Chapter 6

Matrix Diagonalization

6.1 Diagonalization Process

Most calculations we can perform with square matrices are simplified enormously if the matrices we work with are diagonal. Thus, for example, we have:

1. $$\begin{pmatrix} a_1 & & & \\ & a_2 & & \\ & & \ddots & \\ & & & a_n \end{pmatrix} \cdot \begin{pmatrix} b_1 & & & \\ & b_2 & & \\ & & \ddots & \\ & & & b_n \end{pmatrix} = \begin{pmatrix} a_1 \cdot b_1 & & & \\ & a_2 \cdot b_2 & & \\ & & \ddots & \\ & & & a_n \cdot b_n \end{pmatrix}.$$

2. $$\left| \begin{pmatrix} a_1 & & & \\ & a_2 & & \\ & & \ddots & \\ & & & a_n \end{pmatrix} \right| = a_1 \cdot a_2 \cdot \cdots \cdot a_n.$$

3. $$\begin{pmatrix} a_1 & & & \\ & a_2 & & \\ & & \ddots & \\ & & & a_n \end{pmatrix}^{-1} = \begin{pmatrix} a_1^{-1} & & & \\ & a_2^{-1} & & \\ & & \ddots & \\ & & & a^{-1} \end{pmatrix}.$$

4. $$\begin{pmatrix} a_1 & & & \\ & a_2 & & \\ & & \ddots & \\ & & & a_n \end{pmatrix}^{k} = \begin{pmatrix} a_1^k & & & \\ & a_2^k & & \\ & & \ddots & \\ & & & a_n^k \end{pmatrix}.$$

It will therefore be interesting to have methods that allow obtaining diagonal forms for any matrix. Of special importance is the calculation of powers of a square matrix. We have already seen examples of how these powers intervene in iterative phenomena that depend on a transition matrix. In these cases, the concepts we are about to introduce will also allow explaining different aspects of this type of phenomena related to the existence of certain stability states. Finally, we will introduce techniques that allow us to study the trend for future iterations in models with matrix powers.

First, we will establish what we understand by diagonalizing a matrix. Since, given any matrix A, we intend to perform on it the matrix operations we have listed, this diagonalization process must allow to transform A into a diagonal matrix D,

$$\underbrace{A}_{\text{Any matrix}} \xrightarrow[\text{diagonalization}]{} \underbrace{D}_{\text{Diagonal matrix}},$$

in such a way that we can recover for A the operations that we perform more easily on D. Let's see then the definition we give for diagonalization.

Definition 173. Given $A \in \mathcal{M}_n$, we say it is a diagonalizable matrix if there exists $C \in \mathcal{M}_n$ invertible such that the matrix

$$D = C^{-1} \cdot A \cdot C$$

is a diagonal matrix. In such a case, we say that the matrix C diagonalizes the matrix A and we call it the change-of-basis matrix.

We must now see that this definition indeed allows us to perform operations on the diagonalization D that we can later take advantage of for the matrix A. First, for the determinant we have that

$$|D| = |C^{-1} \cdot A \cdot C| = |C^{-1}| \cdot |A| \cdot |C| = |C|^{-1} \cdot |A| \cdot |C| = |A|$$

and therefore the determinants of D and A coincide. Moreover

$$D = C^{-1} \cdot A \cdot C \Rightarrow A = C \cdot D \cdot C^{-1}$$

and then

$$\begin{array}{ll} A^2 = & (C \cdot D \cdot C^{-1}) \cdot (C \cdot D \cdot C^{-1}) = C \cdot D^2 \cdot C^{-1}, \\ A^3 = & A^2 \cdot A = (C \cdot D^2 \cdot C^{-1}) \cdot (C \cdot D \cdot C^{-1}) = C \cdot D^3 \cdot C^{-1}, \\ A^4 = & A^3 \cdot A = (C \cdot D^3 \cdot C^{-1}) \cdot (C \cdot D \cdot C^{-1}) = C \cdot D^4 \cdot C^{-1} \end{array}$$

being able to repeat this process indefinitely. In summary, we have thus demonstrated the following:

Properties 174. *Let $A \in \mathcal{M}_n$ be a diagonalizable matrix such that*

$$D = C^{-1} \cdot A \cdot C,$$

where $C, D \in \mathcal{M}_n$, with D being a diagonal matrix and C an invertible matrix. Then:

i) $|A| = |D|$.

ii) $A^n = C \cdot D^n \cdot C^{-1}$, $n \in \mathbb{N}$. *In particular, if A is invertible this property is also valid for $n \in \mathbb{Z}$, $n < 0$.*

Given a matrix $A \in \mathcal{M}_{n \times n}$, to diagonalize it we must find the change-of-basis matrix C and the diagonalization D. For this, we will make the following considerations:

- We will assume that the column vectors of the change-of-basis matrix C are $v_1, \ldots, v_n \in \mathbb{R}^n$, that is,

  $$C = (v_1 | v_2 | \ldots | v_n),$$

  and that

  $$D = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}.$$

- The matrix C must have an inverse and therefore

  $$\det(C) \neq 0 \Leftrightarrow \{v_1, v_2, \ldots, v_n\} \text{ are all of them independent.}$$

- It is easy to verify that

  $$A \cdot C = A \cdot (v_1 | v_2 | \ldots | v_n) = (Av_1 | Av_2 | \ldots | Av_n)$$

  and also that

  $$C \cdot D = (\lambda_1 v_1 | \lambda_2 v_2 | \ldots | \lambda_n v_n).$$

- If the matrix C diagonalizes A with D being the diagonalization then, taking into account the previous point,

$$D = C^{-1} \cdot A \cdot C \Leftrightarrow A \cdot C = C \cdot D$$

$$\Leftrightarrow (Av_1|Av_2|\dots|Av_n) = (\lambda_1 v_1|\lambda_2 v_2|\dots|\lambda_n v_n)$$

$$\Leftrightarrow \begin{cases} Av_1 &= \lambda_1 v_1 \\ Av_2 &= \lambda_2 v_2 \\ & \vdots \\ Av_n &= \lambda_n v_n \end{cases}$$

Therefore, if we find a basis of vectors of $\mathbb{R}^n$, $\{v_1, v_2, \dots, v_n\}$, satisfying this last property then the matrix A is diagonalizable being $C = (v_1|v_2|\dots|v_n)$ and

$$D = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}.$$

Definition 175. Given $A \in \mathcal{M}_n$ we call:

- **eigenvalue of** A any real number $\lambda \in \mathbb{R}$ such that there exists some non-zero vector, $v \in \mathbb{R}^n$, such that

$$A \cdot v = \lambda v.$$

- **eigenvector of** A **associated with the eigenvalue** λ any vector $v \in \mathbb{R}^n$ such that

$$A \cdot v = \lambda v.$$

- **eigenspace of** A **associated with the eigenvalue** λ the set of all eigenvectors of A associated with the eigenvalue λ,

$$V_\lambda = \{v \in \mathbb{R}^n / A \cdot v = \lambda v\}.$$

  Such set V_λ is a vector subspace of $\mathbb{R}^n$.

The concepts of eigenvector and eigenvalue have important interpretations in different iterative matrix models where they determine the stability positions of a system.

