Linear Algebra and Applications

Presentation at COEP Technological University, Pune

By

Nilesh Khandekar

Department of Mathematics
Savitribai Phule Pune University
Pune-411 007

August 02, 2023

Linear Transformations

$V = \mathbb{R}^n$ (n-finite dimensional vector space over $\mathbb{R}$)

$A = (a_{i,j})$ denotes $n \times n$ matrix with entries from $\mathbb{R}$

Note that matrix A defines a map from

$$A : \mathbb{R}^n \longrightarrow \mathbb{R}^n$$

given by

$$\mathbb{R}^n \ni v \mapsto Av \in \mathbb{R}^n$$

This map satisfies two important properties :

Linear Transformations

(i) $A(v + w) = A(v) + A(w)$ for all $v, w \in \mathbb{R}^n$ and

(ii) $A(\alpha v) = \alpha A(v)$ for all $\alpha \in \mathbb{R}$ and for all $v \in \mathbb{R}^n$.

Any map

$$T : \mathbb{R}^n \longrightarrow \mathbb{R}^n$$

satisfying the properties above is called a Linear Transformation.

There are a lot of interesting examples of linear transformations. This makes the study of linear transformations interesting and useful. Let us look at some examples.

Examples

Example-1 : Let

$$T : \mathbb{R} \longrightarrow \mathbb{R}$$

be given by

$$T(x) = 2x$$

Example-2 : Let

$$T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$$

be given by $T(v) = Av$ for any 2×2 matrix A.

Eigenvalues and Eigenvectors

Let A be $n \times n$ matrix with real entries.

Definition

A non-zero vector $v \in \mathbb{R}^n$ is called eigenvector of A with eigenvalue $\lambda \in \mathbb{R}$ if
$$Av = \lambda v$$

Let
$$A = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$$

Note $A\begin{pmatrix}1\\0\end{pmatrix} = 2\begin{pmatrix}1\\0\end{pmatrix}$ and $A\begin{pmatrix}0\\1\end{pmatrix} = 3\begin{pmatrix}0\\1\end{pmatrix}$.

Thus e_1 and e_2 are eigenvectors with eigenvalues 2 and 3 respectively. It is easy to show that :

λ is an eigenvalue of $A \iff \det(A - \lambda I) = 0$

Application to Markov Chains

Let us now look at an elementary application of Linear Algebra to population dynamics. Suppose if the population of a city and its suburbs were measured each year, then a vector such as

$$x_0 = \begin{pmatrix} 0.60 \\ 0.40 \end{pmatrix}$$

indicates that :

i 60% of the population lives in the city

ii 40% lives in the suburbs

The entries in the vector x_0 add up to 1 because they account for the entire population of the region. A vector with non-negative entries that add up to 1 is called a **probability vector**.

Markov Chain

A **stochastic matrix** is a matrix whose columns are probability vectors.

Definition

A **Markov chain** is a sequence of probability vectors $x_0, x_1, \dots,$ together with a stochastic matrix P such that

$$x_1 = Px_0, x_2 = Px_1, x_3 = Px_2, \dots$$

Markov chain of vectors in $\mathbb{R}^n$ describes probabilities that the outcome of an experiment is one of n possible outcomes. For this reason, x_k is often called a **state vector.**

Let us understand this with the help of an example.

Markov Chain

Fix an initial year- say 2020 and denote the populations of the city and suburbs that year by r_0 and s_0, respectively. Let x_0 be the population vector :

$$x_0 = \begin{pmatrix} r_0 \\ s_0 \end{pmatrix}$$

For 2021 and subsequent years, denote the population of the city and suburbs by the vectors

$$x_1 = \begin{pmatrix} r_1 \\ s_1 \end{pmatrix}, x_2 = \begin{pmatrix} r_2 \\ s_2 \end{pmatrix}, x_3 = \begin{pmatrix} r_3 \\ s_3 \end{pmatrix} \dots$$

Our aim is to describe mathematically how these vectors are related and understand their long term behaviour.

Markov Chain

Suppose a demographic study shows that each year about 5% of the city's population moves to the suburbs and 95% remains in the city. Also the study shows that 3% of the suburban population moves to the city and 97% remains in the suburbs.

