# ELEC40001

# Orthogonal subspaces, Projection and Least Squares

Xin Wang

October 1, 2020

**Abstract**

Sometimes the matrix system $A\underline{\mathbf{x}} = \underline{\mathbf{b}}$ is inconsistent and does not have a solution. For example, satellite trajectory projections shown in the form of matrix systems sometimes do not have exact solutions. There are certain methods available to find an **approximate** solution when an exact solution does not exist.

# Contents

# 1 Orthogonal Subspaces

As previously mentioned, the dot product measures the amount of one vector goes in the direction of each other. The dot product can be used to find if two vectors are orthogonal to each other. Two vectors are **perpendicular** when dot product is zero:
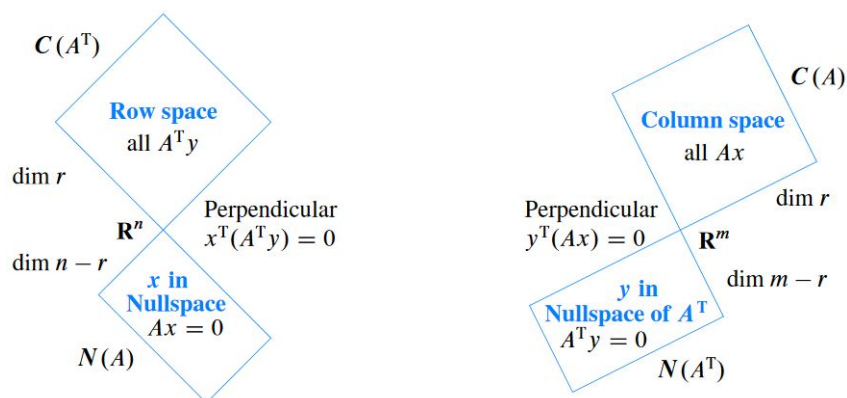
$$\mathbf{v} * \mathbf{w} = \mathbf{v}^T * \mathbf{w} = 0$$

$$\|\mathbf{v}\|^2 * \|\mathbf{w}\|^2 = \|\mathbf{v}+\mathbf{w}\|^2$$

---

Two subspaces V and W of a vector space are **orthogonal** if every vector **v** in set V is perpendicular to every vector **w** in set W.

$$v^T * w = 0$$

for all **v** in V and all **w** in W.

---



The "big picture" of this chapter is that:

- Row space $C(A^T)$ is perpendicular to null space of $N(A)$

- Column space $C(A)$ is perpendicular to null space of $N(A^T)$

Consider a wall and a floor as planes in $\Re^3$. Both are orthogonal to each other but they are not orthogonal subspaces.

- There are diagonal vectors in the wall that are not orthogonal to any vectors in the floor.

- The line of intersection is in both the floor and the wall. A vector can never be orthogonal to itself unless it is the zero vector.

---

For two subspaces to be orthogonal, there must be **only** the zero vector

---

at the intersection:
$$V \cap W = \{\underline{\mathbf{0}}\}$$

## 1.1 Null space perpendicular to row space

The null space of matrix $A$ is the set of all vectors $x$ for which $Ax = 0$. The product of the matrix $A$ and the vector $x$ can be written in terms of the dot product of vectors:

$$A\mathbf{x} = \begin{bmatrix} \mathbf{r}_1 . x \\ \mathbf{r}_2 . x \\ . \\ . \\ . \\ \mathbf{r}_m . x \end{bmatrix}$$

where $r_1, \ldots, r_m$ are the row vectors of $A$.

In order of the definition of null space to be satisfied, $x$ **has to be perpendicular to each of the row vectors of** $A$ **(dot product** $= 0$**) in order for** $Ax = 0$. It follows that the null space of A is the orthogonal complement to the row space.

Every vector $\mathbf{x}$ in null space is perpendicular to every row of $\mathbf{A}$ due to the fundamental definition of the null space where $Ax = 0$. The null space $N(A)$ and row space $C(A^T)$ are orthogonal subspaces of $\Re^n$.