To check if λ is an eigenvalue of A we must find a non-zero vector, $v \in \mathbb{R}^n$, such that $A \cdot v = \lambda v$. Now,

$$A \cdot v = \lambda v \Leftrightarrow A \cdot v - \lambda v = 0 \Leftrightarrow A \cdot v - \lambda I_n \cdot v = 0$$

$$\Leftrightarrow (A - \lambda I_n) \cdot v = 0,$$

so, equivalently, the vector v must verify the equation $(A - \lambda I_n) \cdot v = 0$. Therefore, if we call $\overline{A} = A - \lambda I_n$, what we must find is a non-zero vector $v = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$ such that

$$\overline{A} \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

This last expression constitutes a homogeneous system of n equations with n unknowns expressed in matrix form with coefficient matrix $\overline{A} = A - \lambda I_n$. Homogeneous systems always have at least one solution, the zero solution $(0, 0, \dots, 0) \in \mathbb{R}^n$, and what we intend is to find a non-zero solution of such homogeneous system. For this, it will be necessary that the system has some solution different from the zero solution and therefore

has more than one solution, that is, it is an indeterminate system, which depends on whether its coefficient matrix, $\overline{A}$, is regular or not. For all this we have

$$\lambda \text{ is an eigenvalue of } A \Leftrightarrow \exists\ v \in \mathbb{R}^n, v \neq 0, \text{ such that } A \cdot v = \lambda v$$

$$\Leftrightarrow \overline{A} \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = 0 \text{ is indeterminate} \Leftrightarrow |\overline{A}| = 0$$

$$\Leftrightarrow |A - \lambda I_n| = 0.$$

Let V_λ be the set of all eigenvectors of A associated with the eigenvalue λ. From all the previous reasoning, it follows that

$$\begin{aligned} V_\lambda &= \{v \in \mathbb{R}^n / A \cdot v = \lambda v\} \\ &= \{v = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n / (A - \lambda I_n) \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}\} \end{aligned}$$

so V_λ is a vector subspace with implicit equations given in matrix form by

$$V_\lambda \equiv (A - \lambda I_n) \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

This is the basic technique for calculating eigenvalues and eigenvectors. In the following property, we formulate it precisely along with other results of interest for manipulating eigenvalues and eigenvectors.

Property 176. *Given $A \in \mathcal{M}_n$:*

i) It holds that

$$\lambda \in \mathbb{R} \text{ is an eigenvalue of } A \Leftrightarrow |A - \lambda I_n| = 0$$

and, if $\lambda \in \mathbb{R}$ is an eigenvalue, then V_λ is the vector subspace of $\mathbb{R}^n$ given by

$$V_\lambda \equiv (A - \lambda I_n) \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

and therefore

$$\dim(V_\lambda) = n - \text{rango}(A - \lambda I_n).$$

ii) Suppose that $\lambda_1, \lambda_2, \ldots, \lambda_k \in \mathbb{R}$ are eigenvalues of A distinct from each other. Then, if B_1 is a basis of V_{λ_1}, B_2 is a basis of $V_{\lambda_2}, \ldots$, B_k is a basis of V_{λ_k}, we have that

$$H = B_1 \cup B_2 \cup \cdots \cup B_k$$

is an independent set.

iii) If $\lambda \in \mathbb{R}$ is an eigenvalue of the matrix A and $v \in \mathbb{R}^n$ is an eigenvector of A associated with λ then

$$A^k v = \lambda^k v.$$

Definition 177. Given $A \in \mathcal{M}_n$ we call characteristic polynomial of the matrix A the polynomial

$$p(\lambda) = |A - \lambda I_n| \in \mathbb{P}_n(\lambda)$$

and we call characteristic equation of the matrix A the equation

$$p(\lambda) = 0.$$

Remark. From all the above, it follows that:

- The eigenvalues of a matrix, $A \in \mathcal{M}_n$, are the solutions of its characteristic equation.
- A matrix can be diagonalized if we find a basis formed exclusively by eigenvectors.

We may encounter the following problems that would prevent a matrix from being diagonalizable:

1. The matrix either has no eigenvalues or has an insufficient number of them.
2. The eigenvectors of the matrix do not allow forming a basis.

In what follows, we aim to give a characterization of matrices that are diagonalizable.

Definition 178.

i) Given a polynomial, $p(\lambda) \in \mathbb{P}_n(\lambda)$, we say that $\lambda_0 \in \mathbb{R}$ is a zero of multiplicity k of $p(\lambda)$ if we can express $p(\lambda)$ in the form

$$p(\lambda) = q(\lambda) \cdot (\lambda - \lambda_0)^k,$$

where $q(\lambda) \in \mathbb{P}_{n-k}(\lambda)$ verifies that $q(\lambda_0) \neq 0$.

ii) Given $A \in \mathcal{M}_n$ and $\lambda \in \mathbb{R}$ an eigenvalue of A we say that the algebraic multiplicity of λ is k if λ is a zero of multiplicity k of the characteristic polynomial of the matrix A.

iii) Given $A \in \mathcal{M}_n$ and $\lambda \in \mathbb{R}$ an eigenvalue of A, we call geometric multiplicity of λ the dimension of the eigenspace associated with λ, V_λ, that is, $\dim(V_\lambda)$.

Property 179. *Let $A \in \mathcal{M}_n$ whose eigenvalues are $\lambda_1, \lambda_2, \ldots, \lambda_k$, such that $\forall i = 1, \ldots, k$*

$$\begin{cases} n_i \text{ is the algebraic multiplicity of } \lambda_i. \\ m_i \text{ is the geometric multiplicity of } \lambda_i. \end{cases}.$$

Then it holds that

1. $n_1 + n_2 + \cdots + n_k \leq n$.
2. $1 \leq m_i \leq n_i, \ \forall i = 1, \ldots, k$.
3. A *is diagonalizable* $\Leftrightarrow \begin{cases} n_1 + n_2 + \cdots + n_k = n. \\ m_i = n_i, \ \forall i = 1, \ldots, k. \end{cases}$.

We collect below some important properties. The first three parts of the property indicate that there exist certain types of matrices for which the existence of diagonalization is guaranteed without needing to perform any calculation.

Property 180.

i) If $A \in \mathcal{M}_n$ is a diagonal matrix,

$$A = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix},$$

then A is diagonalizable, and it holds that the matrix I_n diagonalizes the matrix A, the canonical basis B_c of $\mathbb{R}^n$ is a basis of eigenvectors of A, and its eigenvalues are $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{R}$ such that the number of times each eigenvalue is repeated indicates its algebraic and geometric multiplicity.

ii) Every symmetric matrix is diagonalizable.

iii) If $A \in \mathcal{M}_n$ has n eigenvalues, all of them distinct, then A is diagonalizable.

iv) If the columns or the rows of $A \in \mathcal{M}_n$ all sum to the same number $r \in \mathbb{R}$, then $\lambda = r$ is an eigenvalue of A.

v) The characteristic polynomial of $A \in \mathcal{M}_2$ is

$$p(\lambda) = \lambda^2 - trace(A)\lambda + |A|.$$

vi) The characteristic polynomial of $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \in \mathcal{M}_3$ *is*

$$\begin{aligned} p(\lambda) &= -\lambda^3 + trace(A)\lambda^2 \\ &\quad - \left(\left| \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \right| + \left| \begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix} \right| + \left| \begin{pmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{pmatrix} \right| \right) \lambda + |A|. \end{aligned}$$

6.2 Study of Trends in Iterative Processes

The product and power of matrices are fundamental in formulating the most important matrix models.

Suppose we are studying a phenomenon involving several quantities $a_1, a_2, \ldots, a_k$ that vary over time. If we study that phenomenon over several periods, $n = 0$ (initial period), $n = 1$, $n = 2$, etc., the quantities $a_1, a_2, \ldots, a_k$ will take different values. Therefore, we will have different values for them in each period, n: $a_{1,n}, a_{2,n}, \ldots, a_{k,n}$. If we arrange the value of the quantities in each period as a column vector, we have

$$P_n = \begin{pmatrix} a_{1,n} \\ a_{2,n} \\ \vdots \\ a_{k,n} \end{pmatrix}.$$

Thus we have a list of k-tuples, P_0, P_1, ..., P_n, which provide the information of the phenomenon in each period.

