SOME MATRIX MODELS

Ana Paula Garrão, PhD
Margarida Raposo Dias, PhD
Universidade dos Açores, Departamento de Matemática, Portugal
Núcleo Interdisciplinar da Criança e do Adolescente da Universidade dos Açores, Portugal

Abstract

In this paper we present three matrix models that express the relevance of studying Linear Algebra, trying to answer the question *Why study Linear Algebra?*, often asked by students. Linear Algebra is fundamental within the mathematical modelling of phenomena, so we explore examples of models in which it is applied, in a simple manner as possible appropriated to the students' knowledge. The models presented are Leontief Economic Model, Leslie Population Growth Model and Rating Web Pages Model.

Keywords: Leontief matrix; Leslie matrix; PageRank

Introduction

As mathematics teachers in study programmes whose scientific area is not Mathematics, but have Mathematics as a required curricular unit, we are faced with the question: *Why study Linear Algebra?* To answer it we feel the need to demonstrate to the students the relevance of studying Linear Algebra.

An important reason to study Linear Algebra, as is referred in [4], is "Linear Algebra allows and even encourages a very satisfying combination of both elements of mathematics – abstraction and application." Also, Linear Algebra is essential in multiple areas of science in general. Indeed there are many references addressing elementary Linear Algebra applications. For example, in [3] the authors present 20 applications of linear algebra "drawn from business, economics, engineering, physics, computer science, approximation theory, ecology, sociology, demography, and genetics."

In this paper we describe three well known models and how they are worked in the classroom.

The choice of the models was made according to the required math concepts: matrix operations, methods to solve systems of linear equations, matrix inversion, eigenvalues and eigenvectors, matrix diagonalization which are all listed in syllabus of the curricular unit. Furthermore the chosen models: Leontief Economic model, Leslie population growth model and Rating Web pages model, are associated with phenomena of the scientific area of the study programmes we teach, that generally is Biology, Informatics, Economics or Managements.

Usually, in classes, after introducing the mathematical abstract concepts, we describe and explore the models with examples.

Leontief Matrix - Economic model

W. Leontief was awarded the Nobel Prize for his work in 1973. The model we present, based on his idea, is called input-output model.

In this model the economy of a country, region or industry, is divided into n sectors. Each of these sectors uses input from itself and other sectors to produce a product. In addition the sectors must satisfy an outside demand of their goods.

Suppose that there are n sectors, S_i, i=1,2,...,n producing goods or services, which are consumed, marketed or invested; each sector produces a unique and exclusive good, that is, there is a one to one relationship between goods and sectors; each sector produces the corresponding good through the consumption of goods in fixed proportions.

Let us denote by $A=[a_{ij}]$ the *technical coefficients matrix*, where a_{ij} is the value of the output of S_i needed to produce one unit of S_j and by d_i the output of S_i needed to satisfy the outside demand.

The purpose of this model is to find the total output, x_i, of each sector, S_i, i=1,2,...,n, in order to satisfy the intermediate and the outside demand, that is

$$x_i = a_{i1}.x_1 + a_{i2}.x_2 + \cdots + a_{in}.x_n + d_i \ , \text{ for } i = 1, 2, ..., n.$$

This system of equations can be written in matrix form:

$$\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} d_1 \\ \vdots \\ d_n \end{bmatrix} \Leftrightarrow X = AX + D$$

Thus, given A and D, the objective is to find X, production vector, that satisfies $X=AX+D$, with $x_i \geq 0$.

After describing the goal of the model, we explore with the students an example.

Example 1.1

Consider a simple economy with three sectors: manufacturing, agriculture and services. Suppose that:

i. the production of one unit of manufacturing requires 0,1 units of its own, 0,3 units of the agriculture sector and 0,3 units of the services sector.
ii. the production of one unit of agriculture requires 0,2 units of its own, 0,6 units of the manufacturing sector and 0,1 units of the services sector.
iii. the production of one unit of services requires 0,1 units of its own, 0,6 units of the manufacturing sector.

If there is an outside demand of 18 units of manufacturing, we are going to calculate the total output necessary to satisfy it.

