Markov processes

Assume that a system can be in one of the following states, encoded by \( \{1, 2, \ldots, k\} \).
Let \( Y_0, Y_1, \ldots Y_n, \ldots \) be the successive outcomes of an experiment.

**Definition:** We say that \( Y_1, Y_2, \ldots Y_n, \ldots \) is a Markov process (or a Markov chain), if knowing the present state of the system (i.e. knowing \( Y_n \), the outcome of the \( n \)-th experiment), the next state of the system (i.e. \( Y_{n+1} \)) can be predicted with a certain probability.

Note that in a Markov system, given the present state of the system, the next state is independent of the past states. We call such system memoryless.

Note also that a Markov process is a stationary process: the probability to transition from a state \( j \) to a state \( i \) depends only on the states \( j \) and \( i \) and not on the number \( n \) of experiments performed.
The transition matrix, state vectors

Let $Y_0, Y_1, \ldots Y_n, \ldots$ be a Markov process with possible states \{1, 2, \ldots, k\}. Let $p_{ij}$ denote the probability to transition from the present state $j$ to the next state $i$. Then the $k \times k$ matrix

$$P = [p_{ij}] \quad 1 \leq i, j \leq k$$

is called the transition matrix of the given Markov process.

We define the state vectors of the given Markov process to be the vectors $x^{(0)}, x^{(1)}, \ldots, x^{(n)}, \ldots$

$$x^{(n)} = \begin{bmatrix} x_1^{(n)} \\ x_2^{(n)} \\ \vdots \\ x_k^{(n)} \end{bmatrix} \in \mathbb{R}^k$$

where $x_1^{(n)} = \text{probability that } Y_n = 1$, $x_2^{(n)} = \text{probability that } Y_2 = 2$, \ldots, $x_k^{(n)} = \text{probability that } Y_n = k$. 
Stochastic matrices, probability vectors

Definition: A vector

\[ x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \end{bmatrix} \in \mathbb{R}^k \]

is called a probability vector if \( x_i \geq 0 \) for all \( i \) and \( x_1 + x_2 + \ldots + x_k = 1 \).

Definition: A square matrix \( P = [p_{ij}] \), \( 1 \leq i, j \leq k \) is called stochastic (or Markovian) if its column vectors are probability vectors, in other words, if \( p_{i,j} \geq 0 \) for all \( i, j \) and

\[ p_{1j} + p_{2j} + \ldots + p_{kj} = 1 \quad \text{for all } j \]

Given a Markov process, it is clear that its state vectors are probability vectors and its transition matrix is a stochastic matrix.
The transition matrix, state vectors

Let $Y_0, Y_1, \ldots Y_n, \ldots$ be a Markov process with possible states $\{1, 2, \ldots, k\}$. Let $P = [p_{ij}]$ $1 \leq i, j \leq k$ be its transition matrix and let $x^{(0)}, x^{(1)}, \ldots, x^{(n)}, \ldots$ be its state vectors.

**Theorem**: The following relations hold for all times $n$:

$$x^{(n)} = P x^{(n-1)}, \text{ so }$$

$$x^{(n)} = P^n x^{(0)}$$

It turns out that regardless of the initial state vector $x^{(0)}$ of the system, if the transition matrix is “nice enough”, the state vectors of the process converge to a vector called the *steady state vector* of the Markov chain. We explain this in the next slide.
Convergence to a steady state vector

Definition: A transition matrix $P$ is called regular if there is a power $m$ such that $P^m > 0$, meaning that all entries of $P^m$ are positive.

The following is a consequence of an important theorem in linear algebra, called Perron-Frobenius theorem.

Theorem: If $P$ is a regular transition matrix, for any probability vector $x$ we have

$$P^n x \rightarrow q \quad \text{as } n \rightarrow \infty$$

where $q$ is probability vector with all entries positive, called a steady state vector.

The following theorem gives us a simple way to compute the steady state vector of a Markov process given by a regular transition matrix.

Theorem: The vector $q$ in the previous theorem is the unique vector satisfying:

1. $Pq = q$, which is equivalent to $(I - P)q = 0$.
2. The entries $q_1, q_2, \ldots, q_k$ of $q$ are positive and their sum is 1.