The important thing here would be to be able to calculate the k-tuples corresponding to future periods so that we can predict the evolution of the phenomenon. This is where the calculation of matrix powers comes into play, since in numerous situations, if we know the initial situation of the phenomenon, i.e., we know the tuple P_0 corresponding to the initial period $n = 0$, we can calculate the tuple for any period n using a formula of the type

$$P_n = A^n \cdot P_0,$$

where A is a square matrix of order k called the **transition matrix** which governs the changes the phenomenon undergoes from one period to the next.

Consequently, the fundamental elements of these models are:

- The model will describe the situation of a certain phenomenon in successive periods. We will know the initial values which we will collect in a vector P_0 and call P_1, P_2, P_3, in general P_k, the vectors corresponding to the following periods.
- We will have a transition matrix, A, which governs the changes from one period to the next according to the matrix equations

$$P_{k+1} = AP_k \qquad \text{and} \qquad P_k = A^k P_0.$$

Studying the trend involves determining the future behavior of a model of this type, which ultimately means calculating or studying in some way the value of

$$A^k P_0$$

for large values of k (this is equivalent to studying the limit $\lim_{k\to\infty} A^k P_0$). Diagonalization techniques provide us with a direct way to perform this calculation. In fact, **Property 174** shows how to calculate the power A^k once we have diagonalized the matrix A.

Example 181. Suppose that in a certain commercial sector three companies compete, which we will call A, B and C. From one year to the next, the customers of each of them decide to remain loyal or switch to one of the others. A study is conducted on the movements between the three companies and it is observed that year after year the customers show similar behavior determined by the data in the following table:

	Customers of A	**Customers of B**	**Customers of C**
Switch to being customers of A	80%	10%	10%
Switch to being customers of B	10%	60%	20%
Switch to being customers of C	10%	30%	70%

For example, we see that each year 80% of the customers of A remain loyal to A, 10% switch to B and 10% to C.

Suppose also that in the year the study began, company A had 210 customers, B had 190 and C, 320.

Let's set up a matrix model to study this problem. Assuming that year $k = 0$ is the year in which the study of the customers of the three companies began, we will call:

- A_k = number of customers of company A after k years.
- B_k = number of customers of company B after k years.
- C_k = number of customers of company C after k years.

The information for each year will be grouped in a column vector which we will denote as P_k,

$$P_k = \begin{pmatrix} A_k \\ B_k \\ C_k \end{pmatrix}.$$

According to the problem data we have that $A_0 = 210$, $B_0 = 190$ and $C_0 = 320$ so

$$P_0 = \begin{pmatrix} 210 \\ 190 \\ 320 \end{pmatrix}.$$

Applying the transition table it is easy to calculate the customers there will be in each company if we know those of the previous year. Thus, if in year k we have A_k in A, B_k in B and C_k in C, in year $k+1$ we will have:

- $\underbrace{A_{k+1}}_{\text{customers in A in year } k+1} = 80\%$ of A_k+10% of B_k+10% of $C_k = 0.8A_k + 0.1B_k + 0.1C_k$.

- $\underbrace{B_{k+1}}_{\text{customers in B in year } k+1} = 10\%$ of A_k+60% of B_k+20% of $C_k = 0.1A_k + 0.6B_k + 0.2C_k$.

- $\underbrace{C_{k+1}}_{\text{customers in C in year } k+1} = 10\%$ of A_k+30% of B_k+70% of $C_k = 0.1A_k + 0.3B_k + 0.7C_k$

Writing all this information in column and using the definition of matrix product, it is easy to realize that

$$P_{k+1} = \begin{pmatrix} 0.8A_k + 0.1B_k + 0.1C_k \\ 0.1A_k + 0.6B_k + 0.2C_k \\ 0.1A_k + 0.3B_k + 0.7C_k \end{pmatrix} = \begin{pmatrix} 0.8 & 0.1 & 0.1 \\ 0.1 & 0.6 & 0.2 \\ 0.1 & 0.3 & 0.7 \end{pmatrix} \cdot \begin{pmatrix} A_k \\ B_k \\ C_k \end{pmatrix} = \begin{pmatrix} 0.8 & 0.1 & 0.1 \\ 0.1 & 0.6 & 0.2 \\ 0.1 & 0.3 & 0.7 \end{pmatrix} \cdot P_k.$$

Denoting $A = \begin{pmatrix} 0.8 & 0.1 & 0.1 \\ 0.1 & 0.6 & 0.2 \\ 0.1 & 0.3 & 0.7 \end{pmatrix}$, ultimately we have proved that

$$P_{k+1} = AP_k.$$

We therefore have,

$$\begin{aligned} P_1 &= AP_0 \\ P_2 &= AP_1 \\ P_3 &= AP_2 \\ P_4 &= AP_3 \\ &\text{etc.} \end{aligned}$$

Then, if we want to calculate P_4 according to this scheme, since the only data we know are those of the initial year, i.e., P_0, we would have to first calculate P_1, then P_2 and P_3 and finally P_4. However, we have

$$\begin{aligned} P_2 &= AP_1 = A(AP_0) = (AA)P_0 = A^2P_0. \\ P_3 &= AP_2 = (\text{using the previous equation}) = A(A^2P_0) = (AA^2)P_0 = A^3P_0. \\ P_4 &= AP_3 = (\text{using the previous equation}) = A(A^3P_0) = (AA^3)P_0 = A^4P_0. \end{aligned}$$

Therefore using matrix powers we can calculate P_4 without needing to first obtain P_0, P_1, P_2 and P_3. Actually, it is clear that this process can be applied iteratively as many times as we want so, in general,

$$\boxed{P_k = A^kP_0}\,. \tag{6.1}$$

What we see here is that the distribution of customers in year k, P_k, is determined by the initial distribution, P_0, and the k-th power of A. The matrix A regulates the step from one year to the next and is the **transition matrix** for this problem.

Since we know the initial distribution of customers, P_0, we can easily calculate the distribution in successive years. To do this, we compute several powers of A:

$$\begin{aligned} A^2 = AA &= \begin{pmatrix} 0.8 & 0.1 & 0.1 \\ 0.1 & 0.6 & 0.2 \\ 0.1 & 0.3 & 0.7 \end{pmatrix} \begin{pmatrix} 0.8 & 0.1 & 0.1 \\ 0.1 & 0.6 & 0.2 \\ 0.1 & 0.3 & 0.7 \end{pmatrix} = \begin{pmatrix} 0.66 & 0.17 & 0.17 \\ 0.16 & 0.43 & 0.27 \\ 0.18 & 0.4 & 0.56 \end{pmatrix}. \\ A^3 = AA^2 &= \begin{pmatrix} 0.8 & 0.1 & 0.1 \\ 0.1 & 0.6 & 0.2 \\ 0.1 & 0.3 & 0.7 \end{pmatrix} \begin{pmatrix} 0.66 & 0.17 & 0.17 \\ 0.16 & 0.43 & 0.27 \\ 0.18 & 0.4 & 0.56 \end{pmatrix} = \begin{pmatrix} 0.562 & 0.219 & 0.219 \\ 0.198 & 0.355 & 0.291 \\ 0.24 & 0.426 & 0.49 \end{pmatrix}. \\ A^4 = AA^3 &= \begin{pmatrix} 0.8 & 0.1 & 0.1 \\ 0.1 & 0.6 & 0.2 \\ 0.1 & 0.3 & 0.7 \end{pmatrix} \begin{pmatrix} 0.562 & 0.219 & 0.219 \\ 0.198 & 0.355 & 0.291 \\ 0.24 & 0.426 & 0.49 \end{pmatrix} = \begin{pmatrix} 0.4934 & 0.2533 & 0.2533 \\ 0.223 & 0.3201 & 0.2945 \\ 0.2836 & 0.4266 & 0.4522 \end{pmatrix}. \end{aligned}$$

