MATH 304

Linear Algebra

Lecture 38:
Markov chains.

Stochastic process

Stochastic (or **random**) **process** is a sequence of experiments for which the outcome at any stage depends on a chance.

Simple model:

- a finite number of possible outcomes (called **states**);
- discrete time

Let S denote the set of the states. Then the stochastic process is a sequence $s_0, s_1, s_2, \ldots,$ where all $s_n \in S$ depend on chance.

How do they depend on chance?

Bernoulli scheme

Bernoulli scheme is a sequence of independent random events.

That is, in the sequence $s_0, s_1, s_2, \dots$ any outcome s_n is independent of the others.

For any integer $n \geq 0$ we have a probability distribution $p^{(n)}$ on S. This means that each state $s \in S$ is assigned a value $p_s^{(n)} \geq 0$ so that $\sum_{s \in S} p_s^{(n)} = 1$. Then the probability of the event $s_n = s$ is $p_s^{(n)}$.

The Bernoulli scheme is called **stationary** if the probability distributions $p^{(n)}$ do not depend on n.

Examples of Bernoulli schemes:

- Coin tossing

2 states: heads and tails. Equal probabilities: $1/2$.

- Die rolling

6 states. Uniform probability distribution: $1/6$ each.

- Lotto Texas

Any state is a 6-element subset of the set $\{1, 2, \ldots, 54\}$. The total number of states is $25,827,165$. Uniform probability distribution.

Markov chain

Markov chain is a stochastic process with discrete time such that the probability of the next outcome may depend only on the previous outcome.

Let $S = \{1, 2, \ldots, k\}$. The Markov chain is determined by **transition probabilities** $p_{ij}^{(t)}$, $1 \leq i, j \leq k$, $t \geq 0$, and by the **initial** probability distribution q_i, $1 \leq i \leq k$.

Here q_i is the probability of the event $s_0 = i$, and $p_{ij}^{(t)}$ is the conditional probability of the event $s_{t+1} = j$ provided that $s_t = i$. By construction, $p_{ij}^{(t)}, q_i \geq 0$, $\sum_i q_i = 1$, and $\sum_j p_{ij}^{(t)} = 1$.

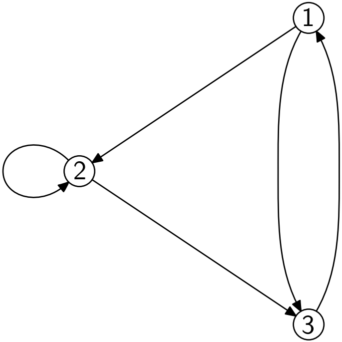
We shall assume that the Markov chain is time-independent, i.e., transition probabilities do not depend on time: $p_{ij}^{(t)} = p_{ij}$.

Then a Markov chain on $S = \{1, 2, \dots, k\}$ is determined by a **probability vector** $\mathbf{x}_0 = (q_1, q_2, \dots, q_k) \in \mathbb{R}^k$ and a $k \times k$ **transition matrix** $P = (p_{ij})$. The entries in each row of P add up to 1.

Let $s_0, s_1, s_2, \dots$ be the Markov chain. Then the vector $\mathbf{x}_0$ determines the probability distribution of the initial state s_0.

Problem. Find the (unconditional) probability distribution for any s_n.

Example: random walk



Transition matrix: $P = \begin{pmatrix} 0 & 1/2 & 1/2 \\ 0 & 1/2 & 1/2 \\ 1 & 0 & 0 \end{pmatrix}$

Problem. Find the (unconditional) probability distribution for any s_n, $n \geq 1$.

The probability distribution of s_{n-1} is given by a probability vector $\mathbf{x}_{n-1} = (a_1, \ldots, a_k)$. The probability distribution of s_n is given by a vector $\mathbf{x}_n = (b_1, \ldots, b_k)$.

