Applications: Markov Chains, Optimizations, Group Representations

Jitendra Rathore

December 1, 2025

Today, we will discuss few applications of certain topics which we learnt in our course. We will mainly discuss three applications one is to Markov Chians, Optimizations problems and the last is to study groups using linear algebra. We have already discussed one application namely Least square approximation in our course. We begin with Stochastic process and Markov Chain.

1 Stochastic Process and Markov Chain

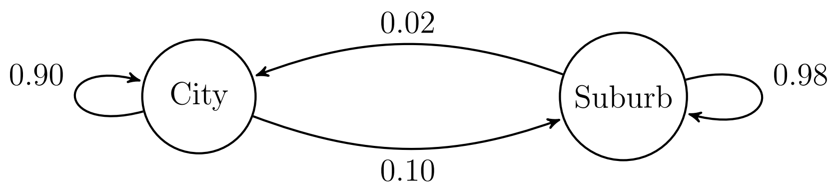
In many real-world situations, a population moves between a finite number of states over time. For example, suppose a metropolitan region consists of two regions: the city and the suburbs. The population remains constant, but people move between these two regions each year.

In the following example there are two states: **residing in the city** and **residing in the suburbs**. The probability that a city resident moves to the suburbs in one year is 0.10, and the probability that a suburb resident moves to the city is 0.02. The probability that a city resident remains in the city for one year is 0.90, and so on.

Probability	City	Suburb
City	0.90	0.02
Suburb	0.10	0.98

.

Visualizing Transitions:



We can write this above data in the form of a matrix, we write

$$A = \begin{pmatrix} 0.90 & 0.02 \\ 0.10 & 0.98 \end{pmatrix}.$$

The entries are nonnegative and each column sums to 1, expressing the assumption of constant population.

Definition. A square matrix is called a **transition matrix** (or **stochastic matrix**) if all its entries are nonnegative and each column sums to 1. In a transition matrix M, the entry M_{ij} denotes the probability of moving from state j to state i in one step.

The system moves between two states: **City (C)** and **Suburb (S)**. The transitions are modeled by the **Transition Matrix** A:

- A is a **Stochastic Matrix**: entries $A_{ij} \geq 0$; each column sums to 1.
- $\mathbf{P}_m$: A **Probability Vector** representing the distribution at time m (entries ≥ 0 and sum to 1).

Let A be the transition matrix and $\mathbf{P}_0$ be the initial probability vector (Year 2000).

$$A = \begin{pmatrix} A_{CC} & A_{SC} \\ A_{CS} & A_{SS} \end{pmatrix} = \begin{pmatrix} 0.90 & 0.02 \\ 0.10 & 0.98 \end{pmatrix} \qquad \mathbf{P}_0 = \begin{pmatrix} P_C \\ P_S \end{pmatrix} = \begin{pmatrix} 0.70 \\ 0.30 \end{pmatrix}$$

- $P_C = 0.70$: Initial proportion in City (C).
- $P_S = 0.30$: Initial proportion in Suburb (S).
- $A_{CC} = 0.90$: Prob. of staying in C (C $\to$ C).
- $A_{CS} = 0.02$: Prob. of moving from S to C (S $\to$ C).

The distribution in the next year (2001) is $\mathbf{P}_1 = A\mathbf{P}_0 = \begin{pmatrix} P'_C \\ P'_S \end{pmatrix}$.

The first coordinate, P'_C (proportion in the City in 2001), is found by multiplying the **first row** of A by the column vector $\mathbf{P}_0$:

$$P'_C = (A_{CC})(P_C) + (A_{CS})(P_S)$$
$$P'_C = (0.90)(0.70) + (0.02)(0.30)$$
$$P'_C = 0.630 + 0.006 = 0.636$$

Hence, similar calculation for P'_S leads the following vector: $\mathrm{P}_1 = \begin{pmatrix} \mathbf{0.636} \\ 0.364 \end{pmatrix}$

Evolution of Distribution: The distribution after m years is given by the matrix product $\mathbf{P}_m = A^m\mathbf{P}_0$.

- $\mathbf{P}_1 = A\mathbf{P}_0$: The distribution after **one** year.
- $\mathbf{P}_2 = A^2\mathbf{P}_0 = A(A\mathbf{P}_0)$: The distribution after **two** years. This represents applying the one-year transition probabilities A to the population distribution $\mathbf{P}_1$. The matrix A^2 itself contains the **two-step transition probabilities**: $(A^2)_{ij}$ is the probability of moving from state j to state i in two steps.

