The Projection Matrix

Introduction to Projection

- Let A be an $n \times d$ matrix.
- In over-determined systems where $d < n$, the system $A\mathbf{x} = \mathbf{b}$ is often inconsistent.
- This occurs when the vector $\mathbf{b}$ does not lie in the span of the columns of A.

Introduction to Projection

- Let A be an $n \times d$ matrix.
- In over-determined systems where $d < n$, the system $A\mathbf{x} = \mathbf{b}$ is often inconsistent.
- This occurs when the vector $\mathbf{b}$ does not lie in the span of the columns of A.
- **Goal:** Find a feasible vector $\mathbf{b}'$ in the column space of A that is closest to $\mathbf{b}$.
- We solve the consistent system $A\mathbf{x} = \mathbf{b}'$.

Introduction to Projection

- Let A be an $n \times d$ matrix.
- In over-determined systems where $d < n$, the system $A\mathbf{x} = \mathbf{b}$ is often inconsistent.
- This occurs when the vector $\mathbf{b}$ does not lie in the span of the columns of A.
- **Goal:** Find a feasible vector $\mathbf{b}'$ in the column space of A that is closest to $\mathbf{b}$.
- We solve the consistent system $A\mathbf{x} = \mathbf{b}'$.
- The mapping from $\mathbf{b}$ to $\mathbf{b}'$ is a linear transformation defined by the **Projection Matrix**.

The Least-Squares Objective

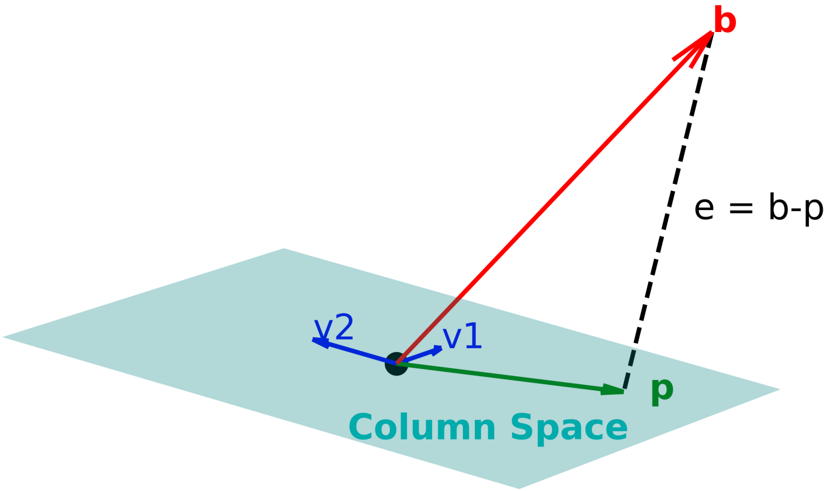
Consider a general $n \times d$ matrix A of full rank.

- We seek the coordinate vector $\mathbf{x}$ such that the projection $\mathbf{b}' = A\mathbf{x}$ minimizes the squared error:

$$\min_{\mathbf{x}} \|\mathbf{b} - A\mathbf{x}\|^2$$

- The vector $\mathbf{x}$ represents the coordinates of the projection $\mathbf{b}'$ in the basis defined by the columns of A.

LS: geometric interpretation



Proof: Derivation of Coordinate Vector $\mathbf{x}$

Geometric Condition: The error vector $(\mathbf{b} - A\mathbf{x})$ represents the shortest path from $\mathbf{b}$ to the hyperplane. Therefore, it must be **orthogonal** to the hyperplane defined by the columns of A.

1. **Orthogonality condition:** The dot product of the error vector with any column of A must be zero.
$$A^T(\mathbf{b} - A\mathbf{x}) = \mathbf{0}$$
2. **Expand the terms:**
$$A^T\mathbf{b} - A^T A\mathbf{x} = \mathbf{0}$$
3. **Rearrange to form the Normal Equation:**
$$A^T A\mathbf{x} = A^T\mathbf{b}$$
4. **Solve for x** (assuming columns of A are linearly independent, $A^T A$ is invertible):
$$\mathbf{x} = (A^T A)^{-1} A^T\mathbf{b}$$

The General Projection Matrix

Using the derived expression for $\mathbf{x}$, we can find the projection $\mathbf{b}'$:

$$\mathbf{b}' = A\mathbf{x} = A\left[(A^T A)^{-1} A^T \mathbf{b}\right]$$

Grouping the matrix terms gives us the definition of the projection matrix P:

$$\mathbf{b}' = \underbrace{A(A^T A)^{-1} A^T}_{P} \mathbf{b}$$

Definition (General Case)

The $n \times n$ projection matrix P is defined as:

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

Case: Orthonormal Columns

Consider the case where the columns of the matrix are orthonormal. We denote this matrix as Q.

- Because the columns are orthonormal, we have $Q^T Q = I_d$ (Identity matrix of size d).
- Substituting Q into the general formula:

$$P = Q(Q^T Q)^{-1} Q^T = Q(I)^{-1} Q^T = QQ^T$$

Definition (Orthonormal Case)

$$P = QQ^T$$

Note: While $Q^T Q = I$, it is generally true that $QQ^T \neq I$ for rectangular matrices where $d < n$.

Idempotent Property

A fundamental property of projection is that projecting a vector already on the plane should not change it. Thus, applying P twice is the same as applying it once ($P^2 = P$).

Proof (Orthonormal Case)

$$P^2 = (QQ^T)(QQ^T)$$

$$P^2 = Q(Q^TQ)Q^T$$

Since $Q^TQ = I$:

$$P^2 = QIQ^T = QQ^T = P$$

General Property: Any symmetric matrix satisfying $P^2 = P$ is a projection matrix.

Using QR Decomposition

Calculating $(A^T A)^{-1}$ can be computationally intensive. We can use **QR Decomposition** ($A = QR$) to simplify the computation.

- Multiplying A by any non-singular matrix does not change the column space, so A and Q share the same projection matrix.
- Therefore, we can compute P using the orthonormal factor Q:

$$P = QQ^T$$

- The projection is computed as $\mathbf{b}' = QQ^T\mathbf{b}$.

Solving for $\mathbf{x}$

Since $\mathbf{b}' = QR\mathbf{x}$, we solve the triangular system:

$$R\mathbf{x} = Q^T\mathbf{b}$$

Orthogonal Complementary Projections

Property

If $P = QQ^T$ is a projection matrix projecting onto a subspace, then the matrix:

$$(I - P)$$

is also a projection matrix.

- It projects onto the **orthogonal complementary vector space** of the column space of A.
- This captures the residual or error component of the vector $\mathbf{b}$.

Summary

1. **Projection Matrix:** Linearly transforms $\mathbf{b}$ to its closest point in the column space of A.
2. **Formula:** $P = A(A^T A)^{-1} A^T$.
3. **Coordinate Vector:** $\mathbf{x} = (A^T A)^{-1} A^T \mathbf{b}$.
4. **Orthonormal Case:** Simplifies to $P = QQ^T$.
5. **Key Properties:** Symmetric ($P^T = P$) and Idempotent ($P^2 = P$).