Imperial College London

Maths for Signals and Systems

Linear Algebra in Engineering

Lecture 12, Friday 6th November 2015

DR TANIA STATHAKI

READER (ASSOCIATE PROFFESOR) IN SIGNAL PROCESSING
IMPERIAL COLLEGE LONDON

Stochastic matrices

- Consider a matrix A with the following properties:
  - ➢ All entries are positive and real.
  - ➢ The elements of each column or each row or both each column and each row add up to 1.
  - ➢ Based on the above a matrix that exhibits the above properties will have all entries ≤ 1.
  - ➢ It is square.
- This is called a **stochastic matrix**.
- Stochastic matrices are also called **Markov, probability, transition, or substitution matrices**.
- The entries of a stochastic matrix usually represent a probability.
- Stochastic matrices are widely used in probability theory, statistics, mathematical finance and linear algebra, as well as computer science and population genetics.

Imperial College London

Stochastic matrices. Types.

- There are several types of stochastic matrices:
  - A **right stochastic matrix** is a matrix of nonnegative real entries, with each row's elements summing to 1.
  - A **left stochastic matrix** is a matrix of nonnegative real entries, with each column's elements summing to 1.
  - A **doubly stochastic matrix** is a matrix of nonnegative real entries with each row's and each column's elements summing to 1.
- A stochastic matrix often describes a so called **Markov chain** X_t over a finite state space S.
- Generally, an $n \times n$ stochastic matrix is related to n "states".
- If the probability of moving from state i to state j in one time step is $P_r(j/i) = [p_{ij}]$, the stochastic matrix P is given by using p_{ij} as the i^{th} row and j^{th} column element:

$$P = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{bmatrix}$$

- Depending on the particular problem the above matrix can be formulated in such a way so that is either right or left stochastic.

Products of stochastic matrices. Stochastic vectors.

- An example of a left stochastic matrix is the following:

$$A = \begin{bmatrix} 0.1 & 0.01 & 0.3 \\ 0.2 & 0.99 & 0.3 \\ 0.7 & 0 & 04 \end{bmatrix}$$

- You can prove that if A and B are stochastic matrices of any type then AB is also a stochastic matrix of the same type.
- Consider two left stochastic matrices A and B with elements a_{ij} and b_{ij} respectively, and $C = AB$ with elements c_{ij}.
- Let us find the sum of the elements of the j^{th} column of C:

$\sum_{i=1}^{n} c_{ij} = \sum_{i=1}^{n} \sum_{k=1}^{n} a_{ik} b_{kj} = \sum_{k=1}^{n} \sum_{i=1}^{n} a_{ik} b_{kj} = \sum_{k=1}^{n} b_{kj} \sum_{i=1}^{n} a_{ik} = 1 \cdot 1 = 1$

- Based on the above the power of a stochastic matrix is a stochastic matrix.
- I am interested in the eigenvalues and eigenvectors of a stochastic matrix.
- Let's call a vector with real, nonnegative entries p_k, for which all the p_k add up to 1, a **stochastic vector**. For a stochastic matrix, every column or row or both is a stochastic vector.

Stochastic matrices and their eigenvalues

- I would like to prove that $\lambda = 1$ is always an eigenvalue of a stochastic matrix.
- Consider again a left stochastic matrix A.
- Since the elements of each column of A add up to 1, the elements of each row of A^T should add up to 1. Therefore:

$$A^T \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} = 1 \cdot \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$$

- Therefore, 1 is an eigenvalue of A^T . The eigenvalues of A^T are obtained through the equation $\det(A^T - \lambda I) = 0$. But:

$$\det(A^T - \lambda I) = \det(A^T - \lambda I^T) = \det[(A - \lambda I)^T] = \det(A - \lambda I)$$

- Hence, the eigenvalues of A and A^T are the same, which implies that 1 is also an eigenvalue of A.
- Since $\det(A - I) = 0$, the matrix $A - I$ is singular, which means that there is a vector x for which

$$(A - I)x = 0 \Rightarrow Ax = x$$

- A vector of the null space of $A - I$ is the eigenvector of A that corresponds to eigenvalue $\lambda = 1$.

Stochastic matrices and their eigenvalues cont.