Example with $\mathbf{P}_0 = \begin{pmatrix} 0.70 \\ 0.30 \end{pmatrix}$ (Year 2000):

Time Step	Calculation	Distribution $\mathbf{P}_m$
$m = 1$ (2001)	$\mathbf{P}_1 = A\mathbf{P}_0$	$\begin{pmatrix} 0.636 \\ 0.364 \end{pmatrix}$
$m = 2$ (2002)	$\mathbf{P}_2 = A^2\mathbf{P}_0$	$\begin{pmatrix} 0.57968 \\ 0.42032 \end{pmatrix}$

Long-Term Distribution: The long-term distribution is the **Steady-State Vector** $\mathbf{v}$, found by taking the limit:

$$\lim_{m\to\infty} A^m \mathbf{P}_0 = \mathbf{v}$$

$\mathbf{v}$ is the unique probability eigenvector corresponding to the eigenvalue $\lambda = 1$:

$$A\mathbf{v} = \mathbf{v}$$

$$\mathbf{v} = \begin{pmatrix} \mathbf{0.1667} \\ \mathbf{0.8333} \end{pmatrix}$$

The population stabilizes at 16.67% in the City and 83.33% in the Suburbs.

The city-suburb problem is an example of a process in which elements of a set are each classified as being in one of several fixed states that can switch over time. In general, such a process is called a **stochastic process**.

If, however, the probability that an object in one state changes to a different state in a fixed interval of time depends only on the two states (and not on the time, earlier states, or other factors), then the stochastic process is called a **Markov process**. If, in addition, the number of possible states is finite, then the Markov process is called a **Markov chain**. Markov chains have many applications as statistical models of real-world processes.

2 Optimization Problems

Definition 2.1. A **quadratic form** Q on $\mathbb{R}^n$ is a function that can be written as:

$$Q(\mathbf{x}) = \mathbf{x}^T A \mathbf{x}$$

where $\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ is a vector in $\mathbb{R}^n$ and A is an $n \times n$ **symmetric matrix**.

Example 2.2. If $A = \begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix}$ and $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, the quadratic form is:

$$Q(\mathbf{x}) = \mathbf{x}^T A \mathbf{x} = a x_1^2 + b x_1 x_2 + c x_2^2$$

A quadratic form $Q(\mathbf{x}) = \mathbf{x}^T A \mathbf{x}$ (with symmetric matrix A) is classified as:

- **Positive Definite:** If $Q(\mathbf{x}) > 0$ for all $\mathbf{x} \neq \mathbf{0}$. (Equivalent to all eigenvalues of A being positive.)
- **Negative Definite:** If $Q(\mathbf{x}) < 0$ for all $\mathbf{x} \neq \mathbf{0}$. (Equivalent to all eigenvalues of A being negative.)
- **Indefinite:** If $Q(\mathbf{x})$ takes on both positive and negative values. (Equivalent to A having both positive and negative eigenvalues.)

Problem: We consider the problem of finding the maximum and minimum values of a quadratic form $Q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ subject to the **unit vector constraint**:

$$\|\mathbf{x}\| = 1 \quad \text{or equivalently} \quad \mathbf{x}^T\mathbf{x} = 1$$

Theorem 2.3. *Let A be an $n \times n$ **symmetric matrix**, and define the minimum value m and the maximum value M of the quadratic form $Q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ on the unit sphere by:*

$$m = \min\{\mathbf{x}^T A\mathbf{x} \mid \|\mathbf{x}\| = 1\}, \quad M = \max\{\mathbf{x}^T A\mathbf{x} \mid \|\mathbf{x}\| = 1\}$$

Then:

1. *M is the **greatest eigenvalue** of A (denoted λ_{max}).*
2. *m is the **least eigenvalue** of A (denoted λ_{min}).*
3. *The value $\mathbf{x}^T A\mathbf{x}$ is M when $\mathbf{x}$ is a **unit eigenvector** corresponding to the greatest eigenvalue λ_{max}.*
4. *The value $\mathbf{x}^T A\mathbf{x}$ is m when $\mathbf{x}$ is a **unit eigenvector** corresponding to the least eigenvalue λ_{min}.*

Note: This theorem provides a powerful result: the extreme values of a quadratic form constrained to the unit sphere are determined entirely by the eigenvalues of its associated symmetric matrix A.

