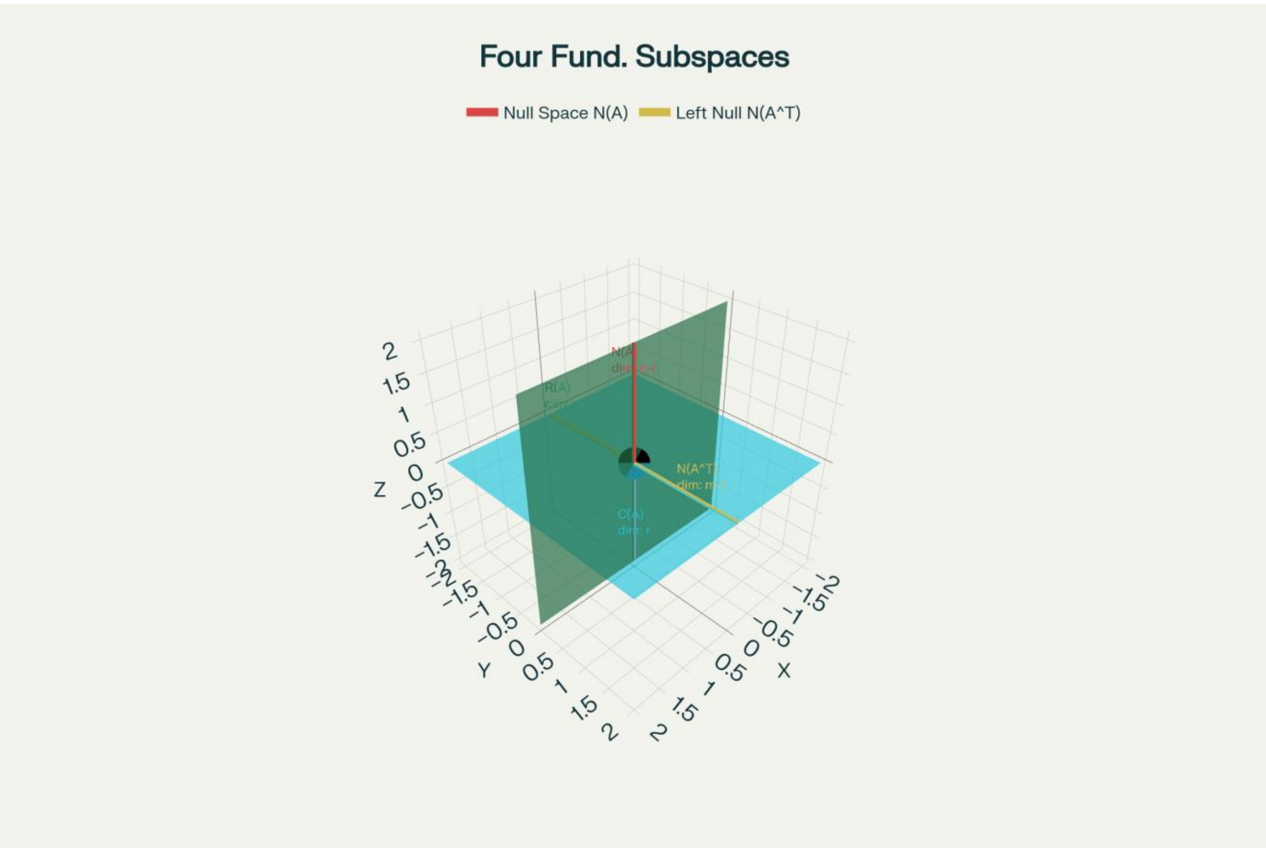
Linear Algebra: Four Fundamental Subspaces, Orthogonality, Projections, and Least Squares

This comprehensive report covers the fundamental concepts of linear algebra related to matrix subspaces, orthogonality, projections, and least squares methods. These topics form the cornerstone of understanding how linear systems work and provide the theoretical foundation for many applications in engineering, data science, and mathematical analysis.

