{pmatrix}$.

Then $AB = \begin{pmatrix} 1 & 0 & -2 \\ 0 & 3 & -1 \end{pmatrix} \begin{pmatrix} 0 & 3 \\ -2 & -1 \\ 0 & 4 \end{pmatrix} = \begin{pmatrix} 0 & -5 \\ -6 & -7 \end{pmatrix}$.

What is about BA? Is it equal to AB?

$$BA = \begin{pmatrix} 0 & 3 \\ -2 & -1 \\ 0 & 4 \end{pmatrix} \begin{pmatrix} 1 & 0 & -2 \\ 0 & 3 & -1 \end{pmatrix} = \begin{pmatrix} 0 & 9 & -3 \\ -2 & -3 & 5 \\ 0 & 12 & -4 \end{pmatrix}$$

Note that in general $AB \neq BA$!

Matrix Multiplication – Example

Now let $A = (1, 2, 3)$ and $B = \begin{pmatrix} 1 \\ -2 \\ 4 \end{pmatrix}$, then

$AB = (1, 2, 3) \begin{pmatrix} 1 \\ -2 \\ 4 \end{pmatrix} = 1 \times 1 + 2 \times (-2) + 3 \times 4 = 9$, i.e. multiplication of row by column gives a scalar.

But $BA = \begin{pmatrix} 1 \\ -2 \\ 4 \end{pmatrix} (1, 2, 3) = \begin{pmatrix} 1 & 2 & 3 \\ -2 & -4 & 6 \\ 4 & 8 & 12 \end{pmatrix}$,
i.e. multiplication of column by row gives a matrix.

Indeed, as we seen before, in general $AB \neq BA$!

Car Rental Example

Example. Suppose a car rental agency has three locations in New York City: Downtown location (labeled A), East end location (labeled B) and a West end location (labeled C). The agency has a group of delivery drivers to serve all three locations. The agency's statistician has determined the following:

1. Of the calls to the Downtown location, 30% are delivered in Downtown area, 30% are delivered in the East end, and 40% are delivered in the West end
2. Of the calls to the East end location, 40% are delivered in Downtown area, 40% are delivered in the East end, and 20% are delivered in the West end
3. Of the calls to the West end location, 50% are delivered in Downtown area, 30% are delivered in the East end, and 20% are delivered in the West end.

Car Rental Example

We model this problem with the following transition probability matrix:

$$P = \left(\begin{array}{c|ccc} & A & B & C \\ \hline A & 0.3 & 0.3 & 0.4 \\ B & 0.4 & 0.4 & 0.2 \\ C & 0.5 & 0.3 & 0.2 \end{array} \right).$$

If you begin at location C, what is the probability that you will be in area B after 2 deliveries?

Car Rental Example

Think about how you can get to B in two steps. We can go from C to C, then from C to B, we can go from C to B, then from B to B, or we can go from C to A, then from A to B.

To solve this problem, let p_{XY} denote the probability of going from X to Y in one delivery (where X,Y can be A,B or C).

From the Law of Total Probability, the probability that a delivery person goes from C to B in 2 deliveries:

$$\begin{aligned} p^2_{CB} &= p_{(CA)}p_{(AB)} + p_{(CB)}p_{(BB)} + p_{(CC)}p_{(CB)} \\ &= (0.5)(0.3) + (0.3)(0.4) + (0.2)(0.3) = 0.33. \end{aligned}$$

This tells us that if we begin at location C, we have a 33% chance of being in location B after 2 deliveries.

Car Rental Example

Let us try this for another pair. If we begin at location B, what is the probability of being at location B after 2 deliveries?

$$\begin{aligned} p^2_{BB} &= p_{(BA)}p_{(AB)} + p_{(BB)}p_{(BB)} + p_{(BC)}p_{(CB)} \\ &= (0.4)(0.3) + (0.4)(0.4) + (0.2)(0.3) = 0.34. \end{aligned}$$

Now it was not so bad calculating where you would be after 2 deliveries, but what if you need to know where you will be after 5, or 15 deliveries? That could take a LONG time!

There must be some simplification...

Markov Chain – Extension to h-steps

Indeed, calculating h-steps transition probabilities is simply based on matrix multiplication of our original transition matrix h times. I.e.,

$$P^{(h)} = P^h.$$

In fact, notice that in our car rental example:

$$P^2 = \left(\begin{array}{c|ccc} & A & B & C \\ \hline A & 0.41 & 0.33 & 0.26 \\ B & 0.38 & 0.34 & 0.28 \\ C & 0.37 & 0.33 & 0.3 \end{array} \right),$$

where $p^2_{CB} = 0.33$ and $p^2_{BB} = 0.34$, i.e. exactly as we found before!

Markov Chain – Distribution of $X(h)$

Let $P_0 = (P_0(1), \dots, P_0(n))$ be probability mass function (pmf) of the initial state $X(0)$.

Similarly, $P_h = (P_h(1), \dots, P_h(n))$ be probability mass function (pmf) of $X(h)$, i.e. state after h transitions.

Then using our matrix approach, we can easily show that distribution of $X(h)$ can be computed as

$$P_h = P_0 P^h.$$

Markov Chain – Limiting Case of $h \to \infty$

What happens if we have very many transitions, i.e. $h \to infty$?

Definition. A collection of limiting probabilities

$$\pi_x = \lim_{h \to \infty} P_h(x)$$

is called a **steady-state distribution** of a Markov chain $X(t)$.

When this limit exists, it can be used as a forecast of the distribution of X after many transitions.

Computing Steady-State Distribution

When a steady-state distribution π exists, it can be computed as follows.

We notice that π is a limit of not only P_h but also P_{h+1}. The latter two are related by the formula

$$P_h P = P_0 P_h P = P_0 P_{h+1} = P_{h+1}$$

.

Taking the limit of P_h and P_{h+1}, as $h \to \infty$, we obtain

$$\pi P = \pi.$$

Then, by solving it for π, we get the steady-state distribution of a Markov chain with a transition probability matrix P.

Jumping Frog Example

Example. Recall a frog who hops on 3 rocks, with one-step **transition** probability matrix:

$$P = \begin{pmatrix} 0 & 1/2 & 1/2 \\ 5/8 & 1/8 & 1/4 \\ 2/3 & 1/3 & 0 \end{pmatrix}$$

Let us find its steady state distribution, i.e. limiting distribution π, by solving $\pi P = \pi$:

$$\begin{aligned} \frac{5}{8}\pi_2 + \frac{2}{3}\pi_3 &= \pi_1 \\ \frac{1}{2}\pi_1 + \frac{1}{8}\pi_2 + \frac{1}{3}\pi_3 &= \pi_2 \\ \frac{1}{2}\pi_1 + \frac{1}{4}\pi_2 &= \pi_3. \end{aligned}$$

Jumping Frog Example

This linear system together with $\sum_{i=1}^{3} \pi_i = 1$ (recall why we have this condition) yields

$$\pi = \left(\frac{38}{97}, \frac{32}{97}, \frac{27}{97} \right).$$