Using these calculations with equation (6.1) we have that

$$\begin{aligned} P_1 = AP_0 &= \begin{pmatrix} 0.8 & 0.1 & 0.1 \\ 0.1 & 0.6 & 0.2 \\ 0.1 & 0.3 & 0.7 \end{pmatrix} \begin{pmatrix} 210 \\ 190 \\ 320 \end{pmatrix} = \begin{pmatrix} 219 \\ 199 \\ 302 \end{pmatrix}. \\ P_2 = A^2P_0 &= \begin{pmatrix} 0.66 & 0.17 & 0.17 \\ 0.16 & 0.43 & 0.27 \\ 0.18 & 0.4 & 0.56 \end{pmatrix} \begin{pmatrix} 210 \\ 190 \\ 320 \end{pmatrix} = \begin{pmatrix} 225.3 \\ 201.7 \\ 293 \end{pmatrix}. \\ P_3 = A^3P_0 &= \begin{pmatrix} 0.562 & 0.219 & 0.219 \\ 0.198 & 0.355 & 0.291 \\ 0.24 & 0.426 & 0.49 \end{pmatrix} \begin{pmatrix} 210 \\ 190 \\ 320 \end{pmatrix} = \begin{pmatrix} 229.71 \\ 202.15 \\ 288.14 \end{pmatrix}. \\ P_4 = A^4P_0 &= \begin{pmatrix} 0.4934 & 0.2533 & 0.2533 \\ 0.223 & 0.3201 & 0.2945 \\ 0.2836 & 0.4266 & 0.4522 \end{pmatrix} \begin{pmatrix} 210 \\ 190 \\ 320 \end{pmatrix} = \begin{pmatrix} 232.797 \\ 201.889 \\ 285.314 \end{pmatrix}. \end{aligned}$$

After performing all these operations, it is clear that the major obstacle is the calculation of the powers of A because it is not a diagonal matrix. Up to the fourth power the calculation could be done manually, but if we wanted higher powers like A^{20} or A^{30}, it seems essential to resort to other techniques. On the other hand, once this model is formulated, several questions arise to be solved:

a) Is it possible to study the future trend in the distribution of customers? Even more interesting than calculating the customers in a specific year would be to be able to describe future behavior by determining whether customers tend to choose one of the three companies with higher priority or if, on the contrary, they distribute homogeneously among them.

b) Do equilibrium distributions exist? For example, the presidents of the three companies could try to agree to distribute the customer market in such a way that it remains constant from one year to the next. For this, we should choose an initial distribution of customers

$$P_0 = \begin{pmatrix} A_0 \\ B_0 \\ C_0 \end{pmatrix}$$

such that the distributions in following years, P_1, P_2, P_3, etc., are always equal. If the customer distributions in year zero and year one are equal we will have $P_0 = P_1$ and since we know that $P_1 = AP_0$ we deduce that

$$\boxed{AP_0 = P_0}\,.$$

If we find an initial distribution, P_0, that satisfies this last condition, it is not difficult to check that the distribution in all subsequent years is always the same since

$$\begin{aligned} P_1 &= AP_0 = P_0, \\ P_2 &= AP_1 = AP_0 = P_0, \\ P_3 &= AP_2 = AP_0 = P_0, \\ P_4 &= AP_3 = AP_0 = P_0, \\ &\text{etc.} \end{aligned}$$

For example, if $A_0 = 600$, $B_0 = 500$, $C_0 = 700$ then $P_0 = (600, 500, 700)$ and it is easy to check that

$$AP_0 = A \begin{pmatrix} 600 \\ 500 \\ 700 \end{pmatrix} = \begin{pmatrix} 600 \\ 700 \\ 500 \end{pmatrix} = P_0.$$

Therefore, since the condition $AP_0 = P_0$ holds, in all following years we will always have the same customer distribution given by the 3-tuple $(600, 500, 700)$.

c) If the total number of customers is increasing or decreasing, it will be impossible for the number of customers of the three companies in successive years to remain constant. In such a case, the presidents could agree that at least the percentages of customers for each company are the same in all years. It is simple to prove that the distributions of customers in the initial year, P_0, and in year one, P_1, represent the same percentages if we can find $\lambda \neq 0$ such that $P_1 = \lambda P_0$. Since $P_1 = AP_0$ we conclude that

$$\boxed{AP_0 = \lambda P_0}\,.$$

In this case, we ask about how to calculate the value λ and the initial distribution P_0.

In fact, the answers to the questions posed at the end of the previous example are the eigenvalues and eigenvectors of the matrix. Let's see next how to calculate them.

Example 182. Let us calculate all eigenvalues and eigenvectors of the matrix

$$A = \begin{pmatrix} 0.8 & 0.1 & 0.1 \\ 0.1 & 0.6 & 0.2 \\ 0.1 & 0.3 & 0.7 \end{pmatrix}.$$

We start by calculating the characteristic polynomial:

$$\begin{aligned} |A - \lambda I_3| &= \left| \begin{pmatrix} 0.8 & 0.1 & 0.1 \\ 0.1 & 0.6 & 0.2 \\ 0.1 & 0.3 & 0.7 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right| = \left| \begin{pmatrix} 0.8-\lambda & 0.1 & 0.1 \\ 0.1 & 0.6-\lambda & 0.2 \\ 0.1 & 0.3 & 0.7-\lambda \end{pmatrix} \right| \\ &= (0.8-\lambda)(0.6-\lambda)(0.7-\lambda) + 0.1 \cdot 0.3 \cdot 0.1 + 0.1 \cdot 0.2 \cdot 0.1 \\ &\quad -\big(0.1(0.6-\lambda)0.1 + 0.3 \cdot 0.2(0.8-\lambda) + 0.1 \cdot 0.1(0.7-\lambda)\big) \\ &= -\lambda^3 + 2.1\lambda^2 - 1.38\lambda + 0.28. \end{aligned}$$

This last expression is the characteristic polynomial of the matrix A (indeed, it is a polynomial of degree three in the variable λ). The characteristic equation of A is

$$-\lambda^3 + 2.1\lambda^2 - 1.38\lambda + 0.28 = 0.$$

Let us solve the characteristic equation. If we consider that we already know that $\lambda = 1$ is an eigenvalue, we know in advance that one of the solutions of the characteristic equation is $\lambda = 1$. If we then apply Ruffini's method for $\lambda = 1$,

$$\begin{array}{c|cccc} & -1 & 2.1 & -1.38 & 0.28 \\ 1 & & -1 & 1.1 & -0.28 \\ \hline & -1 & 1.1 & -0.28 & \underline{\vert 0} \end{array},$$

we confirm that indeed $\lambda = 1$ is a solution of the equation. From here it will be difficult to obtain the other solutions by applying Ruffini's method again. However, the coefficients we obtain in the last line of the previous Ruffini division (-1, 1.1 and -0.28) indicate that the equation left to solve is

$$-\lambda^2 + 1.1\lambda - 0.28 = 0$$

and this is a second-degree equation that we can solve directly by applying the corresponding formula to obtain

$$\lambda = \frac{-1.1 \pm \sqrt{1.1^2 - 4 \cdot (-1) \cdot (-0.28)}}{2 \cdot (-1)} \begin{cases} = 0.4 \\ = 0.7 \end{cases},$$

so finally, the three solutions of the characteristic equation are,

$$\begin{cases} \lambda = 1 \\ \lambda = 0.4 \\ \lambda = 0.7 \end{cases}.$$

In this way, we deduce that the only three eigenvalues of the matrix A are $\lambda = 1$, $\lambda = 0.4$ and $\lambda = 0.7$.