Four Fundamental Subspaces of a Matrix

Every matrix A of size $m \times n$ gives rise to four fundamental subspaces that completely characterize its linear algebraic properties. These subspaces are intimately connected through orthogonality relationships and dimensional constraints that reveal the deep structure underlying linear transformations.



Four Fundamental Subspaces of Matrix A showing orthogonality relationships

Column Space C(A)

The **column space** $C(A)$ represents all possible linear combinations of the columns of matrix A. Mathematically, if $A = [u_1, u_2, \dots, u_n]$, then $C(A) = \text{span}(u_1, u_2, \dots, u_n)$. This subspace is crucial for understanding when the linear system $Ax = b$ has a solution, as the system is solvable if and only if $b \in C(A)$.

The dimension of the column space equals the rank of the matrix, denoted as $\dim(C(A)) = \text{rank}(A) = r$. For systems with more equations than unknowns (overdetermined systems), not every vector b will belong to $C(A)$, leading to inconsistent systems that require alternative approaches like least squares.

Null Space N(A)

The **null space** $N(A)$ consists of all vectors x such that $Ax = 0$. This subspace captures the "kernel" of the linear transformation represented by A and reveals the degrees of freedom in the solution space. The null space is indeed a subspace because it satisfies closure under addition and scalar multiplication: if $x_1, x_2 \in N(A)$ and α is a scalar, then $A(x_1 + x_2) = 0$ and $A(\alpha x) = 0$.

The dimension of the null space is called the **nullity** of A, and it equals $n - r$ where n is the number of columns and r is the rank. This relationship leads to the fundamental rank-nullity theorem: $\text{rank}(A) + \text{nullity}(A) = n$.

Row Space R(A)

The **row space** $R(A)$ is the column space of A^T, i.e., $R(A) = C(A^T)$. It represents the span of all row vectors of A. A remarkable fact in linear algebra is that the column rank always equals the row rank, meaning $\dim(C(A)) = \dim(R(A)) = r$. This equality demonstrates the fundamental symmetry between rows and columns in matrix theory.

Left Null Space N(A^T)

The **left null space** $N(A^T)$ consists of all vectors y such that $A^T y = 0$, or equivalently, $y^T A = 0$. These are vectors that, when viewed as row vectors, produce the zero vector when multiplied by A. The dimension of the left null space is $m - r$, where m is the number of rows.

Dimensional Relationships

The four fundamental subspaces satisfy important dimensional constraints that reflect the structure of the matrix:

- $\dim(C(A)) + \dim(N(A)) = n$ (number of columns)
- $\dim(C(A^T)) + \dim(N(A^T)) = m$ (number of rows)

These relationships ensure that the subspaces partition the entire vector spaces in a meaningful way.

Orthogonality and Vector Relationships

Orthogonality plays a central role in understanding the geometric structure of vector spaces and provides the foundation for projection methods and least squares solutions.

Vector Orthogonality

Two vectors x and y are **orthogonal** (denoted $x \perp y$) if their inner product is zero: $x^T y = 0$[2]. This condition extends the familiar geometric notion of perpendicularity to higher dimensions. The length or **norm** of a vector x is defined as $\|x\|^2 = x_1^2 + x_2^2 + \cdots + x_n^2 = x^T x$.

The Pythagorean theorem generalizes to higher dimensions: if $x \perp y$, then $\|x\|^2 + \|y\|^2 = \|x + y\|^2$. This relationship provides a powerful tool for verifying orthogonality and understanding geometric relationships in vector spaces.

Properties of Orthogonal Sets

Orthogonal sets of vectors possess remarkable properties that make them particularly useful in linear algebra:

1. The zero vector is orthogonal to every vector
2. Any set of mutually orthogonal non-zero vectors is linearly independent

The second property is particularly important: if $\{v_1, v_2, \dots, v_k\}$ are mutually orthogonal non-zero vectors, then any linear combination $c_1 v_1 + c_2 v_2 + \cdots + c_k v_k = 0$ implies $c_i = 0$ for all i. This follows from taking the inner product with each v_i and using orthogonality.