**Example 1**: Show the null space and row space are orthogonal subspaces of $\Re^3$ in matrix $A$:

$$A = \begin{bmatrix} 1 & 2 & 5 \\ 2 & 4 & 10 \end{bmatrix}$$

1. Find the rank of the matrix:

   Matrix has rank of 1 since rows below first row are all multiples of the first row.

   $$\text{Row space} = \mathbf{C}(A^T) = k * \begin{bmatrix} 1 \\ 2 \\ 5 \end{bmatrix}$$

2. Find the dimension of the row space:

   $$\dim \mathbf{C}(A^T) = 1 = \text{rank}(A)$$

3. Find the dimension of the null space:

   $$\dim \mathbf{N}(A) = n - r = \text{rank}(A)$$

4. Find the null space:

- Find the reduced row echelon form:

$$A \sim \begin{bmatrix} 1 & 2 & 5 \\ 0 & 0 & 0 \end{bmatrix} \Rightarrow A\underline{\mathbf{x}} = \underline{\mathbf{0}} \to x_1 + 2x_2 + 5x_3 = 0$$

- Find the special solutions:

$$\begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} \text{ and } \begin{bmatrix} -5 \\ 0 \\ 1 \end{bmatrix}$$

- Find the linear linear combinations that form the null space:

$$\begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}^T * \begin{bmatrix} 1 \\ 2 \\ 5 \end{bmatrix} = 0 \text{ and } \begin{bmatrix} -5 \\ 0 \\ 1 \end{bmatrix}^T * \begin{bmatrix} 1 \\ 2 \\ 5 \end{bmatrix} = 0$$

5. Results show every vector in row space is orthogonal to every vector in the null space.

## 1.2   Left null space orthogonal to column space

Every vector $\mathbf{y}$ in the null space of $A^T$ i.e. left null space is perpendicular to every column of $A$. The left null space $N(A^T)$ and the column space $C(A)$ are orthogonal in $\Re^m$.

# 2   Projection

When $b$ is projected onto a line, its **projection** $p$ is the part of $b$ along that line. If $\mathbf{b}$ is projected onto a plane, $p$ is the part in that plane denoted $P\,\mathbf{b}$
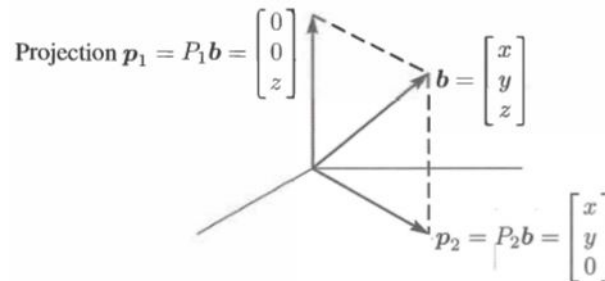


Figure 1: Projection visualised

## 2.1 Properties of $A^T A$

A matrix that was covered previously. This matrix is a square matrix of dimensions $n \times n$ and symmetric. Its null space is of particular interest especially if the square matrix is not invertible.

Given:

$$A\underline{\mathbf{x}} = \underline{\mathbf{b}}$$

multiply by $A^T$ to form a new equation:

$$A^T \times A\underline{\hat{\mathbf{x}}} = A^T \times \underline{\mathbf{b}}$$
$$\Rightarrow A^T A\underline{\hat{\mathbf{x}}} = A^T \underline{\mathbf{b}}$$

Note $\underline{\mathbf{x}}$ is changed into $\underline{\hat{\mathbf{x}}}$ because the result is **an approximation** as the exact solution $\underline{\mathbf{x}}$ cannot be found.