The students when constructing the technical coefficients matrix

$$A = \begin{bmatrix} 0,1 & 0,6 & 0,6 \\ 0,3 & 0,2 & 0 \\ 0,3 & 0,1 & 0,1 \end{bmatrix}$$

recognize the importance of placing the matrix entries in the correct position.

To obtain the total output we have to solve the system *X*=*AX*+*D* with $D = \begin{bmatrix} 18 \\ 0 \\ 0 \end{bmatrix}$.

To solve this system of linear equations the students can use one of these methods: Gaussian Elimination, Cramer's rule and Matrix inversion. In the class the system is solved using the previous methods.

Using matrix inversion we have:
$X = AX + D \Leftrightarrow X - AX = D \Leftrightarrow (I - A)X = D \Leftrightarrow X = (I - A)^{-1}D,$ if *I-A* is invertible.

In this case $I - A = \begin{bmatrix} 0,9 & -0,6 & -0,6 \\ -0,3 & 0,8 & 0 \\ -0,3 & -0,1 & 0,9 \end{bmatrix}$, which inverse is $(I - A)^{-1} = \begin{bmatrix} \frac{20}{9} & \frac{50}{27} & \frac{40}{27} \\ \frac{5}{6} & \frac{35}{18} & \frac{5}{9} \\ \frac{5}{6} & \frac{5}{6} & \frac{5}{3} \end{bmatrix}$.

Will be obtained $X = (I - A)^{-1}D = \begin{bmatrix} 40 \\ 15 \\ 15 \end{bmatrix}$.

Thus the output of the manufacturing sector will be 40 units (4 for its own, 9 for agriculture, 9 for services and 18 for the outside demand); the output of the agriculture sector will be 15 units (12 for manufacturing and 3 for its own) and the output of the services sector will be 15 units (12 for manufacturing, 1,5 for agriculture and 1,5 for its own).

After we suppose that the outside demand is 60, 20 and 20 units of manufacturing, agriculture and services, respectively.

Then students conclude that, to obtain the new outputs, it is enough to multiply $(I-A)^{-1}$ by the new *D*, since we already have calculated $(I-A)^{-1}$.

So the total output *X* required will be $\begin{bmatrix} \frac{20}{9} & \frac{50}{27} & \frac{40}{27} \\ \frac{5}{6} & \frac{35}{18} & \frac{5}{9} \\ \frac{5}{6} & \frac{5}{6} & \frac{5}{3} \end{bmatrix} \begin{bmatrix} 60 \\ 20 \\ 20 \end{bmatrix} = \begin{bmatrix} 200 \\ 100 \\ 100 \end{bmatrix}.$

Naturally some questions arise:

- is *I-A* always invertible?
- is the production vector *X* nonnegative?

In order to answer this question we present the next result.

Theorem 1.1 *If A is a square matrix with nonnegative entries whose entries in each column sum less than one then* $(I-A)^{-1}$ exists and $(I-A)^{-1} \geq 0$.

Proof: Let λ be the largest column sum in *A*. By hypothesis $0 < \lambda < 1$. Therefore each column sum is less than or equal to λ.

We need to show that $(I-A)^{-1}$ exists and $(I-A)^{-1} \geq 0$. To do so we are going to prove that the inverse of $I-A$ is $I+A+A^2+\cdots$.

The j^{th} column of A^2 is a linear combination of the columns of A with coefficients the entries of the j^{th} column of *A*. Therefore, since each column sum in *A* is less than or equal to λ, we conclude that each column sum in A^2 is less than or equal to λ^2. Similarly, each column sum in A^m is less than or equal to $\lambda^m, m \in \mathbb{N}$. Hence, since A^m has nonnegative entries, we can conclude that every entry of A^m is less than or equal to λ^m.

Since $0 < \lambda < 1$ then $\lambda^m \to 0$ as $m \to \infty$ therefore $A^m \to 0$ (zero matrix).

In addition, we have $I - A^{m+1} = (I-A)(I+A+A^2+\cdots+A^m)$, consequently $I = (I-A)(I+A+A^2+\cdots+A^m)$ as $m \to \infty$ that is that $(I-A)^{-1}$ exists and furthermore $(I-A)^{-1} \geq 0$.

□

Being the technical coefficients matrix *A* a nonnegative matrix with each column sum less than one, then $(I-A)^{-1}$ exists, $(I-A)^{-1} \geq 0$ and the production vector $X \geq 0$ since $D \geq 0$.