Let us calculate the eigenspaces corresponding to each of the three eigenvalues:

- The eigenspace associated with $\lambda = 1$ is the vector subspace with implicit equations

$$V_1 \equiv (A - 1I_3) \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Rightarrow V_1 \equiv \begin{pmatrix} -0.2 & 0.1 & 0.1 \\ 0.1 & -0.4 & 0.2 \\ 0.1 & 0.3 & -0.3 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

  Solving the system, it is easy to check that a basis for this subspace is $B_1 = \{(6, 5, 7)\}$.

- For $\lambda = 0.4$ the eigenspace is the vector subspace

$$V_{0.4} \equiv (A - 0.4I_3)\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Rightarrow V_{0.4} \equiv \begin{pmatrix} 0.4 & 0.1 & 0.1 \\ 0.1 & 0.2 & 0.2 \\ 0.1 & 0.3 & 0.3 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

  A basis for this subspace is $B_{0.4} = \{(0, -1, 1)\}$.

- For $\lambda = 0.7$ the eigenspace is the vector subspace

$$V_{0.7} \equiv (A - 0.7I_3)\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Rightarrow V_{0.7} \equiv \begin{pmatrix} 0.1 & 0.1 & 0.1 \\ 0.1 & -0.1 & 0.2 \\ 0.1 & 0.3 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

  A basis for this last subspace is $B_{0.7} = \{(-3, 1, 2)\}$.

Part *ii)* of **Property 176** guarantees that by combining the elements of B_1, $B_{0.4}$ and $B_{0.7}$ we obtain a set of independent vectors

$$B = \{(6, 5, 7), (0, -1, 1), (-3, 1, 2)\}.$$

Since three independent vectors in $\mathbb{R}^3$ form a basis, B is a basis formed by eigenvectors associated, in that order, with the eigenvalues $\lambda = 1$, $\lambda = 0.4$ and $\lambda = 0.7$. Therefore, the initial matrix, A, is diagonalizable with change-of-basis matrix C and diagonalization D given by

$$C = \begin{pmatrix} 6 & 0 & -3 \\ 5 & -1 & 1 \\ 7 & 1 & 2 \end{pmatrix}, \qquad D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0.4 & 0 \\ 0 & 0 & 0.7 \end{pmatrix}.$$

Once we know the eigenvalues and eigenvectors, there are calculations we can perform more easily.

Example 183. Part *iii)* of **Property 176** shows that a calculation of the type $A^k v$ is greatly simplified if v is an eigenvector.

If in the previous customer example, the initial data are $P_0 = (6, 5, 7)$, to calculate $A^k P_0$ since we know that $(6, 5, 7)$ is an eigenvector of A associated with the eigenvalue $\lambda = 1$, directly for any k we have

$$A^k P_0 = A^k \begin{pmatrix} 6 \\ 5 \\ 7 \end{pmatrix} = 1^k \begin{pmatrix} 6 \\ 5 \\ 7 \end{pmatrix}$$

and we do not need to perform the explicit calculation of the power A^k.

However, when we know the initial data tuple P_0 or when we want to study the trend for large values of k, the method we will see next called the 'power method' is more appropriate.

6.2.1 The Power Method

Suppose we want to perform the calculation

$$A^k P_0$$

for some matrix $A \in \mathcal{M}_n$, the initial data n-tuple P_0 and $k \in \mathbb{N}$. Assume that the matrix A is diagonalizable. Then, we can calculate for A a basis of eigenvectors $\{v_1, v_2, \ldots, v_n\}$ each of them associated respectively with the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ in the form

Eigenvector	Associated Eigenvalue
v_1	λ_1
v_2	λ_2
$\vdots$	$\vdots$
v_n	λ_n

Since the eigenvectors $v_1, v_2, \ldots, v_n$ form a basis of $\mathbb{R}^n$, any n-tuple can be obtained as a linear combination of them. In particular, the n-tuple P_0 can be written in the form

$$P_0 = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n$$

for certain coefficients $\alpha_1, \alpha_2, \ldots, \alpha_n \in \mathbb{R}$ that can be calculated by solving the corresponding system. Now, we can take advantage of the expression of P_0 as a linear combination of the eigenvectors $v_1, v_2, \ldots, v_n$ to perform the calculation $A^k P_0$. Indeed, we have that:

$$A^k P_0 = A^k \left(\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n\right) = \left(\begin{array}{c}\text{using the}\\ \text{distributive property}\\ \text{of matrix product}\end{array}\right) = A^k\alpha_1 v_1 + A^k\alpha_2 v_2 + \cdots + A^k\alpha_n v_n$$

$$= \alpha_1 \underline{A^k v_1} + \alpha_2 \underline{A^k v_2} + \cdots + \alpha_n \underline{A^k v_n}.$$

In this last expression, the calculations we have underlined remain to be done. But if we take into account that $v_1, v_2, \ldots, v_n$ are eigenvectors, it is possible to apply part *iii)* of **Property 176** to reach

$$A^k v_1 = \lambda_1^k v_1, \qquad A^k v_2 = \lambda_2^k v_2, \ldots \qquad A^k v_n = \lambda_n^k v_n$$

and in this way we have carried out the most complicated part of the calculation since after using the last equalities the matrix power A^k disappears and in its place we have the powers $\lambda_1^k, \lambda_2^k, \ldots, \lambda_n^k$ which are all simple powers of numbers (a number raised to a number and not a matrix raised to a number). In this way, gathering all the calculations we have that

$$A^k P_0 = \alpha_1 \underbrace{A^k v_1}_{\lambda_1^k v_1} + \alpha_2 \underbrace{A^k v_2}_{\lambda_2^k v_2} + \cdots + \alpha_n \underbrace{A^k v_n}_{\lambda_n^k v_n} = \alpha_1 \lambda_1^k v_1 + \alpha_2 \lambda_2^k v_2 + \cdots + \alpha_n \lambda_n^k v_n.$$

$$\Rightarrow \boxed{A^k P_0 = \alpha_1 \lambda_1^k v_1 + \alpha_2 \lambda_2^k v_2 + \cdots + \alpha_n \lambda_n^k v_n}.$$

As we have already commented, we see how the calculation of the matrix power A^k reduces to the simpler calculation of the numerical powers $\lambda_1^k, \lambda_2^k, \ldots, \lambda_n^k$.

Example 184. Suppose three investment groups, which we will call A, B and C, manage most of their capital themselves but diversify their investment by allocating a percentage to one of the other two groups. From one year to the next, they keep the investment percentages fixed according to the following table:

		invests in		
		A	**B**	**C**
Group	**A**	90%	30%	30%
	B	10%	70%	20%
	C	10%	10%	60%

From the data in the table, it follows that group A manages 90% of its funds itself and invests 30% in B and another 30% in C. If we sum the percentages, we see that the total investment of group A is (90+30+30)% = 150% and therefore, each year group A receives 50% profits which it again allocates to investment. The same analysis performed for the other two groups reveals that the investment of group B amounts to 100% of its capital while group C reinvests only 80% (the other 20% could be losses or capital allocated to other purposes).

Suppose that initially the capital in each group is, in millions of euros, as follows:

	Group A	**Group B**	**Group C**
Capital	17	27	21

Let us study the capital in subsequent years. To do this, we will set up a matrix model for this problem. We will start by calling

$$P_0 = \begin{pmatrix} 17 \\ 27 \\ 21 \end{pmatrix}.$$

the initial data 3-tuple. It is clear that if in year k, we have capital A_k in group A, B_k in group B and C_k in group C, in the next year (year $k+1$) we will have:

A_{k+1}=Capital in A in year $k+1$= 90% of A_k+10% of B_k+10% of C_k=$0.9\mathrm{A}_k$+$0.1\mathrm{B}_k$+$0.1\mathrm{C}_k$,

B_{k+1}=Capital in B in year $k+1$= 30% of A_k+70% of B_k+10% of C_k=$0.3\mathrm{A}_k$+$0.7\mathrm{B}_k$+$0.1\mathrm{C}_k$,

C_{k+1}=Capital in C in year $k+1$= 30% of A_k+20% of B_k+60% of C_k=$0.3\mathrm{A}_k$+$0.2\mathrm{B}_k$+$0.6\mathrm{C}_k$.

