Section 7.5 Stochastic Matrices

Stochastic Matrices are used to model transitions in sequences of states. For example, Markov chains.

These systems are "memoryless" - moving from one state to the next is only governed by probabilities.

We will often be concerned with the long term behaviour of such a system, in which case we will need to understand how to take limits.

Check out Google's PageRank algorithm.

Definitions: A probability vector is a column of non-negative entries which add up to 1.

$$\begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad \begin{bmatrix} 0.8 \\ 0.05 \\ 0.15 \end{bmatrix}$$

A stochastic matrix M is a real square matrix whose columns are probability vectors.

$$A = \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix} \qquad B = \begin{bmatrix} 0.9 & 0.2 \\ 0.1 & 0.8 \end{bmatrix} \qquad I_n$$

$$J_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \qquad E = \begin{bmatrix} 0.8 & 0.3 & 0.2 \\ 0.1 & 0.2 & 0.6 \\ 0.1 & 0.5 & 0.2 \end{bmatrix}$$

Note: Some people require rows to also add to 1. However, we'd like to be able to distinguish between the two types. So:

More Defintions: A doubly stochastic matrix has both rows and columns adding to 1.

For example: A, J_2, I_n

A steady-state vector is a vector, $\underline{v}$, such that $A\underline{v}=\underline{v}$.

But we will mostly talk about probability steady state vectors.

An idempotent matrix, A, is a square matrix satisfying: $A^2=A$.

For example:

$$A^2=\begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}\begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}=\begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}=A$$

Facts:

1. Idempotent matrices have eigenvalues 0 and 1 only.
2. If A is an idempotent matrix, $A^n=A$ for all $n \geq 1$.

Exercise: prove these.

In practice, it is more difficult to find an expression for a high power of a matrices. For example, the matrix B (from above) is not idempotent. In fact, B^n, is quite messy. However, we can still find the long term behaviour using $\lim_{n\to\infty} B^n$.

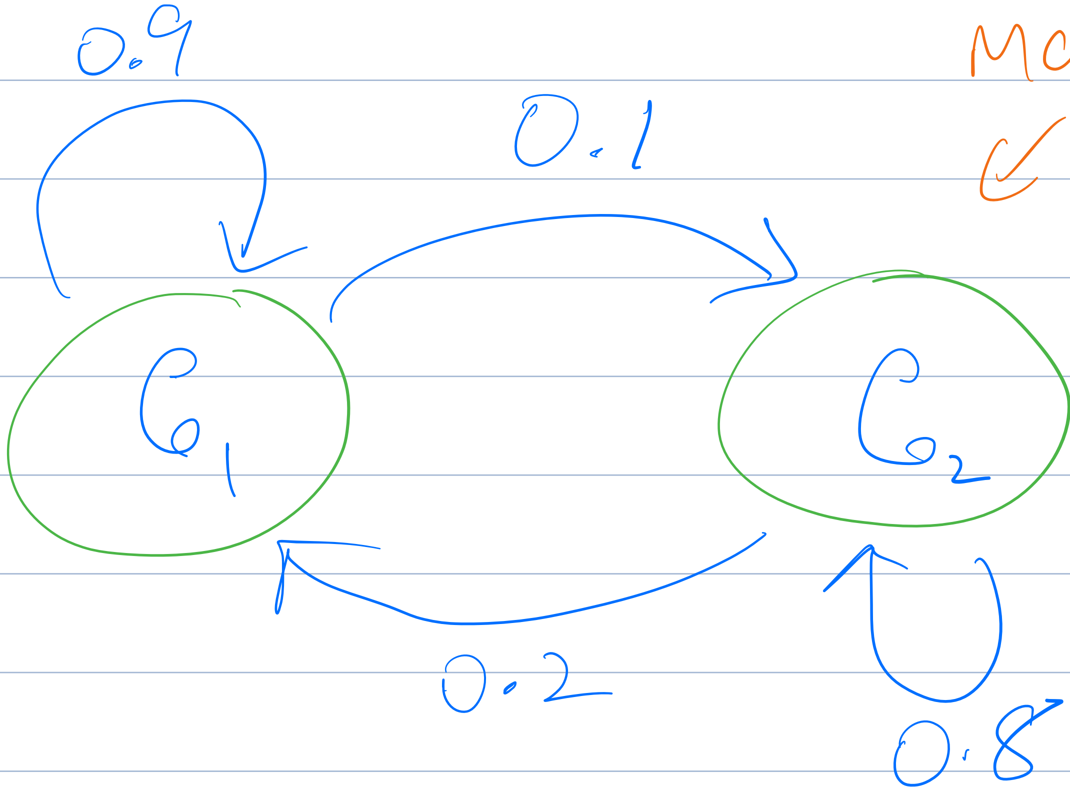
Example: Two companies, C_1 & C_2, control some particular market (eg. Coles and Woolworths).