---

The result of $A^T A$ is a square matrix and symmetric but **not always invertible**. The invertibility of $A^T A$ depends on $A$ having independent columns i.e. $N(A) = \underline{\mathbf{0}}$

$$N(A^T A) = N(A) \text{ and } rank(A^T A) = rank(A)$$

---

## 2.2 Projection in 1D

The interest in projection is due to the point given by $\mathbf{p}$, the projection of $\underline{\mathbf{b}}$ onto $\underline{\mathbf{a}}$, this is the point on $\underline{\mathbf{a}}$ nearest to $\underline{\mathbf{b}}$. As mentioned previously if the exact solution $\underline{\mathbf{b}}$ is not available, an approximation can be made by taking the nearest point on $\underline{\mathbf{a}}$ as an approximation and $\underline{\mathbf{e}} \neq \underline{\mathbf{0}}$ is the resulting error.
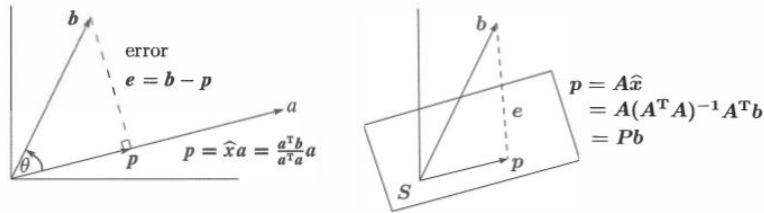


Figure 2: Projection onto a line

The key to projection is te orthogonality property - the point $\mathbf{p}$ on line $\underline{\mathbf{a}}$ that is closest to $\underline{\mathbf{b}}$. As $\underline{\mathbf{a}}$ is only an approximation of $\underline{\mathbf{b}}$, there is an resulting error $\underline{\mathbf{e}}$: $\underline{\mathbf{e}} = \underline{\mathbf{b}} - \hat{\mathbf{x}}\underline{\mathbf{a}}$

Projection is usually defined as: $\mathbf{p} = \lambda\underline{\mathbf{a}}$ with $\lambda$ can be defined as:

$$\lambda = \frac{\underline{\mathbf{a}} * \underline{\mathbf{b}}}{\underline{\mathbf{a}} * \underline{\mathbf{a}}} = \frac{\underline{\mathbf{a}}^T * \underline{\mathbf{b}}}{\underline{\mathbf{a}}^T * \underline{\mathbf{a}}}$$

hence simplify:

$$\mathbf{p} = \frac{\underline{\mathbf{a}}^T * \underline{\mathbf{b}}}{\underline{\mathbf{a}}^T * \underline{\mathbf{a}}} * \underline{\mathbf{a}} = \frac{(\underline{\mathbf{a}}^T * \underline{\mathbf{b}}) * \underline{\mathbf{a}}}{\underline{\mathbf{a}}^T * \underline{\mathbf{a}}} = \frac{\underline{\mathbf{a}} * \underline{\mathbf{a}}^T}{\underline{\mathbf{a}}^T * \underline{\mathbf{a}}} * \underline{\mathbf{b}} = P * \underline{\mathbf{b}}$$

where $P$ is the **Projection Matrix** defined as:

$$P = \frac{\underline{\mathbf{a}} * \underline{\mathbf{a}}^T}{\underline{\mathbf{a}}^T * \underline{\mathbf{a}}}$$

## 2.3   Projection in a subspace (2D or higher)

As stated earlier, if $A\underline{\mathbf{x}} = \underline{\mathbf{b}}$ has no solutions, the closest solution can be found by considering $A\underline{\mathbf{x}}$ in the column space $C(A)$ and find $\mathbf{p}$ which is the closest vector in $C(A)$ to $\underline{\mathbf{b}}$:

$$A\underline{\hat{\mathbf{x}}} = \mathbf{p}$$

where:

- $\mathbf{p}$ is the nearest point to $\underline{\mathbf{b}}$ in the plane

- The error is defined as $\underline{\mathbf{e}} = \underline{\mathbf{b}} - \mathbf{p} = \underline{\mathbf{b}} - A\underline{\hat{\mathbf{x}}}$

---

If $\mathbf{p}$ is in the column space, the error $\underline{\mathbf{e}}$ must be in $N(A^T)$, orthogonal to the column space.