Orthonormal Vectors

Vectors that are both orthogonal and of unit length are called **orthonormal**. A set $\{u, v\}$ is orthonormal if $u^T v = 0$ and $\|u\| = \|v\| = 1$. Classic examples include the standard basis vectors and rotation matrices like $\begin{pmatrix} \cos\,\theta \\ \sin\,\theta \end{pmatrix}, \begin{pmatrix} -\sin\,\theta \\ \cos\,\theta \end{pmatrix}$.

Orthogonal Subspaces

Two subspaces U and V are **orthogonal** if every vector in U is orthogonal to every vector in V: $x^T y = 0$ for all $x \in U, y \in V$. This concept extends pairwise vector orthogonality to entire subspaces and reveals fundamental structural relationships.

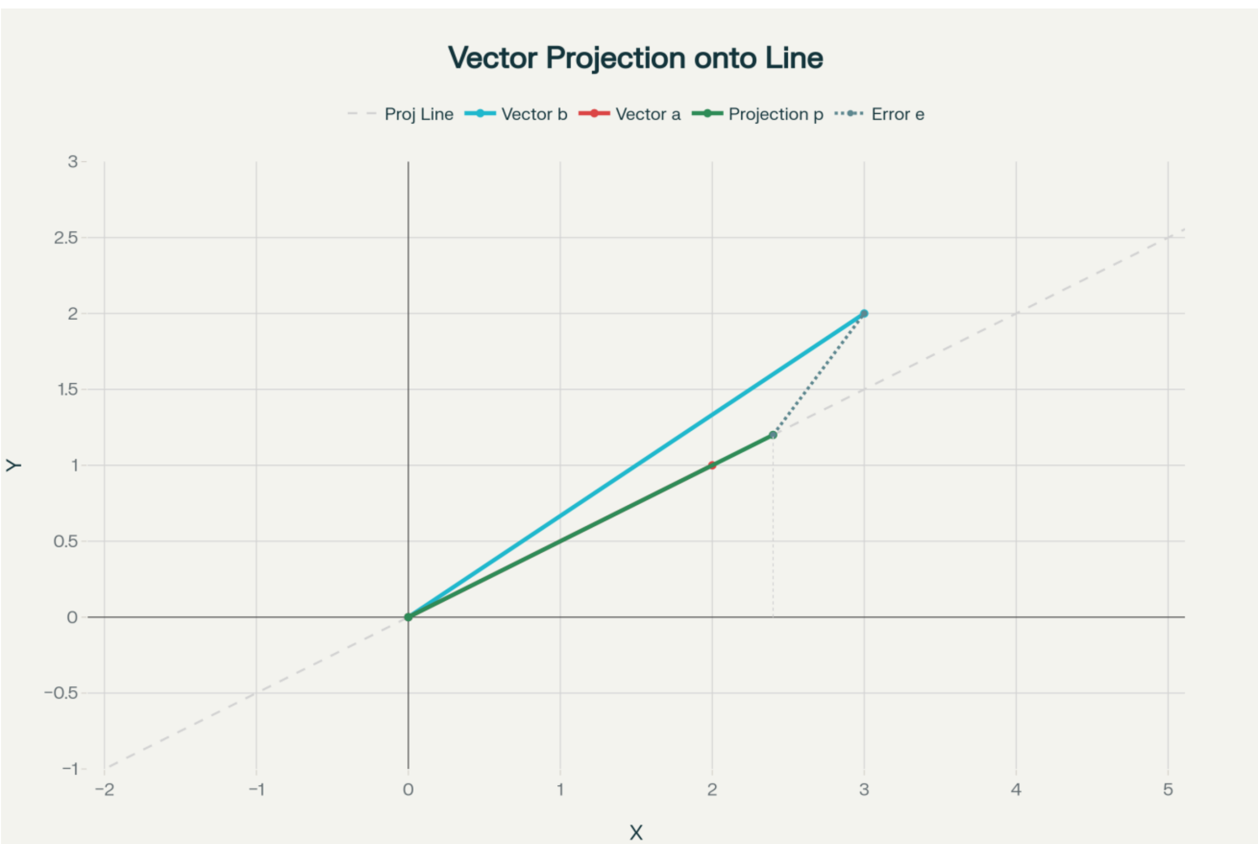
The most important orthogonality relationships among the four fundamental subspaces are:[2]

- $R(A) \perp N(A)$: The row space is orthogonal to the null space
- $C(A) \perp N(A^T)$: The column space is orthogonal to the left null space

These relationships can be proven directly from the definitions. For instance, if $x \in N(A)$, then $Ax = 0$, which means each row of A is orthogonal to x. Since $R(A)$ is the span of the rows, every vector in $R(A)$ is orthogonal to x.