Suppose the following are reasonably accurate:

1. Each season, C_1 retains 0.9 of it's custom but loses 0.1 to C_2.
2. Each season, C_2 retains 0.8 of it's custom but loses 0.2 to C_1.

It looks like C_1 has the edge over C_2. Will C_2 go out of business? Let's find out.

Markov Process



Let's convert the problem into matrix arithmetic.

At any given time, put

x = the proportion of the market held by C_1, and

y = the proportion of the market held by C_2.

Then x+y=1 since there are no other companies in the system.

Put $B = \begin{bmatrix} 0.9 & 0.2 \\ 0.1 & 0.8 \end{bmatrix}.$

Then the proportions of the market share after one season are

$$B\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0.9 & 0.2 \\ 0.1 & 0.8 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0.9x+0.2y \\ 0.1x+0.8y \end{bmatrix}.$$

proportion held by C_1 after one season.

proportion held by C_2 after one season.

After two seasons the proportions would be given by $B^2\begin{bmatrix} x \\ y \end{bmatrix}.$

After three seasons the proportions would be given by $B^3 \begin{bmatrix} x \\ y \end{bmatrix}$.

In general, the proportions after n seasons are given by $B^n \begin{bmatrix} x \\ y \end{bmatrix}$.

So, we want to understand B^n, for large n.

Let's diagonalise B.

$\boxed{\text{Eigenvalues}}$

$$\det(\lambda I - B) = \begin{vmatrix} \lambda - 0.9 & -0.2 \\ -0.1 & \lambda - 0.8 \end{vmatrix}$$

$$= (\lambda - 0.9)(\lambda - 0.8) - 0.02$$

$$= \lambda^2 - 0.17\lambda + 0.72 - 0.02$$

$$= \lambda^2 - \frac{17}{10}\lambda + \frac{7}{10}$$

$$= (\lambda - 1)(\lambda - 7/10)$$

So $\lambda = 1$ and $7/10$.

$\boxed{\text{Eigenspaces}}$

$\boxed{\lambda = 1}$ $I - B = \begin{bmatrix} 0.1 & -0.2 \\ -0.1 & 0.2 \end{bmatrix} \sim \begin{array}{c} \begin{array}{cc} x & y \end{array} \\ \begin{bmatrix} 1 & -2 \\ 0 & 0 \end{bmatrix} \end{array}$

Let $y = t \in \mathbb{R}$.

Then $x - 2t = 0 \Longrightarrow x = 2t$.

So eigenspace is $\left\{ t \begin{bmatrix} 2 \\ 1 \end{bmatrix} \,\middle|\, t \in \mathbb{R} \right\}$

Let's choose a particular eigenvector of $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$

$\boxed{\lambda = 7/10}$ $\frac{7}{10}I - B = \begin{bmatrix} -0.2 & -0.2 \\ -0.1 & -0.1 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$ let $y = s \in \mathbb{R}$

$\qquad\qquad\qquad\ \ x \quad y$

Then $x = -y = -s$.

So the eigenspace is $\left\{ s \begin{bmatrix} -1 \\ 1 \end{bmatrix} \,\middle|\, s \in \mathbb{R} \right\}$

Choose $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$ as our eigenvector.

