Markov Chains 1

Mihir Arjunwadkar

Centre for Modeling and Simulation
Savitribai Phule Pune University

Stochastic process

Stochastic process $\{X_t : t \in T\}$ :: A collection of random variables

Often written as $X(t) \equiv X_t$

State space $\mathcal{X}$ :: The set of values taken by the RVs X

Index set :: The set T, often interpreted/referred to as "time"

The state space and the index set can be discrete or continuous depending on the application

All of Statistics by Larry Wasserman, Springer (2004) Ch 23

- A sequence of IID RVs $X_1, X_2, \ldots \sim f$ is a stochastic process. State space $\mathcal{X}$: discrete or continuous depending on the RVs. Index set $T = \{1, 2, \ldots\}$.

All of Statistics by Larry Wasserman, Springer (2004) Ch 23

Examples of stochastic processes

- Suppose weather is modeled as having states $\mathcal{X} = \{$sunny, cloudy$\}$. Index set is $T = \{1, 2, \ldots\}$ in days. A realization of this stochastic process might be

  $$\text{sunny, sunny, cloudy, sunny, cloudy, cloudy, } \ldots$$

All of Statistics by Larry Wasserman, Springer (2004) Ch 23

- Empirical distribution function

$$\widehat{F}_n(t) = \frac{1}{n}\sum_{i=1}^{n} I(X_i \leq t)$$

State space $\mathcal{X}$: set [0,1] of values taken by $\widehat{F}_n(t)$; continuous.
Index set T: set of values taken by t; continuous.

All of Statistics by Larry Wasserman, Springer (2004) Ch 23

Joint PDF of a finite-length stochastic process

$$
\begin{aligned}
f(x_1, x_2, \ldots, x_n) &= f(x_1) \\
&\times f(x_2|x_1) \\
&\times f(x_3|x_2, x_1) \\
&\times \ldots \\
&\times f(x_n|x_{n-1}, \ldots, x_2, x_1) \\
&= \prod_{i=1}^{n} f(x_i|\text{past}_i)
\end{aligned}
$$

where

$$\text{past}_i = x_{i-1}, \ldots, x_2, x_1$$

and

$$\text{past}_1 = \{\}$$

We will assume

discrete & finite state space

$$\mathcal{X} = \{1, \ldots, k\}$$

index set

$$T = \{1, 2, \ldots\}$$

Definition. The process $\{X_n : n \in T\}$ is a called *Markov chain* if

$$P(x_n|x_1, \ldots, x_{n-1}) = P(x_n|x_{n-1})$$

for all $n \in T$ and for all $x \in \mathcal{X}$.

The present (n) depends only on the immediate past ($n-1$).

The present is *conditionally independent* of the past.

A Markov chain is *memoryless*.

All of Statistics by Larry Wasserman, Springer (2004) Ch 23

Markov chain

RA Howard, Dynamic Probabilistic Systems, vol. 1 (NY: John Wiley and Sons, 1971)

Joint PMF of a Markov chain $X_1, \ldots, X_n$

$$
\begin{aligned}
P(x_1, \ldots, x_n) &= P(x_1)P(x_2|x_1)P(x_3|x_2)\ldots P(x_n|x_{n-1}) \\
&= P(x_1)\prod_{i=2}^{n} P(x_i|x_{i-1}).
\end{aligned}
$$

All of Statistics by Larry Wasserman, Springer (2004) Ch 23

Complete specification of a finite portion of a Markov chain requires specifying

the initial PMF $P(x_1)$

and the

transition probabilities $P(x_2|x_1)$, $P(x_3|x_2)$, …, $P(x_n|x_{n-1})$.

All of Statistics by Larry Wasserman, Springer (2004) Ch 23

Definition. If $P(x_n|x_{n-1})$ is independent of n (i.e., does not change from time to time), the Markov chain is called a *homogeneous* Markov chain.

This means that

$$P(x_n = i | x_{n-1} = j) = P(x_{n-1} = i | x_{n-2} = j) = \ldots = P(x_2 = i | x_1 = j)$$

for every pair of states i, j.

That is, the transition probabilities do not change with time.

All of Statistics by Larry Wasserman, Springer (2004) Ch 23

A homogeneous Markov is completely specified by the time-independent state-to-state *transition probability matrix* (AKA *Markov matrix*, *stochastic matrix*)

$$\mathcal{P} = \begin{array}{c} \\ 1 \\ 2 \\ \vdots \\ k \end{array} \begin{array}{c} \begin{array}{cccc} 1 & 2 & \dots & k \end{array} \\ \begin{bmatrix} P(X_2 = 1|X_1 = 1) & P(X_2 = 2|X_1 = 1) & \dots & P(X_2 = k|X_1 = 1) \\ P(X_2 = 1|X_1 = 2) & P(X_2 = 2|X_1 = 2) & \dots & P(X_2 = k|X_1 = 2) \\ \vdots & \vdots & \vdots & \vdots \\ P(X_2 = 1|X_1 = k) & P(X_2 = 2|X_1 = k) & \dots & P(X_2 = k|X_1 = k) \end{bmatrix} \end{array}$$

Rows :: X_1 :: the immediate past (the present)

Columns :: X_2 :: the present (immediate future)

Row & column labels :: the states of the Markov chain

i, jth element

$$\mathcal{P}_{ij} = P(X_2 = j | X_1 = i) \geq 0$$

represents the probability that the Markov chain makes a transition from state i to state j in one time step.