Given a technical coefficients matrix *A*, whose entries in each column sum less than one, the matrix $(I-A)^{-1}$ is called the *Leontief inverse matrix*, which allows us to calculate the total outputs, *X*, of each sector for any outside demand *D*.

Remark: This proof suggests another way to calculate the approximate value of the production vector *X* doing $X \cong D + AD + A^2D + \cdots A^mD$ for m sufficiently large.

2. Leslie Matrix – Population Model

In the study of the growth of a given population it is essential having in account the survival and reproduction rates of the individuals of that population. It is known, however, that these characteristics differ according to the age of the individuals, their size body or any other individual characteristics influencing survival and fertility.

The Leslie matrix model was invented by P. H. Leslie and describes the growth of the female portion of a human or animal population. In this model, the population is divided into groups based on age classes of equal duration.

The purpose is to project the population from time *t* to time *t*+1, in years or some other time unit, assuming that the unit of time is the same as the age class width (*projection interval*). A model with projection interval of one year will differ from one that projects from month to month or decade to decade.

Supposing that the individuals of the population are classified into *k* age classes, the *population projection matrix*, often referred as a *Leslie matrix*, is:

$$L = \begin{bmatrix} R_1 & R_2 & R_3 & \dots & R_{k-1} & R_k \\ S_1 & 0 & 0 & \dots & 0 & 0 \\ 0 & S_2 & 0 & \dots & 0 & 0 \\ & \vdots & & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & S_{k-1} & S_k \end{bmatrix}$$

Where R_i is the reproduction rate and S_i is the survival rate, of age class *i*, for *i*=1,..., *k*.

In some models the last age class is assumed to be removed from the population after a time unit, so the entry, S_k , is 0 in the matrix *L*.

If the population at time *t*, distributed into the *k* classes is:

$$X^{(t)} = \begin{bmatrix} x_1^{(t)} \\ x_2^{(t)} \\ \vdots \\ x_k^{(t)} \end{bmatrix}$$ where $x_i^{(t)}$ is the population, at time *t*, in age class *i*=1,..., *k* ,

then the population at time *t*+1, in age class *i*=1,..., *k* will be:

$$x_1^{(t+1)} = R_1 x_1^{(t)} + R_2 x_2^{(t)} + \cdots + R_k x_k^{(t)} ,$$

that is, the population in age class one must have originated from reproduction, and not be survivors of any other age class, between times *t* and *t*+1.

$x_{i+1}^{(t+1)} = S_i\, x_i^{(t)}$, for *i*=1,..., *k*-1,

that is the population in age class *i*+1 will be the survivors of class *i*, between times *t* and *t*+1.

Using matrix notation the previous equations can be written as $X^{(t+1)} = LX^{(t)}$.
After describing the model, we explore with the students an example.

Example 2.1

Suppose that a female population has three age classes and the unit of time is five years. Assume that the reproduction rate in the second age class is 4 and in the third age class is 3. In addition suppose that the survival rate in the first class is 0,5 and 0,25 in the second class.

With this information the students can obtain the Leslie matrix for this population:

$$L = \begin{bmatrix} 0 & 4 & 3 \\ 0{,}5 & 0 & 0 \\ 0 & 0{,}25 & 0 \end{bmatrix}.$$

Assuming that the initial population is 10 in all age classes, when we ask the students to calculate the population in the next 15 years they do:

$$X^{(1)} = LX^{(0)} = \begin{bmatrix} 0 & 4 & 3 \\ 0{,}5 & 0 & 0 \\ 0 & 0{,}25 & 0 \end{bmatrix} \begin{bmatrix} 10 \\ 10 \\ 10 \end{bmatrix} = \begin{bmatrix} 70 \\ 5 \\ 2{,}5 \end{bmatrix}$$

$$X^{(2)} = LX^{(1)} = \begin{bmatrix} 0 & 4 & 3 \\ 0{,}5 & 0 & 0 \\ 0 & 0{,}25 & 0 \end{bmatrix} \begin{bmatrix} 70 \\ 5 \\ 2{,}5 \end{bmatrix} = \begin{bmatrix} 27{,}5 \\ 35 \\ 1{,}25 \end{bmatrix}$$

$$X^{(3)} = LX^{(2)} = \begin{bmatrix} 0 & 4 & 3 \\ 0{,}5 & 0 & 0 \\ 0 & 0{,}25 & 0 \end{bmatrix} \begin{bmatrix} 27{,}5 \\ 35 \\ 1{,}25 \end{bmatrix} = \begin{bmatrix} 143{,}75 \\ 13{,}75 \\ 8{,}75 \end{bmatrix}$$