Then $\qquad P = \begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix}, \; D = \begin{bmatrix} 1 & 0 \\ 0 & \frac{7}{10} \end{bmatrix}$

and then $\qquad P^{-1} = \frac{1}{2+1} \begin{bmatrix} 1 & 1 \\ -1 & 2 \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 1 & 1 \\ -1 & 2 \end{bmatrix}.$

So we can conjugate $B = PDP^{-1}$

which we can use to find $B^n = (PDP^{-1})^n = PD^nP^{-1}$

$$= \left(\begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \frac{7}{10}^n \end{bmatrix} \right) \frac{1}{3} \begin{bmatrix} 1 & 1 \\ -1 & 2 \end{bmatrix}$$

$$= \frac{1}{3} \begin{bmatrix} 2 & -\left(\frac{7}{10}^n\right) \\ 1 & \left(\frac{7}{10}\right)^n \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -1 & 2 \end{bmatrix}$$

$$= \frac{1}{3} \begin{bmatrix} 2 + \cancel{\left(\frac{7}{10}\right)^n}^{\,0} & 2 - 2\cancel{\left(\frac{7}{10}\right)^n}^{\,0} \\ 1 - \cancel{\left(\frac{7}{10}\right)^n}^{\,0} & 1 + 2\cancel{\left(\frac{7}{10}\right)^n}^{\,0} \end{bmatrix}$$

Hence $\qquad \lim_{n \to \infty} B^n = \frac{1}{3} \begin{bmatrix} 2 & 2 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} \frac{2}{3} & \frac{2}{3} \\ \frac{1}{3} & \frac{1}{3} \end{bmatrix}.$

So the long term effect on the initial proportions $\begin{bmatrix} x \\ y \end{bmatrix}$ is

$$\begin{bmatrix} \frac{2}{3} & \frac{2}{3} \\ \frac{1}{3} & \frac{1}{3} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \frac{2}{3}x + \frac{2}{3}y \\ \frac{1}{3}x + \frac{1}{3}y \end{bmatrix} = \begin{bmatrix} \frac{2}{3}(x+y) \\ \frac{1}{3}(x+y) \end{bmatrix} = \begin{bmatrix} \frac{2}{3} \\ \frac{1}{3} \end{bmatrix}$$

probability vector

So, regardless of the initial proportions, in the long term

C_1 captures 2/3 of the market, and

C_2 captures 1/3 of the market.

Note that since $B \begin{bmatrix} \frac{2}{3} \\ \frac{1}{3} \end{bmatrix} = \begin{bmatrix} 0.9 & 0.2 \\ 0.1 & 0.8 \end{bmatrix} \begin{bmatrix} \frac{2}{3} \\ \frac{1}{3} \end{bmatrix} = \begin{bmatrix} \frac{2}{3} \\ \frac{1}{3} \end{bmatrix}$,

equilibrium has been reached and $\begin{bmatrix} \frac{2}{3} \\ \frac{1}{3} \end{bmatrix}$ is a steady-state vector.

Theorems: Let A be a stochastic matrix.

1. 1 is an eigenvalue of A, and so A has at least one steady-state probability vector, $\underline{v}$, which is fixed by A. $A\underline{v} = 1\underline{v}$

2. All eigenvalues of A are less than or equal to 1 in magnitude.

3. If A is regular (some power of A has all positive entries), then:

   A. $\underline{v}$ is unique There is only 1 steady state vector. up to scalar multiple $A(\lambda \underline{v}) = \lambda A \underline{v} = \lambda \underline{v}$

   B. $\lim_{n\to\infty} A^n = [\underline{v}\ \underline{v} \cdots \underline{v}]$

   C. For any probability vector, $\underline{x}$, $\lim_{n\to\infty} A^n \underline{x} = \underline{v}$.

   D. 1 is the dominant eigenvalue.

Better A: There is only one probability steady state vector.

Proof: We only prove Theorem 1.

First note that any square matrix has the same eigenvalues as its transpose.

Let $A = [a_{ij}]$ be a $n \times n$ stochastic matrix.

Then $[1\ 1 \cdots 1]A = [\sum_{i=1}^{n} a_{i1}\ \sum_{i=1}^{n} a_{i2} \cdots \sum_{i=1}^{n} a_{in}]$

$=$