---

The process involving projection is summarised in three steps:
1. Find the vector $\hat{\mathbf{x}}$:

$$A^T * A * \hat{\mathbf{x}} = A^T * \underline{\mathbf{b}}$$

$$\hat{\mathbf{x}} = (A^T * A)^{-1} * A^T * \mathbf{b}$$

- The matrix $A^T * A$ is symmetric
- Invertible if columns are independent
2. Projection $\mathbf{p}$ of $\mathbf{b}$ onto subspace:

$$\mathbf{p} = A * \hat{\mathbf{x}} = A(A^T * A)^{-1} * A^T * \mathbf{b}$$

3. Projection Matrix $\mathbf{P}$:

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

---

## 2.4   Least-squares method - Linear Regression

When $Ax = b$ has no solution, this is usually due to there being too many equations e.g. $A$ has more rows than columns i.e. more equations than unknowns.

When $Ax = b$ has no solution, multiply by $A^T$ and solve the equation:

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

where $\hat{x}$ is the **least-square solution** i.e. where the error $e$ is the smallest.

**Example 1:** Find the quadratic line of best fit with the following points:

- $(x, y) = (-1, 1)$
- $(x, y) = (0, 3)$
- $(x, y) = (1, 4)$
- $(x, y) = (2, 0)$

where $y(x)$ is the trajectory of the object thrown at an angle - a parabolic curve: $y = ax^2 + bx + c$ - subjected to gravity.

1. Formulate the problem - identify the parameters:

   - at $(-1, 1)$: $1 = a - b + c$
   - at $(0, 3)$: $3 = 0a + 0b + c$
   - at $(1, 4)$: $4 = a + b + c$
   - at $(2, 0)$: $0 = 4a + 2b + c$

2. Form matrix notation:

$$A\underline{x} = \underline{b}$$

$$\Rightarrow \begin{bmatrix} 1 & -1 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \\ 4 & 2 & 1 \end{bmatrix} \underline{x} = \begin{bmatrix} 1 \\ 3 \\ 4 \\ 0 \end{bmatrix}$$

3. Verify that it has no solutions: Not full rank.

4. Form equation: $A^T A\underline{\hat{x}} = A^T * \underline{b}$:

$$A^T * A * \underline{\hat{x}} = \begin{bmatrix} 1 & 0 & 1 & 4 \\ -1 & 0 & 1 & 2 \\ 1 & 1 & 1 & 1 \end{bmatrix} * \begin{bmatrix} 1 & -1 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \\ 4 & 2 & 1 \end{bmatrix} * \underline{\hat{x}} = \begin{bmatrix} 1 & 0 & 1 & 4 \\ -1 & 0 & 1 & 2 \\ 1 & 1 & 1 & 1 \end{bmatrix} * \begin{bmatrix} 1 \\ 3 \\ 4 \\ 0 \end{bmatrix} = A^T * \underline{b}$$

5. Simplify:

$$\begin{bmatrix} 18 & 8 & 6 \\ 8 & 6 & 2 \\ 6 & 2 & 4 \end{bmatrix} * \underline{\hat{x}} = \begin{bmatrix} 5 \\ 3 \\ 8 \end{bmatrix}$$

6. Verify the columns are independent: Not multiples of each other.

7. Invert to obtain $\hat{\underline{x}}$:

$$\hat{\underline{x}} = (A^T * A)^{-1} * (A^T * \underline{b}) = \frac{1}{20} \begin{bmatrix} 5 & -5 & -5 \\ -5 & 9 & 3 \\ -5 & 3 & 11 \end{bmatrix} * \begin{bmatrix} 5 \\ 3 \\ 8 \end{bmatrix} = \begin{bmatrix} -1.5 \\ 1.3 \\ 3.6 \end{bmatrix}$$

8. Form the resulting equation:

$$y = -1.5x^2 + 1.3x + 3.6$$

9. $\underline{p}$ can be found and confirmed to be in column space of $A$

10. The error vector $\underline{e}$ can be found and can be confirmed to be in left nullspace $N(A)^T$