Projections and Geometric Interpretations

Projection is a fundamental operation that finds the "closest" point in a subspace to a given vector, providing the geometric foundation for solving inconsistent linear systems.



Vector Projection onto a Line showing geometric interpretation

Projection onto a Line

Given a vector b and a line through the origin defined by vector a, the **projection** of b onto the line is the point on the line closest to b. The projection p is given by the formula:

$$p = \frac{a^T b}{a^T a} a$$

This formula ensures that the error vector $e = b - p$ is orthogonal to the line, i.e., $a^T e = 0$. The scalar $\frac{a^T b}{a^T a}$ represents the coordinate of the projection along the direction a.

Projection Matrix

The projection operation can be represented by a **projection matrix** P. For projection onto a line through vector a, the projection matrix is:

$$P = \frac{a a^T}{a^T a}$$

This matrix has several important properties:

1. **Symmetry**: $P^T = P$
2. **Idempotency**: $P^2 = P$ (projecting twice gives the same result)
3. **Rank**: $\text{rank}(P) = 1$ for line projection

The idempotency property reflects the geometric fact that if a vector is already on the line, projecting it again doesn't change it.

Projection onto General Subspaces

For projection onto the column space $C(A)$ of a matrix A, the projection formula generalizes to:[4]

$$P = A(A^T A)^{-1} A^T$$

This formula assumes that $A^T A$ is invertible, which occurs when the columns of A are linearly independent. The projection of vector b onto $C(A)$ is then $Pb = A(A^T A)^{-1} A^T b$.

Least Squares Method and Applications

The least squares method provides a systematic approach to solving overdetermined systems by finding the solution that minimizes the sum of squared errors.

The Least Squares Problem

Consider an inconsistent system $Ax = b$ where A is $m \times n$ with $m > n$. Since no exact solution exists, we seek the vector $\hat{x}$ that minimizes the **sum of squared residuals**:

$$E^2 = \|Ax - b\|^2 = (Ax - b)^T(Ax - b)$$

The least squares solution $\hat{x}$ is characterized by the **normal equations**:

$$A^T A\hat{x} = A^T b$$

This system always has a solution when $A^T A$ is invertible, giving $\hat{x} = (A^T A)^{-1} A^T b$.

Connection to Projections

The least squares method is intimately connected to projections. The solution $\hat{x}$ produces $A\hat{x}$, which is the projection of b onto the column space $C(A)$. The error vector $e = b - A\hat{x}$ is orthogonal to $C(A)$, which leads directly to the normal equations.