- I would like now to prove that the eigenvalues of a stochastic matrix have magnitude smaller or equal to 1.
- Assume that $v = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$ is an eigenvector of a stochastic matrix A with an eigenvalue $|\lambda| > 1$. Then $A^n v = \lambda^n v$ implies $\sum_{j=1}^{n} [A^n]_{ij} v_j = \lambda^n v_i$
- λ^n has exponentially growing length for $n \to \infty$. The maximum value that $\sum_{j=1}^{n} [A^n]_{ij} v_j$ can take is $v_{\max} \cdot \max_i \sum_{j=1}^{n} [A^n]_{ij} \le v_{\max} \cdot n$ since the entries of A^n are ≤ 1 .
- The relationship $\sum_{j=1}^{n} [A^n]_{ij} v_j = \lambda^n v_i$ must be valid for any n. There is always a minimum n for which $\sum_{j=1}^{n} [A^n]_{ij} v_j \le v_{\max} \cdot n < \lambda^n v_i$ if $|\lambda| > 1$.
- Based on the above, the assumption of an eigenvalue being larger than 1 can not be valid.

Application of stochastic matrices

- Consider again the system described by an equation of the form $u_k = A^k u_0$, where A is now a stochastic matrix.
- Previously, we managed to write $u_k = A^k u_0 = c_1 {\lambda_1}^k x_1 + c_2 {\lambda_2}^k x_2 + \cdots$ where λ_i and x_i are the eigenvalues and eigenvectors of matrix A, respectively.
- Note that the above relationship requires a complete set of eigenvectors.
- If $\lambda_1 = 1$ and $|\lambda_i| < 1, i > 1$ then the steady state of the system is $c_1 x_1$ (which is part of the initial condition u_0).
- I will use an example where A is a 2×2 matrix. Generally, an $n \times n$ stochastic matrix is related to n "states". Assume that the 2 "states" are 2 UK cities.
- I take London and Oxford. I am interested in the population of the two cities and how it evolves.
- I assume that people who inhabit these two cities move between them only.

$$\begin{bmatrix} u_{\text{ox}} \\ u_{\text{lon}} \end{bmatrix}_{t=k+1} = \begin{bmatrix} 0.9 & 0.2 \\ 0.1 & 0.8 \end{bmatrix} \begin{bmatrix} u_{\text{ox}} \\ u_{\text{lon}} \end{bmatrix}_{t=k}$$

- It is now obvious that the column elements are positive and also add up to 1 because they represent probabilities.

Application of stochastic matrices (cont.)

- I assume that $\begin{bmatrix} u_{\text{ox}} \\ u_{\text{lon}} \end{bmatrix}_{t=k=0} = \begin{bmatrix} 0 \\ 1000 \end{bmatrix}$. Then $\begin{bmatrix} u_{\text{ox}} \\ u_{\text{lon}} \end{bmatrix}_{k=1} = \begin{bmatrix} 0.9 & 0.2 \\ 0.1 & 0.8 \end{bmatrix} \begin{bmatrix} 0 \\ 1000 \end{bmatrix} = \begin{bmatrix} 200 \\ 800 \end{bmatrix}$.

Problem:

What is the population of the two cities after a long time?

Solution:

Consider the matrix $\begin{bmatrix} 0.9 & 0.2 \\ 0.1 & 0.8 \end{bmatrix}$. The eigenvalues are $\lambda_1 = 1$ and $\lambda_2 = 0.7$.

(Notice that the second eigenvalue is found by the trace of the matrix.)

The eigenvectors of this matrix are $x_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ and $x_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$

$$\begin{bmatrix} u_{\text{ox}} \\ u_{\text{lon}} \end{bmatrix}_k = c_1 {\lambda_1}^k \begin{bmatrix} 2 \\ 1 \end{bmatrix} + c_2 {\lambda_2}^k \begin{bmatrix} -1 \\ 1 \end{bmatrix} = c_1 \begin{bmatrix} 2 \\ 1 \end{bmatrix} + c_2 0.7^k \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

I find $c_1, \ c_2$ from the initial condition $\begin{bmatrix} u_{\text{ox}} \\ u_{\text{lon}} \end{bmatrix}_{t=k=0} = \begin{bmatrix} 0 \\ 1000 \end{bmatrix}$

$\begin{bmatrix} 0 \\ 1000 \end{bmatrix} = c_1 \begin{bmatrix} 2 \\ 1 \end{bmatrix} + c_2 \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ and therefore, $c_1 = \frac{1000}{3}$ and $c_2 = \frac{2000}{3}$

Application of stochastic matrices (cont.)

- Stochastic models facilitate the modeling of various real life engineering applications.

- An example is the modeling of the movement of people without gain or loss: total number of people is conserved.