We have

$$b_j = a_1 p_{1j} + a_2 p_{2j} + \cdots + a_k p_{kj}, \quad 1 \leq j \leq k.$$

That is,

$$(b_1, \ldots, b_k) = (a_1, \ldots, a_k) \begin{pmatrix} p_{11} & \ldots & p_{1k} \\ \vdots & \ddots & \vdots \\ p_{k1} & \ldots & p_{kk} \end{pmatrix}.$$

$$\mathbf{x}_n = \mathbf{x}_{n-1}P \implies \mathbf{x}_n^T = (\mathbf{x}_{n-1}P)^T = P^T\mathbf{x}_{n-1}^T.$$

Thus $\mathbf{x}_n^T = Q\mathbf{x}_{n-1}^T$, where $Q = P^T$ and the vectors are regarded as row vectors.

Then $\mathbf{x}_n^T = Q\mathbf{x}_{n-1}^T = Q(Q\mathbf{x}_{n-2}^T) = Q^2\mathbf{x}_{n-2}^T$.

Similarly, $\mathbf{x}_n^T = Q^3\mathbf{x}_{n-3}^T$, and so on.

Finally, $\boxed{\mathbf{x}_n^T = Q^n\mathbf{x}_0^T.}$

Example. Very primitive weather model:

Two states: "sunny" (1) and "rainy" (2).

Transition matrix: $P = \begin{pmatrix} 0.9 & 0.1 \\ 0.5 & 0.5 \end{pmatrix}$.

Suppose that $\mathbf{x}_0 = (1, 0)$ (sunny weather initially).

Problem. Make a long-term weather prediction.

The probability distribution of weather for day n is given by the vector $\mathbf{x}_n^T = Q^n \mathbf{x}_0^T$, where $Q = P^T$.

To compute Q^n, we need to diagonalize the matrix $Q = \begin{pmatrix} 0.9 & 0.5 \\ 0.1 & 0.5 \end{pmatrix}$.

$$\det(Q - \lambda I) = \begin{vmatrix} 0.9-\lambda & 0.5 \\ 0.1 & 0.5-\lambda \end{vmatrix} =$$

$$= \lambda^2 - 1.4\lambda + 0.4 = (\lambda - 1)(\lambda - 0.4).$$

Two eigenvalues: $\lambda_1 = 1$, $\lambda_2 = 0.4$.

$$(Q - I)\mathbf{v} = \mathbf{0} \iff \begin{pmatrix} -0.1 & 0.5 \\ 0.1 & -0.5 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\iff (x, y) = t(5, 1), \; t \in \mathbb{R}.$$

$$(Q - 0.4I)\mathbf{v} = \mathbf{0} \iff \begin{pmatrix} 0.5 & 0.5 \\ 0.1 & 0.1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\iff (x, y) = t(-1, 1), \; t \in \mathbb{R}.$$

$\mathbf{v}_1 = (5, 1)^T$ and $\mathbf{v}_2 = (-1, 1)^T$ are eigenvectors of Q belonging to eigenvalues 1 and 0.4, respectively.

$$\mathbf{x}_0^T = \alpha\mathbf{v}_1 + \beta\mathbf{v}_2 \iff \begin{cases} 5\alpha - \beta = 1 \\ \alpha + \beta = 0 \end{cases} \iff \begin{cases} \alpha = 1/6 \\ \beta = -1/6 \end{cases}$$

Now $\mathbf{x}_n^T = Q^n\mathbf{x}_0^T = Q^n(\alpha\mathbf{v}_1 + \beta\mathbf{v}_2) =$

$$= \alpha(Q^n\mathbf{v}_1) + \beta(Q^n\mathbf{v}_2) = \alpha\mathbf{v}_1 + (0.4)^n\beta\mathbf{v}_2,$$

which converges to the vector $\alpha\mathbf{v}_1 = (5/6, 1/6)^T$ as $n \to \infty$.

The vector $\mathbf{x}_\infty = (5/6, 1/6)$ gives the **limit distribution**. Also, it is a **steady-state** vector.

Remarks. In this example, the limit distribution does not depend on the initial distribution, but it is not always so. However 1 is always an eigenvalue of the matrix P (and hence Q) since $P\,(1, 1, \ldots, 1)^T = (1, 1, \ldots, 1)^T$.