Thus, since the unit of time is 5 years, after 15 years there are, nearly 143 females in age class one, 13 in age class two and 8 in age class three.

After we observe that since $X^{(n)} = LX^{(n-1)} = L\left(LX^{(n-2)}\right) = \cdots = L^n X^{(0)}$ we can calculate the population after *n* projection intervals knowing the initial population and powers of the Leslie matrix.

So, if we want to project the population after, for example 30 years, since the unit of time is 5 years, we can calculate $L^6 X^{(0)}$.

To calculate L^6 we propose to use Cayley-Hamilton Theorem [95] that states that any matrix satisfies its characteristic equation.

The characteristic equation of L is:

$$\det(L - \lambda I) = 0 \Leftrightarrow \begin{vmatrix} -\lambda & 4 & 3 \\ 0{,}5 & -\lambda & 0 \\ 0 & 0{,}25 & -\lambda \end{vmatrix} = 0 \Leftrightarrow -\lambda^3 + 2\lambda + 0{,}375 = 0.$$

By Cayley-Hamilton Theorem $-L^3 + 2L + 0{,}375\, I = 0$, where I is the identity matrix of the same size of L, that is, $L^3 = 2L + 0{,}375\, I$.

[95] See, for exemple, [2], page 86

We have $L^6 = 4L^2 + 1{,}5L + 0{,}140625\, I = \begin{bmatrix} 8{,}140625 & 9 & 4{,}5 \\ 0{,}75 & 8{,}140625 & 6 \\ 0{,}5 & 0{,}375 & 0{,}140625 \end{bmatrix}$.

Then $X^{(6)} = L^6 X^{(0)} = \begin{bmatrix} 216{,}40625 \\ 148{,}90625 \\ 10{,}15625 \end{bmatrix}$, that is after 30 years there are, nearly 216 females in age class one, 148 in age class two and 10 in age class three. Following we observe that since we know $X^{(3)}$ and $X^{(6)} = L^6 X^{(0)} = L^3(L^3 X^{(0)}) = L^3 X^{(3)}$, another way to obtain $X^{(6)}$ is:

$$X^{(6)} = (2L + 0{,}375\, I)X^{(3)} = \begin{bmatrix} 0{,}375 & 8 & 6 \\ 1 & 0{,}375 & 0 \\ 0 & 0{,}5 & 0{,}375 \end{bmatrix} \begin{bmatrix} 143{,}75 \\ 13{,}75 \\ 8{,}75 \end{bmatrix} = \begin{bmatrix} 216{,}40625 \\ 148{,}90625 \\ 10{,}15625 \end{bmatrix}.$$

After we suppose we want to project the population after a long period of time

To do this we introduce the following result.

Given a square matrix *A* we define the positive real number $\rho(A) = max\{|\lambda|: \lambda \in \sigma(A)\}$ where $\sigma(A)$ is the set of the eigenvalues of *A*.

Theorem 2.1 *If A is a nonnegative square matrix, then* $\rho(A)$ *is an eigenvalue of A, often called the dominant eigenvalue of A, and there is a positive vector X such that AX=* $\rho(A)X$.

The proof can be found in [2] page 503.

Since a Leslie matrix, *L*, is always nonnegative then $\rho(L)$ is the dominant eigenvalue of *L*, and exists a positive vector *V* such that *LV*= $\rho(L)V$.

Assuming that *L* is diagonalizable, then exists *P* and *D* such that $L = PDP^{-1}$, where *D* is a diagonal matrix whose entries are the eigenvalues of *L* and *P* is a matrix whose columns are the eigenvectors, therefore $L^n = (PDP^{-1})^n = PD^nP^{-1}$.