Notice the opposite ordering of i, j on the two sides of the above equation. We will stick to this convention.

A property (aka *row normalization*) of the TPM:

$$\sum_{j=1}^{k} \mathcal{P}_{ij} = 1.$$

All of Statistics by Larry Wasserman, Springer (2004) Ch 23

Homogeneous Markov chains: Example 1

On any day, a machine can be in one of the two states *operational*, or *broken*. If operational, it may break down the next day with probability p $(0 \leq p \leq 1)$. If broken, it may be repaired by the next day with probability q $(0 \leq q \leq 1)$. This behaviour does not change on a day-to-day basis, nor does it depend on the history of the machine.

State space $\mathcal{X} = \{$ operational, broken $\}$ :: discrete, finite

Index set $T = \{ 1, 2, \ldots \}$, denoting days from the start date :: discrete

Let us assume that this is a Markov chain (i.e., only the immediate past determines the present).

This is a homogeneous Markov chain because the probabilities q, p do not depend on the time index (day).

$$\mathcal{P} = \begin{array}{r} \\ \text{operational} \\ \text{broken} \end{array} \begin{array}{c} \begin{array}{cc} \text{operational} & \text{broken} \end{array} \\ \left[\begin{array}{cc} 1-p & p \\ q & 1-q \end{array} \right] \end{array}$$

Introduction to Stochastic Processes by Hoel, Port & Stone, Houghton Mifflin (1972)

Homogeneous Markov chains: Example 2

1D discrete random walk over integers $\{1, \ldots, 5\}$ with absorbing end points: Except for the two end points $(1, 5)$, the probabilities for left and right moves (1 unit) from the current location are p and $1 - p$ respectively. The random walker stays at an end point forever if it reaches there.

State space $\mathcal{X} = \{\ 1,\ \ldots,\ 5\ \}$ :: discrete, finite

Index set $T = \{\ 1,\ 2,\ \ldots\}$ :: discrete

This is a Markov chain because only the current location determines the next location of the random walker.

This is a homogeneous Markov chain because the probability p does not depend on the time index.

$$\mathcal{P} = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{array} \begin{array}{c} \begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \end{array} \\ \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ p & 0 & 1-p & 0 & 0 \\ 0 & p & 0 & 1-p & 0 \\ 0 & 0 & p & 0 & 1-p \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \end{array}$$

Homogeneous Markov chains: Example 3

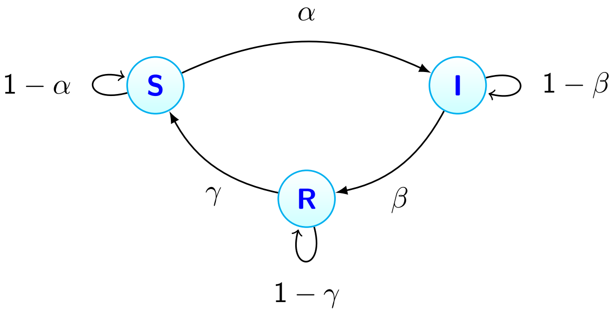
Each time a certain horse runs in a three-horse race, he has probability 1/2 of winning, 1/5 of coming in second, and 3/10 of coming in third, independent of the outcome of any previous race. This is an independent trials process, but it can also be considered a homogeneous Markov chain with the TPM

$$\mathcal{P} = \begin{array}{r} \\ \text{win} \\ \text{second} \\ \text{third} \end{array} \begin{array}{c} \begin{array}{ccc} \text{win} & \text{second} & \text{third} \end{array} \\ \begin{bmatrix} 1/2 & 1/5 & 3/10 \\ 1/2 & 1/5 & 3/10 \\ 1/2 & 1/5 & 3/10 \end{bmatrix} \end{array}$$

Independence $\rightarrow$ identical rows of the TPM

Introduction to Probability by Grinstead & Snell, American Mathematical Society (1997)

A Markov model of an SIR epidemic



$$\begin{array}{c} \\ \mathbf{S} \\ \mathbf{I} \\ \mathbf{R} \end{array} \begin{array}{c} \begin{array}{ccc} \mathbf{S} & \mathbf{I} & \mathbf{R} \end{array} \\ \begin{bmatrix} 1-\alpha & \alpha & 0 \\ 0 & 1-\beta & \beta \\ \gamma & 0 & 1-\gamma \end{bmatrix} \end{array}.$$

More examples of homogeneous Markov chains

- *Introduction to Probability* by Grinstead & Snell, American Mathematical Society (1997)
- *Introduction to Stochastic Processes* by Hoel, Port & Stone, Houghton Mifflin (1972)
- Umpteen resources on the internet!