Overview of Markov Chains

Sergio Pino

Outline

- Will talk about
  - Motivation: There is no motivation just math.
  - A very brief introduction to Markov chains
  - Weather model as an example.
  - Sampling the Markov chain from traces.
- Will not talk about
  - Chapman-Kolmogorov equations, classification of states, asymptotic analysis (limiting probabilities),...

(The slides content are based on [2-7])

Motivation [1]

SIAM Rev., 48(3), 569–581. (13 pages)

The $25,000,000,000 Eigenvector: The Linear Algebra behind Google

Kurt Bryan and Tanya Leise

DOI:10.1137/050623280

Google's success derives in large part from its PageRank algorithm, which ranks the importance of web pages according to an eigenvector of a weighted link matrix. Analysis of the PageRank formula provides a wonderful applied topic for a linear algebra course. Instructors may assign this article as a project to more advanced students or spend one or two lectures presenting the material with assigned homework from the exercises. This material also complements the discussion of Markov chains in matrix algebra. Maple and *Mathematica* files supporting this material can be found at www.rose-hulman.edu/~bryan.

Related Database

Web of Science

You must be logged in wit
active subscription to view

Article Data

History

Published online: 03 Augu

Keywords

Intro to Markov chains [2, 3, 6]

Def: A stochastic process $\{X_t, t \in T\}$ is a collection of random variables X_t, $t \in T$. It is determined by
(i). State space S, i.e. the range of possible value of X_t.
(ii). index set T which can be discrete (finite or countable) or continuous.
(iii). dependency relations between variables.

Def: A stochastic process $\{X_n, n = 0, 1, 2 \cdots\}$ is a Markov Chain if $S = \{0, 1, 2, \cdots\}$ and

$$\mathbb{P}\left(X_{n+1} = j | X_n = i, X_{n-1} = i_{n-1}, \cdots, X_0 = i_0\right) = P_{ij}$$

for all $i_0, i_1, \cdots, i_{n-1}, i, j \in S$ and $n \geq 0$.

- $X_n = i$ means that the process is in the state $i \in S$ at the time n.

Intro to Markov chains

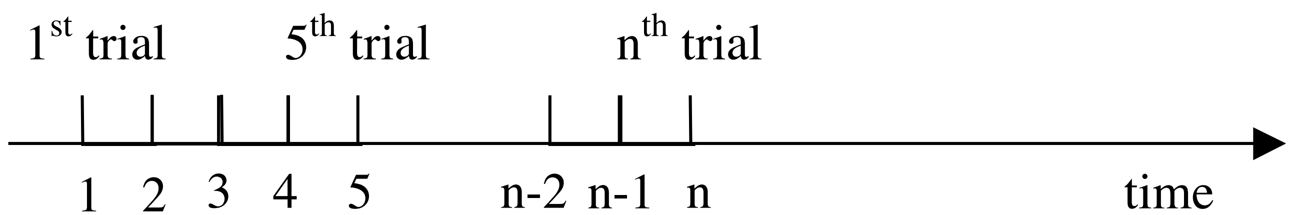
$(X_n)_{n\geq 1}$ is said to be a Markov chain if and only if:

$$\boxed{\forall\ n, i_1, i_2, \ldots, i_n \quad P\left\{X_n = i_n / \bigcap_{k=1}^{n-1} \{X_k = i_k\}\right\} = P\{X_n = i_n / X_{n-1} = i_{n-1}\}}$$

This relation is known as the Markov property.

Intro to Markov chains

If we associate a time scale to the sequence of trials,



n corresponds to the future
n-1 corresponds to the present
1 to(n-2) corresponds to the past

Then, the Markov property can be stated as follows:

$$\boxed{P\{Future/Present\ and\ Past\} = P\{Future / Present\}}$$

Intro to Markov chains

Furthermore $(X_n)_{n\geq 1}$ is homogeneous if and only if:

$\forall\, n, j, i \quad P\{X_n = j / X_{n-1} = i\}$ does not depend on n.

So we can denote $p_{ij} = P\{X_n = j / X_{n-1} = i\}$

Intro to Markov chains

These conditional probabilities p_{ij} are called transition probabilities. If the number of states is finite (for instance n_0), they can be arranged in a transition probability matrix T so that the first subscript (i) stands for row and the second (j) for column. T is a square matrix ($n_0 \times n_0$) with non negative elements and unit row sums.

$$\forall\, i,j \quad 0 \le p_{ij} \le 1 \quad and \quad \forall\, i \quad \sum_{j=1}^{n_0} p_{ij} = 1$$

$$T = \begin{array}{c} \\ E_1 \\ E_2 \\ \\ E_{n0} \end{array} \begin{array}{c} \begin{array}{ccc} E_1 & E_2 & E_{n0} \end{array} \\ \begin{pmatrix} p_{11} & p_{12} & p_{1n0} \\ p_{21} & p_{22} & p_{2n0} \\ & & \\ p_{n01} & p_{n02} & p_{n0n0} \end{pmatrix} \end{array}$$

Example [2, 6]

Rain or no rain, two states Markov Chain.

$$\mathbf{P} = \begin{pmatrix} \alpha & 1-\alpha \\ \beta & 1-\beta \end{pmatrix}$$

$$\mathbf{P} = \begin{bmatrix} 0.7 & 0.3 \\ 0.4 & 0.6 \end{bmatrix}$$

Graph sketch

seq: **S R R R R R R R R R R R S R R S**
R R R R R R S R R S S S S S S ...

Sampling the Markov chain from traces [4, 5, 7]

Let's do this! Use the following equations to estimate:

$$\pi_i = \frac{c_i}{L} \quad i = \overline{1,N} \tag{1.3}$$

$$a_{ij} = \frac{c_{ij}}{\sum_{j=1}^{N} c_{ij}} \quad i,j = \overline{1,N} \qquad \sum_{i=1}^{N}\sum_{j=1}^{N} c_{ij} = L-1 \tag{1.4}$$

Training sequence:

S R R R R R R R R R R R S R R S R R R R R R S R R S S S S S S

Questions

References

- [1] Bryan, Kurt, and Tanya Leise. "The $25,000,000,000 eigenvector: The linear algebra behind Google." Siam Review 48.3 (2006): 569-581.
- [2] Wenbo Li. Math 630: Probability Theory and Application (class notes). University of Delaware 2012.
- [3] http://www-public.it-sudparis.eu/~uro/
- [4] http://www.eecis.udel.edu/~lliao/cis841s06/hmmtutorialpart1.pdf
- [5] S. Jha, K. Tan, and R. A. Maxion, “Markov chains, classifiers, and intrusion detection,” in Proceedings. 14th IEEE Computer Security Foundations Workshop, 2001., 2001, pp. 206–219. http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=930147&tag=1
- [6] Ross, Sheldon M. Introduction to probability models. Academic press, 2006.
- [7] Jones, Matthew T. Estimating Markov transition matrices using proportions data: an application to credit risk. International Monetary Fund, 2005.https://www.imf.org/external/pubs/cat/longres.aspx?sk=18387.0