Markov Chain

The vector representing the population distribution in the year 2021 is

$$\begin{pmatrix} r_1 \\ s_1 \end{pmatrix} = r_0 \begin{pmatrix} 0.95 \\ o.05 \end{pmatrix} + s_0 \begin{pmatrix} 0.03 \\ 0.97 \end{pmatrix} = \begin{bmatrix} 0.95 & 0.03 \\ 0.05 & 0.97 \end{bmatrix} \begin{pmatrix} r_0 \\ s_0 \end{pmatrix}$$

That is

$$x_1 = Mx_0,$$

where M is the migration matrix determined by the following table

$$\begin{array}{cc} \begin{array}{cc} city & suburb \end{array} & \\ \begin{pmatrix} 0.95 & 0.03 \\ 0.05 & 0.97 \end{pmatrix} & \begin{array}{c} city \\ suburb \end{array} \end{array}$$

The matrix above is a stochastic matrix. The equation above describes how population changes from 2020 to 2021.

Markov Chain

If the migartion percentages remain constant then the change from 2021 to 2022 is given by

$$x_2 = Mx_1$$

and similarly for 2022 to 2023 and subsequent years. In general

$$x_{k+1} = Mx_k \text{ for } k = 0, 1, 2, \ldots$$

The sequence of vectors $\{x_0, x_1, x_2, \ldots\}$ describes the population of the city/suburban region over a period of years.

Markov Chain

The most interesting aspect of Markov chains is the study of a chain's long-term behavior. For example what can be said about the population distribution in the long-run?

Suppose the 2020 population of the region is $6,00,000$ in the city and $4,00,000$ in the suburbs. Then the initial distribution in the region is given by

$$x_0 = \begin{pmatrix} 0.60 \\ 0.40 \end{pmatrix}$$

What is the population distribution in 2021? In 2022? More generally we can ask for the long term behavior of the population changes?

Markov Chain

Note that the population vector $\begin{pmatrix} 6,00,000 \\ 4,00,000 \end{pmatrix}$ changes to

$$\begin{bmatrix} 0.95 & 0.03 \\ 0.05 & 0.97 \end{bmatrix} \begin{pmatrix} 6,00,000 \\ 4,00,000 \end{pmatrix} = \begin{pmatrix} 5,82,000 \\ 4,18,000 \end{pmatrix}$$

Similarly the population distribution in 2022 is described by a vector x_2, where

$$x_2 = Mx_1 = \begin{bmatrix} 0.95 & 0.03 \\ 0.05 & 0.97 \end{bmatrix} \begin{pmatrix} 5,82,000 \\ 4,18,000 \end{pmatrix} = \begin{pmatrix} 5,65,000 \\ 4,35,000 \end{pmatrix}$$

Thus we have a sequence of population distribution vectors :

$$\begin{pmatrix} 6,00,000 \\ 4,00,000 \end{pmatrix} \to \begin{pmatrix} 5,82,000 \\ 4,18,000 \end{pmatrix} \to \begin{pmatrix} 5,65,000 \\ 4,35,000 \end{pmatrix} \to \cdots$$

Markov Chain

Computing a few more vectors in the sequence we get

$$\begin{pmatrix} 6,00,000 \\ 4,00,000 \end{pmatrix} \to \begin{pmatrix} 5,82,000 \\ 4,18,000 \end{pmatrix} \to \begin{pmatrix} 5,65,000 \\ 4,35,000 \end{pmatrix} \to \begin{pmatrix} 5,38,000 \\ 4,50,000 \end{pmatrix} \to$$

$$\to \begin{pmatrix} 5,13,000 \\ 4,63,400 \end{pmatrix} \to \begin{pmatrix} 4,90,980 \\ 4,74,760 \end{pmatrix} \to \begin{pmatrix} 4,70,854 \\ 4,85,066 \end{pmatrix} \to \cdots$$

Where does this sequence lead to?

Note for example that city population is decreasing.

Does it mean that the city population will diminish to zero eventually?

We answer this question as follows.