Expressing these equalities in matrix form we have that

$$\begin{pmatrix} \mathrm{A}_{k+1} \\ \mathrm{B}_{k+1} \\ \mathrm{C}_{k+1} \end{pmatrix} = \begin{pmatrix} 0.9 & 0.1 & 0.1 \\ 0.3 & 0.7 & 0.1 \\ 0.3 & 0.2 & 0.6 \end{pmatrix} \cdot \begin{pmatrix} \mathrm{A}_k \\ \mathrm{B}_k \\ \mathrm{C}_k \end{pmatrix}$$

so that, as we have seen in previous examples, we arrive at

$$\begin{pmatrix} \mathrm{A}_k \\ \mathrm{B}_k \\ \mathrm{C}_k \end{pmatrix} = \begin{pmatrix} 0.9 & 0.1 & 0.1 \\ 0.3 & 0.7 & 0.1 \\ 0.3 & 0.2 & 0.6 \end{pmatrix}^k \cdot \begin{pmatrix} \mathrm{A}_0 \\ \mathrm{B}_0 \\ \mathrm{C}_0 \end{pmatrix}. \tag{6.2}$$

If we denote

$$P_k = \begin{pmatrix} \mathrm{A}_k \\ \mathrm{B}_k \\ \mathrm{C}_k \end{pmatrix} \quad \text{and} \quad A = \begin{pmatrix} 0.9 & 0.1 & 0.1 \\ 0.3 & 0.7 & 0.1 \\ 0.3 & 0.2 & 0.6 \end{pmatrix}$$

abbreviatedly the matrix equation (6.2) is written as

$$P_k = A^k P_0.$$

The 3-tuple P_k that contains the data on the capital of the three groups in year k can be obtained through the calculation of the matrix power A^k but in this case we will use the power method described on page 225. For this, we start by calculating the eigenvalues and eigenvectors of the matrix A.

The characteristic polynomial of A is

$$p(\lambda) = |A - \lambda I_3| = \left| \begin{pmatrix} 0.9-\lambda & 0.1 & 0.1 \\ 0.3 & 0.7-\lambda & 0.1 \\ 0.3 & 0.2 & 0.6-\lambda \end{pmatrix} \right| = -\lambda^3 + 2.2\lambda^2 - 1.51\lambda + 0.33.$$

To calculate the eigenvalues we must solve the equation (we have multiplied by -1 so that the coefficient accompanying λ^3 is positive)

$$\lambda^3 - 2.2\lambda^2 + 1.51\lambda - 0.33 = 0.$$

However, it is easy to check that the sum of all rows of A is equal to 1.1 so part *iv)* of **Property 180** allows us to assert that $\lambda = 1.1$ is an eigenvalue of A. In this way we already know that one of the solutions of the characteristic equation is $\lambda = 1.1$. If we divide using Ruffini's method we have

	1	−2.2	1.51	−0.33
1.1		1.1	−1.21	0.33
	1	−1.1	0.3	0

By performing the division by Ruffini we confirm that indeed $\lambda = 1.1$ is a solution of the equation (since we obtain remainder equal to zero) and moreover the non-zero coefficients in the last row (i.e., 1, -1.1 and 0.3) indicate that the equation left to solve is

$$1 \cdot \lambda^2 - 1.1\lambda + 0.3 = 0.$$

But this last one is a second-degree equation that can be solved directly, obtaining as a result

$$\lambda = \frac{1.1 \pm \sqrt{1.1^2 - 4 \cdot 1 \cdot 0.3}}{2 \cdot 1} \Rightarrow \lambda = 0.6 \quad \text{and} \quad \lambda = 0.5.$$

In this way we have that the matrix A has the following eigenvalues

$$\lambda_1 = 1.1, \quad \lambda_2 = 0.6, \quad \lambda_3 = 0.5.$$

Since A is a square matrix of order three with three distinct eigenvalues, applying part *iii)* of **Property 180**, we know that the matrix A is diagonalizable and we will be able to obtain a basis of eigenvectors. Let us next calculate the eigenvectors corresponding to the calculated eigenvalues:

- **Eigenvectors associated with $\lambda_1 = 1.1$:** The eigenvectors associated with $\lambda_1 = 1.1$ form the eigenspace $V_{1.1}$ which has implicit equations

$$V_{1.1} \equiv (A - 1.1I_3) \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Rightarrow V_{1.1} \equiv \begin{pmatrix} -0.2 & 0.1 & 0.1 \\ 0.3 & -0.4 & 0.1 \\ 0.3 & 0.2 & -0.5 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

  It is easy to check that $V_{1.1} = \langle (1, 1, 1) \rangle$ and therefore $B_{1.1} = \{(1, 1, 1)\}$ is a basis for $V_{1.1}$.

- **Eigenvectors associated with $\lambda_2 = 0.6$:** The eigenvectors associated with $\lambda_2 = 0.6$ form the eigenspace $V_{0.6}$ which has implicit equations

$$V_{0.6} \equiv (A - 0.6I_3) \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Rightarrow V_{0.6} \equiv \begin{pmatrix} 0.3 & 0.1 & 0.1 \\ 0.3 & 0.1 & 0.1 \\ 0.3 & 0.2 & 0 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

  In this case $V_{0.6} = \langle (-2, 3, 3) \rangle$ and $B_{0.6} = \{(-2, 3, 3)\}$ is a basis for $V_{0.6}$.

- **Eigenvectors associated with $\lambda_3 = 0.5$:** The eigenvectors associated with $\lambda_3 = 0.5$ form the eigenspace $V_{0.5}$ which has implicit equations

$$V_{0.5} \equiv (A - 0.5I_3) \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Rightarrow V_{0.5} \equiv \begin{pmatrix} 0.4 & 0.1 & 0.1 \\ 0.3 & 0.2 & 0.1 \\ 0.3 & 0.2 & 0.1 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

  Now $V_{0.5} = \langle (1, 1, -5) \rangle$ and $B_{0.5} = \{(1, 1, -5)\}$ is a basis of $V_{0.5}$.

Part *ii)* of **Property 176** guarantees that by combining the vectors from $B_{1.1}$, $B_{0.6}$ and $B_{0.5}$ we obtain a set of independent vectors so We obtain a basis of eigenvectors of A formed by the vectors

$v_1 = (1, 1, 1)$ associated with the eigenvalue $\lambda_1 = 1.1$,
$v_2 = (-2, 3, 3)$ associated with the eigenvalue $\lambda_2 = 0.6$,
$v_3 = (1, 1, -5)$ associated with the eigenvalue $\lambda_3 = 0.5$.