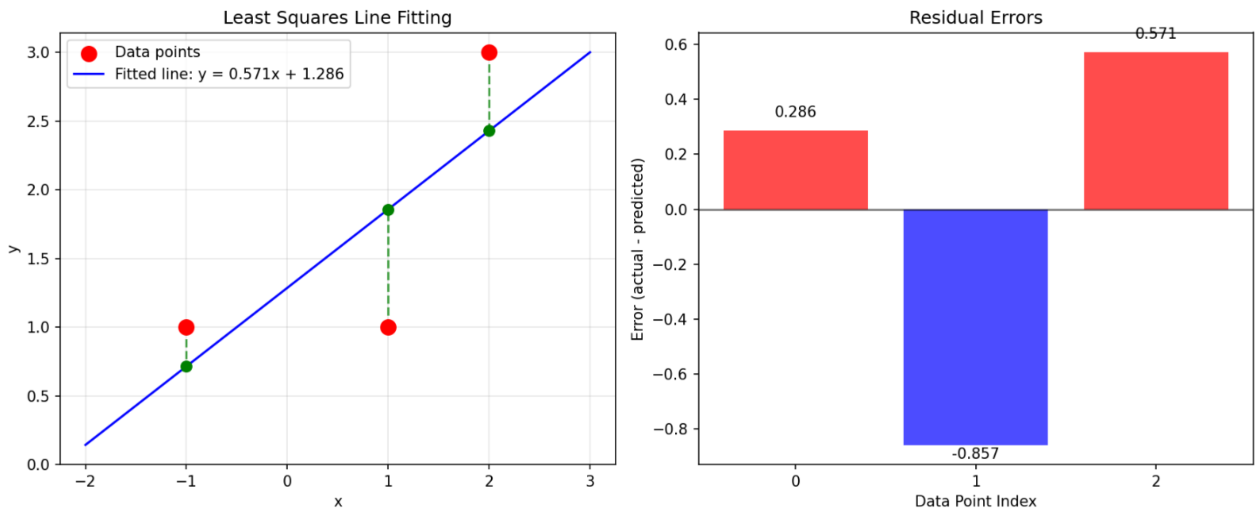
The geometric interpretation is that we're finding the point in $C(A)$ that is closest to b in the Euclidean sense[4]. This projection minimizes the distance $\|b - Ax\|$ over all possible x.

Linear Regression Example

A classic application is fitting a straight line to data points. Given data points $(x_1, b_1), (x_2, b_2), \dots, (x_m, b_m)$, we seek a line $y = \theta' x + \theta''$ that best fits the data in the least squares sense.

This can be formulated as the matrix equation:

$$\begin{pmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_m & 1 \end{pmatrix} \begin{pmatrix} \theta' \\ \theta'' \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$



Worked Example

Consider the specific example from the materials with data points $(-1,1)$, $(1,1)$, and $(2,3)$. The matrix equation becomes:

$$\begin{pmatrix} -1 & 1 \\ 1 & 1 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} \theta' \\ \theta'' \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix}$$

Computing the normal equations:

- $A^T A = \begin{pmatrix} 6 & 2 \\ 2 & 3 \end{pmatrix}$
- $A^T b = \begin{pmatrix} 6 \\ 5 \end{pmatrix}$

Solving gives $\theta' = \frac{4}{7}$ and $\theta'' = \frac{9}{7}$, so the best-fit line is $y = \frac{4}{7}x + \frac{9}{7}$.

The projections are $p_1 = \frac{5}{7}$, $p_2 = \frac{13}{7}$, $p_3 = \frac{17}{7}$, and the error vector is $e = (\frac{2}{7}, -\frac{6}{7}, \frac{4}{7})$. The sum of squared errors is $E^2 = \|e\|^2 = \frac{56}{49} = \frac{8}{7}$.

Properties of Projection Matrices

General projection matrices satisfy fundamental properties that make them essential tools in linear algebra:

1. **Symmetry**: $P^T = P$

2. **Idempotency**: $P^2 = P$
3. **Complementary projector**: $I - P$ is also a projection matrix

These properties ensure that projections behave geometrically as expected and provide computational advantages in numerical implementations.

Theoretical Foundations and Advanced Concepts

The concepts covered in this report form the theoretical backbone of modern linear algebra and have far-reaching applications across mathematics, engineering, and data science. The interplay between the four fundamental subspaces reveals the deep structure of linear transformations, while orthogonality provides the geometric framework for understanding these relationships.

The least squares method, emerging naturally from projection theory, represents one of the most important computational tools in applied mathematics. Its applications extend from simple curve fitting to complex machine learning algorithms, statistical inference, and optimization problems.

The dimensional relationships among subspaces, the orthogonality conditions, and the projection formulas all work together to provide a complete characterization of how linear systems behave. When exact solutions don't exist, the least squares approach gives us the "best" approximate solution in a precise mathematical sense.

Understanding these fundamental concepts is essential for anyone working with linear systems, whether in pure mathematics, engineering applications, or data analysis. The geometric intuition provided by projections, combined with the algebraic precision of the normal equations, creates a powerful framework for solving real-world problems involving linear relationships and optimization.