Steady State Vectors

Definition

If P is a stochastic matrix, then a **steady-state vector** (or **equilibrium vector**) for P is a probability vector q such that

$$Pq = q$$

It can be shown that every stochastic matrix has a steady-state vector. In the example above the probability vector

$$q = \begin{pmatrix} 3,75,000 \\ 6,25,000 \end{pmatrix}$$

is a steady state vector. It has the property that

$$\begin{bmatrix} 0.95 & 0.03 \\ 0.05 & 0.97 \end{bmatrix} \begin{pmatrix} 3,75,000 \\ 6,25,000 \end{pmatrix} = \begin{pmatrix} 3,75,000 \\ 6,25,000 \end{pmatrix}$$

Finding Steady State Vectors

A steady-state vector $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ for migration matrix M is a vector x such that $Mx = x$. Thus to find steady-state vector x we need to solve the matrix equation

$$Mx = x \ i.e \ (M - I)x = 0$$

This is a system of two equations in two variables x_1 and x_2. On solving this we get $x = \begin{pmatrix} 3 \\ 5 \end{pmatrix}$. This is a steady-state vector. Any scalar multiple of this vector is also a steady-state vector as is any scalar multiple of an eigenvector is eigenvector.

Steady State Vectors

A stochastic matrix P is said to be **regular** if some matrix power P^k contains only strictly positive entries.

Theorem

If P is an $n \times n$ regular stochastic matrix then P has a unique steady-state vector q.

Further if x_0 is any initial state and $x_{k+1} = Px_k$ for $k = 0, 1, 2, \dots$ then the Markov chain $\{x_k\}$ converges to q as $k \to \infty$.

Next we try to understand why this theorem is true?

Steady State Vectors

For the matrix

$$\begin{bmatrix} 0.95 & 0.03 \\ 0.05 & 0.97 \end{bmatrix}$$

the characteristic polynomial is

$$P(\lambda) = \lambda^2 - 1.92\lambda + 0.91 = (\lambda - 1)(\lambda - 0.92)$$

So the eigenvalues are : $\lambda = 1$ and $\lambda = 0.92$.

Eigenvectors corresponding to $\lambda = 1$ & $\lambda = 0.92$ are multiples of

$$v_1 = \begin{pmatrix} 3 \\ 5 \end{pmatrix} \quad \text{and} \quad v_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

Steady State Vectors

Recall that in our example the initial population distribution is

$$\begin{pmatrix} 6,00,000 \\ 4,00,000 \end{pmatrix}$$

Normalize the vector to

$$\begin{pmatrix} 0.6 \\ 0.4 \end{pmatrix}$$

As noted earlier we have $x_k = A^k x_0$.

The problem now is to compute x_k efficiently. To do this we write

$$x_0 = c_1 v_1 + c_2 v_2$$

and solve the following system for c_1 and c_2

$$\begin{pmatrix} 0.6 \\ 0.4 \end{pmatrix} = c_1 \begin{pmatrix} 3 \\ 5 \end{pmatrix} + c_2 \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

Steady State Vectors

Solving the system we obtain

$$c_1 = 0.125 \quad \text{and} \quad c_2 = 0.225$$

Thus we have

$$\begin{pmatrix} 0.6 \\ 0.4 \end{pmatrix} = 0.125 \begin{pmatrix} 3 \\ 5 \end{pmatrix} + 0.225 \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

Now

$$x_1 = Ax_0 = c_1 Av_1 + c_2 Av_2 = c_1 v_1 + c_2 (0.92) v_2$$

$$x_2 = Ax_1 = c_1 Av_1 + c_2 (0.92) Av_2 = c_1 v_1 + c_2 (0.92)^2 v_2$$

In general,

$$x_k = c_1 v_1 + c_2 (0.92)^k v_2 \quad \text{for} \quad k = 0, 1, 2 \ldots$$

Steady State Vectors

$$x_k = c_1 v_1 + c_2 (0.92)^k v_2 \quad \text{for} \quad k = 0, 1, 2 \ldots$$

So using values of c_1 and c_2 we get

$$x_k = 0.125 \begin{pmatrix} 3 \\ 5 \end{pmatrix} + 0.225 (0.92)^k \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

Now letting $k \to \infty$ we get,

$$\lim_{k \to \infty} x_k = 0.125 \begin{pmatrix} 3 \\ 5 \end{pmatrix} = \begin{pmatrix} 0.375 \\ 0.625 \end{pmatrix}$$

Thus the steady state population vector is

$$\begin{pmatrix} 3,75,000 \\ 6,25,000 \end{pmatrix}$$

Steady State Vectors

Under the given assumptions of rates of transfers of populations the city population eventually stabilizes to $3, 75, 000$ while the suburban population stabilizes to $6, 25, 000$

Notice here that the largest eigenvalue is 1. Hence the contribution of eigenvectors corresponding to smaller eigenvalues decreases as k increases.

References

1. Linear Algebra and its Applications by David Lay

Thank You!