Suppose *L* is a Leslie matrix, 3×3, with eigenvalues $\lambda_1 = \rho(L)$, λ_2 ,λ_3 and the corresponding eigenvectors v_1, v_2 and v_3.

Then $X^{(n)} = L^n X^{(0)} =$

$$[v_1 \quad v_2 \quad v_3] \begin{bmatrix} {\lambda_1}^n & 0 & 0 \\ 0 & {\lambda_2}^n & 0 \\ 0 & 0 & {\lambda_3}^n \end{bmatrix} [v_1 \quad v_2 \quad v_3]^{-1} \begin{bmatrix} x_1^{(0)} \\ x_2^{(0)} \\ x_3^{(0)} \end{bmatrix}.$$

Dividing by $(\lambda_1)^n$ we have:

$$\frac{1}{{\lambda_1}^n}X^{(n)} = [v_1 \quad v_2 \quad v_3]\begin{bmatrix} 1 & 0 & 0 \\ 0 & \left(\frac{\lambda_2}{\lambda_1}\right)^n & 0 \\ 0 & 0 & \left(\frac{\lambda_3}{\lambda_1}\right)^n \end{bmatrix}[v_1 \quad v_2 \quad v_3]^{-1}\begin{bmatrix} x_1^{(0)} \\ x_2^{(0)} \\ x_3^{(0)} \end{bmatrix}$$

Since $\left|\frac{\lambda_2}{\lambda_1}\right|, \left|\frac{\lambda_3}{\lambda_1}\right| < 1$ then $lim_{n\to\infty}\left(\frac{\lambda_2}{\lambda_1}\right)^n = lim_{n\to\infty}\left(\frac{\lambda_3}{\lambda_1}\right)^n = 0$, therefore

$$lim_{n\to\infty}\frac{1}{{\lambda_1}^n}X^{(n)} = [v_1 \quad v_2 \quad v_3]\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}[v_1 \quad v_2 \quad v_3]^{-1}\begin{bmatrix} x_1^{(0)} \\ x_2^{(0)} \\ x_3^{(0)} \end{bmatrix} \Leftrightarrow$$

$$\Leftrightarrow lim_{n\to\infty}\frac{1}{{\lambda_1}^n}X^{(n)} = [v_1 \quad 0 \quad 0][v_1 \quad v_2 \quad v_3]^{-1}\begin{bmatrix} x_1^{(0)} \\ x_2^{(0)} \\ x_3^{(0)} \end{bmatrix} \Leftrightarrow$$

$$\Leftrightarrow lim_{n\to\infty}\frac{1}{{\lambda_1}^n}X^{(n)} = [v_1 \quad 0 \quad 0]\begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} \Leftrightarrow lim_{n\to\infty}\frac{1}{{\lambda_1}^n}X^{(n)} = c_1[v_1] .$$

Thus
$X^{(n)} \cong {\lambda_1}^n c_1[v_1]$.

Also $X^{(n)} \cong \lambda_1 \times {\lambda_1}^{n-1}c_1[v_1] \cong \lambda_1 X^{(n-1)}$, that is, for large values of time, the population is directly proportional to the preceding population, at each age class.

According to the value of λ_1, the dominant eigenvalue of the Leslie matrix, we can observe:

If $\lambda_1 > 1$ the population is eventually increasing.
If $\lambda_1 < 1$ the population is eventually decreasing.
If $\lambda_1 = 1$ the population eventually stabilizes.

Returning to our example, the eigenvalues of *L* are the solutions of the charact eristic equation:

$$-\lambda^3 + 2\lambda + 0{,}375 = 0 \Leftrightarrow \lambda = 1{,}5 \vee \lambda = \frac{-3 \pm \sqrt{5}}{4}$$

and the dominant eigenvalue of *L* is 1,5.

Since $\lambda_1 = 1{,}5 > 1$ the population is increasing, about 50%, in each of the three age classes, as well the total number of females.