The vectors v_1, v_2 and v_3 are a basis of $\mathbb{R}^3$ and to apply the power method we need to express the initial values tuple, P_0, as a linear combination of them. That is, we need to find the coefficients α_1, α_2 and α_3 such that

$$P_0 = \alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3 \Rightarrow \begin{pmatrix} 17 \\ 27 \\ 21 \end{pmatrix} = \alpha_1 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + \alpha_2 \begin{pmatrix} -2 \\ 3 \\ 3 \end{pmatrix} + \alpha_3 \begin{pmatrix} 1 \\ 1 \\ -5 \end{pmatrix} \Rightarrow \begin{pmatrix} 17 \\ 27 \\ 21 \end{pmatrix} = \begin{pmatrix} \alpha_1 - 2\alpha_2 + \alpha_3 \\ \alpha_1 + 3\alpha_2 + \alpha_3 \\ \alpha_1 + 3\alpha_2 - 5\alpha_3 \end{pmatrix}$$

$$\Rightarrow \begin{cases} \alpha_1 - 2\alpha_2 + \alpha_3 = 17 \\ \alpha_1 + 3\alpha_2 + \alpha_3 = 27 \\ \alpha_1 + 3\alpha_2 - 5\alpha_3 = 21 \end{cases}$$

and solving this system we obtain $\alpha_1 = 20$, $\alpha_2 = 2$, $\alpha_3 = 1$ and therefore

$$P_0 = 20v_1 + 2v_2 + v_3 \quad \text{or equivalently} \quad \begin{pmatrix} 17 \\ 27 \\ 21 \end{pmatrix} = 20 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + 2 \begin{pmatrix} -2 \\ 3 \\ 3 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \\ -5 \end{pmatrix}.$$

In this way, to calculate $A^k P_0$ we proceed as on page 226 in the form

$$A^k P_0 = 20A^k v_1 + 2A_k v_2 + A^k v_3 = 20 \cdot 1.1^k v_1 + 2 \cdot 0.6^k v_2 + 0.5^k v_3$$

or, equivalently,

$$\begin{pmatrix} 0.9 & 0.1 & 0.1 \\ 0.3 & 0.7 & 0.1 \\ 0.3 & 0.2 & 0.6 \end{pmatrix}^k \cdot \begin{pmatrix} 17 \\ 27 \\ 21 \end{pmatrix} = 20 \cdot 1.1^k \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + 2 \cdot 0.6^k \begin{pmatrix} -2 \\ 3 \\ 3 \end{pmatrix} + 0.5^k \begin{pmatrix} 1 \\ 1 \\ -5 \end{pmatrix}.$$

Now, by means of the expressions we have obtained, we can calculate the capital in each group after any number of years. For example:

- After $k = 3$ years, the capitals in each group will be determined by the tuple $P_3 = A^3 P_0$ which can be calculated via

$$P_3 = A^3 P_0 = 20 \cdot 1.1^3 v_1 + 2 \cdot 0.6^3 v_2 + 0.5^3 v_3 = 26.62v_1 + 2 \cdot 0.432v_2 + 0.125v_3$$

$$= 26.62 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + 0.432 \begin{pmatrix} -2 \\ 3 \\ 3 \end{pmatrix} + 0.125 \begin{pmatrix} 1 \\ 1 \\ -5 \end{pmatrix} = \begin{pmatrix} 25.881 \\ 28.041 \\ 27.291 \end{pmatrix}.$$

  Therefore, after three years, the capital in group A is 25.881 million euros, in group B 28.041 million and in group C 27.291 million.

- After $k = 10$ years, the capitals in each group will be determined by the tuple $P_{10} = A^{10} P_0$ which we can calculate as:

$$P_{10} = A^{10} P_0 = 20 \cdot 1.1^{10} v_1 + 2 \cdot 0.6^{10} v_2 + 0.5^{10} v_3 = 51.8748v_1 + 2 \cdot 0.0120932v_2 + 0.000976563v_3$$

$$= 51.8748 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + 0.0120932 \begin{pmatrix} -2 \\ 3 \\ 3 \end{pmatrix} + 0.000976563 \begin{pmatrix} 1 \\ 1 \\ -5 \end{pmatrix} = \begin{pmatrix} 51.8516 \\ 51.9121 \\ 51.9062 \end{pmatrix}.$$

  Thus, after ten years, the capital in group A rises to 51.8516 million euros, in group B is 51.9121 million and in group C 51.9062 million.

Using this method it is equally easy to calculate the capitals after any number of years.

We will now focus on the study of the trend for iterative matrix models. We assume then that we continue with a matrix model in which the tuple that provides the values for period k, P_k, is calculated via the matrix equation

$$P_k = A^k P_0,$$

where $A \in \mathcal{M}_n$ is the transition matrix, P_0 is the n-tuple of initial values and $k \in \mathbb{N}$ the number of elapsed periods.

Earlier we saw that if the matrix A is diagonalizable and we have a basis formed by the eigenvectors $v_1, v_2, \ldots, v_n$ associated respectively with the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$, it is easy to perform the calculation $A^k P_0$ if we have an expression of the initial values tuple P_0 as a linear combination of the eigenvectors,

$$P_0 = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n.$$

Then, the calculation of the power $A^k P_0$ was simple through the identity

$$A^k P_0 = \alpha_1 \lambda_1^k v_1 + \alpha_2 \lambda_2^k v_2 + \cdots + \alpha_n \lambda_n^k v_n. \tag{6.3}$$

If we study the right-hand side of this equality we observe that all the elements involved in it are constants ($\alpha_1, \alpha_2, \ldots, \alpha_n$ or $v_1, v_2, \ldots, v_n$ are coefficients or eigenvectors that we will have calculated previously). As k increases, the only elements of the expression that vary are the numerical powers $\lambda_1^k, \lambda_2^k, \ldots, \lambda_n^k$ which we underline below,

$$\alpha_1 \underline{\lambda_1^k} v_1 + \alpha_2 \underline{\lambda_2^k} v_2 + \cdots + \alpha_n \underline{\lambda_n^k} v_n.$$

Among the underlined powers, as k increases, the one corresponding to the largest eigenvalue will grow more rapidly. It is for this reason that the largest eigenvalue determines (in the sense we will see later) the behavior of the matrix model when k grows. This motivates us to give a name to that largest eigenvalue in the following definition.

Definition 185. An eigenvalue of a matrix A is said to be the dominant eigenvalue if its absolute value is greater than that of the rest of the eigenvalues of the matrix. An eigenvector associated with the dominant eigenvalue is said to be a dominant eigenvector.

Examples 186.
1) The eigenvalues of the matrix

$$A = \begin{pmatrix} 25 & -40 & -31 \\ 2 & 1 & -2 \\ 18 & -36 & -24 \end{pmatrix}$$

are $\lambda_1 = -6$, $\lambda_2 = 5$ and $\lambda_3 = 3$. If we calculate the absolute value of these eigenvalues we have that

$$|\lambda_1| = 6, \quad |\lambda_2| = 5, \quad |\lambda_3| = 3.$$

The absolute value of the eigenvalue $\lambda_1 = -6$ is greater than that of the other eigenvalues. Therefore, $\lambda_1 = -6$ is the dominant eigenvalue of the matrix A. It is possible to calculate by the usual methods the eigenspace associated with the eigenvalue $\lambda_1 = -6$ obtaining that $V_{-6} = \langle (1, 0, 1) \rangle$. The vector $(1, 0, 1)$ (and all its linear combinations) is an eigenvector associated with the dominant eigenvalue $\lambda_1 = -6$ and therefore we will say that $(1, 0, 1)$ is a dominant eigenvector for the matrix A.

2) The eigenvalues of the matrix

$$A = \begin{pmatrix} 25 & -38 & -31 \\ 5 & -4 & -5 \\ 14 & -28 & -20 \end{pmatrix}$$

are $\lambda_1 = -6$, $\lambda_2 = 6$ and $\lambda_3 = 1$. The corresponding absolute values are

$$|\lambda_1| = 6, \quad |\lambda_2| = 6, \quad |\lambda_3| = 1.$$

The absolute value of the first two eigenvalues coincides. In that case none of the eigenvalues has an absolute value strictly greater than that of all the others and the matrix has no dominant eigenvalue.