Example 2.4. Find the maximum and minimum values of the quadratic form $Q(\mathbf{x}) = 3x_1^2 + 2x_2^2$ subject to the constraint $x_1^2 + x_2^2 = 1$.

Solution: The quadratic form is $Q(\mathbf{x}) = 3x_1^2 + 2x_2^2 + 0x_1x_2$. The constraint is $\|\mathbf{x}\|^2 = 1$. The associated symmetric matrix A is:

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

and the form is $Q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x} = \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$.

Since A is a diagonal matrix, so we have $\lambda_1 = 3 \quad \text{and} \quad \lambda_2 = 2$. Applying the theorem, we get the maximum value ($\mathbf{M}$), which is also the greatest eigenvalue is $\lambda_{\max} = 3$. The minimum value is $\mathbf{m}$, which is also the least eigenvalue is $\lambda_{\min} = 2$.

Maximum Occurs at: The unit eigenvector corresponding to $\lambda_1 = 3$.

$$A\mathbf{x} = 3\mathbf{x} \implies \begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 3 \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \implies \begin{cases} 3x_1 = 3x_1 \\ 2x_2 = 3x_2 \end{cases}$$

This requires $x_2 = 0$. Since $\|\mathbf{x}\| = 1$, the optimal unit vectors are $\mathbf{v}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\mathbf{v}_1 = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$.

Minimum Occurs at: The unit eigenvector corresponding to $\lambda_2 = 2$.

$$A\mathbf{x} = 2\mathbf{x} \implies \begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 2 \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \implies \begin{cases} 3x_1 = 2x_1 \\ 2x_2 = 2x_2 \end{cases}$$

This requires $x_1 = 0$. The optimal unit vectors are $\mathbf{v}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ and $\mathbf{v}_2 = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$.

3 Group Representations

An important concept in mathematics, a **representation** ρ of a group G is a structure-preserving map, or **homomorphism**, G to the group of invertible linear transformations on a vector space V. If $V \cong \mathbb{C}^n$, then a homomorphism is a map

$$\rho : G \to GL_n(\mathbb{C}) \text{ such that } \rho(g_1 \cdot g_2) = \rho(g_1)\rho(g_2) \,\forall\, g_1, g_2 \in G. \tag{1}$$

where, $GL_n(\mathbb{C})$ is the group all $n \times n$ invertible matrices.

Via this map, we can view represent elements of group G as matrices in $GL(n, K)$, which is the group of all invertible $n \times n$ matrices over the field K (often $\mathbb{R}$ or $\mathbb{C}$).

This realization is the key link between abstract group theory and concrete linear algebra:

$$\text{Abstract Group Element } (g \in G) \quad \longrightarrow \quad \text{Concrete Matrix } (\mathbf{M}_g \in GL(n, K))$$

The homomorphism property, when translated into matrix language, ensures that the matrix representation of a product of group elements is the matrix product of their individual representations:

$$\mathbf{M}_{g_1 g_2} = \mathbf{M}_{g_1}\mathbf{M}_{g_2} \tag{2}$$

Table 1: Analogies and Similarities in Groups and Representation Theory

Group Theory Concept	**Linear Algebra Equivalent**
Group Representation ρ	Homomorphism from G to $GL(n, K)$
Group Element $g \in G$	Invertible $n \times n$ Matrix $\mathbf{M}_g \in GL(n, K)$
Group Action (g acting on $v \in V$)	Matrix-Vector Product ($\mathbf{M}_g\mathbf{v}$)
Equivalence of Representations ($\rho_1 \sim \rho_2$)	Similarity of Matrices ($\mathbf{M}'_g = \mathbf{P}\mathbf{M}_g\mathbf{P}^{-1}$)
Decomposition of a Representation ($\rho = \rho_1 \oplus \rho_2 \oplus \cdots$)	Block Diagonalization of the Matrix ($\mathbf{M}_g$ into irreducible blocks)
Irreducible Sub-representation (ρ_i)	Invariant Subspace of the representation space V
Character of the Representation ($\chi_\rho(g)$)	Trace of the Matrix ($\mathrm{Tr}(\mathbf{M}_g)$)