In addition the eigenvectors are

$v_1 = \begin{bmatrix} 18 \\ 6 \\ 1 \end{bmatrix}$ $v_2 = \begin{bmatrix} 7+3\sqrt{5} \\ -3-\sqrt{5} \\ 1 \end{bmatrix}$ $v_3 = \begin{bmatrix} 7-3\sqrt{5} \\ -3+\sqrt{5} \\ 1 \end{bmatrix}$ and $[v_1 \quad v_2 \quad v_3]^{-1} = \begin{bmatrix} \frac{1}{38} & \frac{3}{38} & \frac{1}{19} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix}$.[96]

Since $\begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} \frac{1}{38} & \frac{3}{38} & \frac{1}{19} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix} \begin{bmatrix} 10 \\ 10 \\ 10 \end{bmatrix}$, therefore $c_1 = \frac{30}{19}$.

So $X^{(n)} \cong 1{,}5^n \frac{30}{19} \begin{bmatrix} 18 \\ 6 \\ 1 \end{bmatrix}$.

If, for example, n=20 we have $X^{(20)} \cong \begin{bmatrix} 94500 \\ 31500 \\ 5250 \end{bmatrix}$, that is, after 100 years there are, nearly, 94500 females in age class one, 31500 in age class two and 5250 in age class three.

Furthermore since $X^{(n)} \cong k \begin{bmatrix} 18 \\ 6 \\ 1 \end{bmatrix}$, $k \in \mathbb{R}$, for a long period of time we have:

$\frac{18}{25} = 72\%$ of the females are in age class one

$\frac{6}{25} = 24\%$ of the females are in age class two

$\frac{1}{25} = 4\%$ of the females are in age class three

3. Rating Web Pages Model

In order to measure the relative importance of web pages, Sergey Brin and Larry Page proposed in 1998 a method, *PageRank*, for computing a ranking for every web page based on the graph of the web. To test the utility of PageRank, they built a web search engine called Google.

The purpose of this model is to calculate the importance score (PageRank) of each web page. So the higher is the PageRank of one page, the higher is its chance of being found on Google.

Suppose we have a *strongly connected web* (that is, you can get from any page to any other page in a finite number of links) with n pages, P_i , i=1,…,n.

Denote by x_i, the importance score of page P_i , i=1,…,n, in the web.

[96] We only present the first row of the inverse of the eigenvectors matrix, since the remaining entries are not used in calculations.

If page P_j contains n_j links, one of which links to page P_k , then page P_j contributes for the importance score of page P_k with $x_j \frac{1}{n_j}$.

Therefore $x_k = \sum_{P_j \in L_k} \frac{x_j}{n_j}$ where L_k is the set of the pages which links to page P_k, k=1,…, n.

Using matrix notation the previous equations can be written as X=AX, where $A = [a_{ij}]$ with

$$a_{ij} = \begin{cases} \frac{1}{n_j}, & if \ P_j \text{ links to } P_i \\ 0, & \text{otherwise} \end{cases} \quad , \text{for } i,j=1,\ldots,n.$$

This matrix is called the *link matrix* for this web.

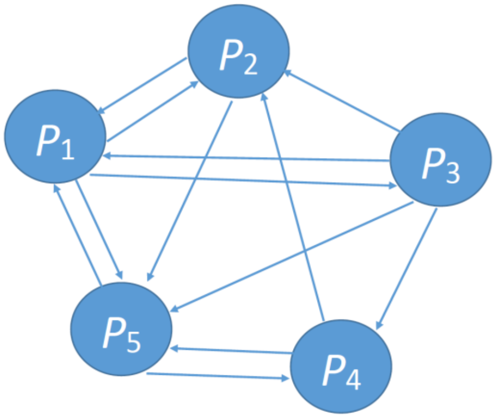
Notice that the j^{th} column of A has n_j non-zero entries, each equal to $\frac{1}{n_j}$, thus, the entries in each column of A sum to 1. Furthermore, since all its entries are nonnegative, A is a *column-stochastic matrix*, and therefore has 1 as an eigenvalue. In fact if we consider the column vector v with all its entries equal to one, we have $A^T v = v$, that is 1 is an eigenvalue of A^T and therefore also an eigenvalue of A. Also if we denote by $V_1(A)$ the eigenspace for eigenvalue 1 of A then dim $V_1(A)$=1, since the web is strongly connected.[1]

Then we are sure that the system, X=AX , has non-zero solutions, the eigenvectors associated to the eigenvalue 1, producing a unique ranking.
Then we ask the students to represent in a directed graph a web with five pages (the vertices of the graph) and where an arrow from page P_i to page P_j, i=1,…, 5, indicates a link from page P_i to page P_j. [97]

Example 3.1

Suppose, for example, we have a web with five pages P_i, i=1,…,5, illustrated in the graph:

[97] The students have some knowledge of graph theory obtained in another curricular unit.