Suppose that in the identity (6.3) the eigenvalue λ_1 is the dominant eigenvalue of the matrix A and that

$$|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \cdots \geq |\lambda_n|.$$

Consequently, we will have that v_1 is a dominant eigenvector of A. Taking the dominant eigenvalue as a common factor on the right-hand side of (6.3) we have that

$$A^k P_0 = \lambda_1^k \left(\alpha_1 v_1 + \alpha_2 \left(\frac{\lambda_2}{\lambda_1} \right)^k v_2 + \cdots + \alpha_n \left(\frac{\lambda_n}{\lambda_1} \right)^k v_n \right).$$

Since λ_1 is the dominant eigenvalue it is clear that

$$\left| \frac{\lambda_2}{\lambda_1} \right|, \ldots, \left| \frac{\lambda_n}{\lambda_1} \right| < 1$$

but for large values of k it is easy to check that if a number $r \in \mathbb{R}$ has an absolute value less than one ($|r| < 1$) then $r^k \approx 0$. In this way when k is large,

$$\left(\frac{\lambda_2}{\lambda_1} \right)^k \approx 0, \quad \left(\frac{\lambda_3}{\lambda_1} \right)^k \approx 0, \quad \left(\frac{\lambda_n}{\lambda_1} \right)^k \approx 0$$

and therefore when k becomes large we will have that

$$A^k P_0 = \lambda_1^k \left(\alpha_1 v_1 + \alpha_2 \underbrace{\left(\frac{\lambda_2}{\lambda_1} \right)^k}_{\approx 0} v_2 + \cdots + \alpha_n \underbrace{\left(\frac{\lambda_n}{\lambda_1} \right)^k}_{\approx 0} v_n \right)$$

$$\Rightarrow \boxed{A^k P_0 \approx \lambda_1^k \alpha_1 v_1}.$$

From this we draw the following conclusions:

- For large values of k, the behavior of $A^k P_0$ depends solely on the dominant eigenvalue and the dominant eigenvector.
- Depending on the value of λ_1, the expression $\alpha_1 \lambda_1^k v_1$ will have one behavior or another. Specifically, we have:
  - If $|\lambda_1| < 1$, for large values of k we will have that $\lambda_1^k \approx 0$ and in that case

    $$\alpha_1 \lambda_1^k v_1 \approx 0.$$

    In other words, the values of P_k in successive periods tend to vanish.
  - If $|\lambda_1| > 1$, for large values of k we will have that $\lambda^k \approx \pm\infty$ and then

    $$\alpha_1 \lambda_1^k v_1 \approx \pm\infty$$

    which means that the values in successive periods will grow or decrease without limit.
  - If $\lambda_1 = 1$, for large values of k we will have that

    $$\alpha_1 \lambda_1^k v_1 = \alpha_1 v_1$$

    and the data tuples in successive periods will tend to a constant equilibrium value given by αv_1.
- We have that for large values of k, the data in period k, P_k, can be calculated approximately by

  $$P_k = A^k P_0 \approx \alpha_1 \lambda_1^k v_1.$$

  In many situations it will be of interest to calculate the vector of percentages of P_k and then we will have that

  $$\text{vector of percentages of } P_k \approx \text{vector of percentages of } \alpha_1 \lambda_1^k v_1.$$

However, it is simple to check that

$$\text{vector of percentages of } \underbrace{\alpha_1\lambda_1^k}_{\text{number}}\ \underbrace{v_1}_{\text{vector}} = \text{vector of percentages of } v_1$$

so that

$$\text{vector of percentages of } P_k \approx \text{vector of percentages of } v_1.$$

In other words, for large values of k, the percentages represented by the data in different periods will be approximately equal to those of the dominant eigenvector v_1.

Examples 187.

1) In **Example 184** we were studying the problem of three financial groups that invest according to a certain fixed annual investment table that led to a matrix model for the calculation of the capitals of the three groups in successive periods of the form

$$P_k = \underbrace{\begin{pmatrix} 0.9 & 0.1 & 0.1 \\ 0.3 & 0.7 & 0.1 \\ 0.3 & 0.2 & 0.6 \end{pmatrix}^k}_{=A} \cdot \underbrace{\begin{pmatrix} 17 \\ 27 \\ 21 \end{pmatrix}}_{=P_0}.$$

We saw that the transition matrix A has eigenvalues

$$\lambda_1 = 1.1, \quad \lambda_2 = 0.6, \quad \lambda_3 = 0.5$$

so the dominant eigenvalue is $\lambda_1 = 1.1$ and the corresponding dominant eigenvector is $v_1 = (1, 1, 1)$. On the other hand, the expression of the initial data tuple P_0 in the basis of eigenvectors v_1, v_2 and v_3 calculated on page 228 is

$$P_0 = \underbrace{20}_{=\alpha_1} v_1 + 2v_2 + v_3.$$

Recalling the reasoning from page 231 we have that:

- For large values of k we have that
$$P_k \approx 20 \cdot 1.1^k v_1.$$
For example:
  - After $k = 3$ years, the capital tuple, P_3, can be calculated approximately as
$$P_3 \approx 20 \cdot 1.1^3 v_1 = 26.62 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 26.62 \\ 26.62 \\ 26.62 \end{pmatrix}.$$
  - After $k = 10$ years, the capital tuple, P_{10}, can be calculated approximately as
$$P_{10} \approx 20 \cdot 1.1^{10} v_1 = 51.8748 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 51.8748 \\ 51.8748 \\ 51.8748 \end{pmatrix}.$$

  It can be checked how even for not excessively high values of k the approximations provide results very similar to the exact data we obtained on page 229.

- Since the dominant eigenvalue satisfies $|\lambda_1| = |1.1| = 1.1 > 1$, we have that
$$P_k \approx \alpha_1 1.1^k v_1 = 20 \cdot 1.1^k \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$
and the capitals of the three groups grow without limit during the course of successive years.

- The percentages represented by the capitals for year k, when k is sufficiently large, will be approximately the same as those represented by the dominant eigenvector v_1. The vector of percentages of v_1 is

$$\frac{100}{1+1+1}\begin{pmatrix}1\\1\\1\end{pmatrix}=\begin{pmatrix}33.\overline{3}\\33.\overline{3}\\33.\overline{3}\end{pmatrix}.$$

  Therefore, the future trend is that:

  $33.\overline{3}$ of the total capital will belong to group A.
  $33.\overline{3}$ of the total capital will belong to group B.
  $33.\overline{3}$ of the total capital will belong to group C.

  It is observed that the trend, after a sufficiently large number of years, is that the three groups accumulate capital of the same amount.

2) If we analyze **Example 182** we have that the eigenvalues of the transition matrix are

$$\lambda_1=1,\quad \lambda_2=0.4,\quad \lambda_3=0.7.$$

Therefore the dominant eigenvalue is $\lambda_1=1$. We had also calculated the eigenvectors associated with these eigenvalues; in particular, we saw that $(6,5,7)$ is an eigenvector associated with the dominant eigenvalue $\lambda_1=1$ so $v_1=(6,5,7)$ is a dominant eigenvector. If we also consider the eigenvectors associated with the other two eigenvalues we obtain the following basis of eigenvectors:

$$B=\{(6,5,7),(0,-1,1),(-3,1,2)\}.$$

When we first formulated this example on page 221 we saw that the initial data were: 120 customers in company A, 190 in B and 320 in C. This corresponded to an initial vector

$$P_0=\begin{pmatrix}210\\190\\320\end{pmatrix}.$$

If we calculate the coordinates of P_0 in B we obtain the following expression:

$$P_0=35v_1+30v_2+15v_3.$$

Then we have that the situation after k periods is approximated by

$$A^kP_0\approx 35\cdot 1^kv_1=35v_1.$$

Since the dominant eigenvalue is equal to one, we have a stability situation in which the distribution of the companies will stabilize around the limiting value $35v_1=35(6,5,7)$. On page **??** we saw that this vector represented the percentages

$$(33.\overline{3}\%,27.\overline{7}\%,38.\overline{8}\%)$$

and the distribution we can expect for the future will be:

- Customers in company A $=33.\overline{3}\%$ of the total.
- Customers in company B $=27.\overline{7}\%$ of the total.
- Customers in company C $=38.\overline{8}\%$ of the total.