In this example we have $n_1 = 3$, $n_2 = 2$, $n_3 = 4$, $n_4 = 2$ and $n_5 = 2$ (in graph language n_i is the out-degree of vertex P_i).

In first place we verify that the web (graph) is strongly connected, that is for every pair (P_i , P_j) of vertices there is a path from P_i to P_j .

To calculate x_i, i=1,…,5, the importance score of page P_i in this web, we have to solve the system of linear equations.

$$\begin{cases} x_1 = \frac{1}{2}x_2 + \frac{1}{4}x_3 + \frac{1}{2}x_5 \\ x_2 = \frac{1}{3}x_1 + \frac{1}{4}x_3 + \frac{1}{2}x_4 \\ x_3 = \frac{1}{3}x_1 \\ x_4 = \frac{1}{4}x_3 + \frac{1}{2}x_5 \\ x_5 = \frac{1}{3}x_1 + \frac{1}{2}x_2 + \frac{1}{4}x_3 + \frac{1}{2}x_4 \end{cases} \Leftrightarrow \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{4} & 0 & \frac{1}{2} \\ \frac{1}{3} & 0 & \frac{1}{4} & \frac{1}{2} & 0 \\ \frac{1}{3} & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{4} & 0 & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{2} & \frac{1}{4} & \frac{1}{2} & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix}$$

We already know that the system has more than one solution, so we can't use Cramer's Rule neither matrix inversion. The solutions of the previous system are

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = k \begin{bmatrix} \frac{10}{11} \\ \frac{2}{3} \\ \frac{10}{33} \\ \frac{19}{33} \\ 1 \end{bmatrix}, k \in \mathbb{R}.$$

Since the eigenvectors are just scalar multiples of each other, we can choose any of them to be our PageRank vector.

If we choose $k = \frac{33}{114}$ we obtain $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} \cong \begin{bmatrix} 0,26 \\ 0,19 \\ 0,09 \\ 0,17 \\ 0,29 \end{bmatrix}$, the unique eigenvector with the sum of all entries equal to 1.

In this web the ranking is: P_5 (29%), P_1 (26%), P_2 (19%),P_4 (17%), P_3(9%) .

Remark

If the web is not strongly connected, that is, there is, at least, one page P_k such that, there is no other page linking to P_k , the link matrix has, at least, one null line, and therefore yields non unique rankings.

If the web has *dangling nodes*, that is, there is, at least, one page with no links, then the link matrix has one or more columns of zeros.

In order to solve these problems, the model described is adapted, using, instead of A the matrix $M = (1 - p)A + pB$, with $0 < p < 1$ (usually $p = 0{,}15$), A the link matrix and B the matrix whose entries are all equal to $\frac{1}{n}$. [1]

Conclusion:

From our experience as mathematics teachers in Economics, Natural Sciences and Computer Science programs, we noticed that the students understand the mathematical concepts through the exploration of mathematical models, illustrated with examples of the respective study areas. In particular the students:

- recognize the importance of placing the matrix entries in the correct position when constructing the matrix models of the examples ;
- understand the meaning of matrix operations;
- identify properties of a matrix just “looking” at its entries;
- understand that eigenvalues are not only roots of characteristic equation;
- admit the importance of knowing different ways of solving linear equation systems;
- recognize the usefulness of matrix inversion;
- realize the potential of matrix diagonalization;
- apply matrix theory and graph theory creating connections between two curricular units,

combining abstraction and application.

References:

Bryan, K. and Leise, T. The $25,000,000,000 Eigenvector: The Linear Algebra Behind Google. SIAM Review 48, 3, 569-581, 2006

Horn, Roger and Johnson, Charles. Matrix Analysis. Cambridge University Press, 1985

Howard Anton, Chris Rorres. Elementary Linear Algebra: Applications Version. John Wiley & Sons, 1994

Strang, Gilbert. Linear Algebra and its Applications (second edition). Academic Press, 1980