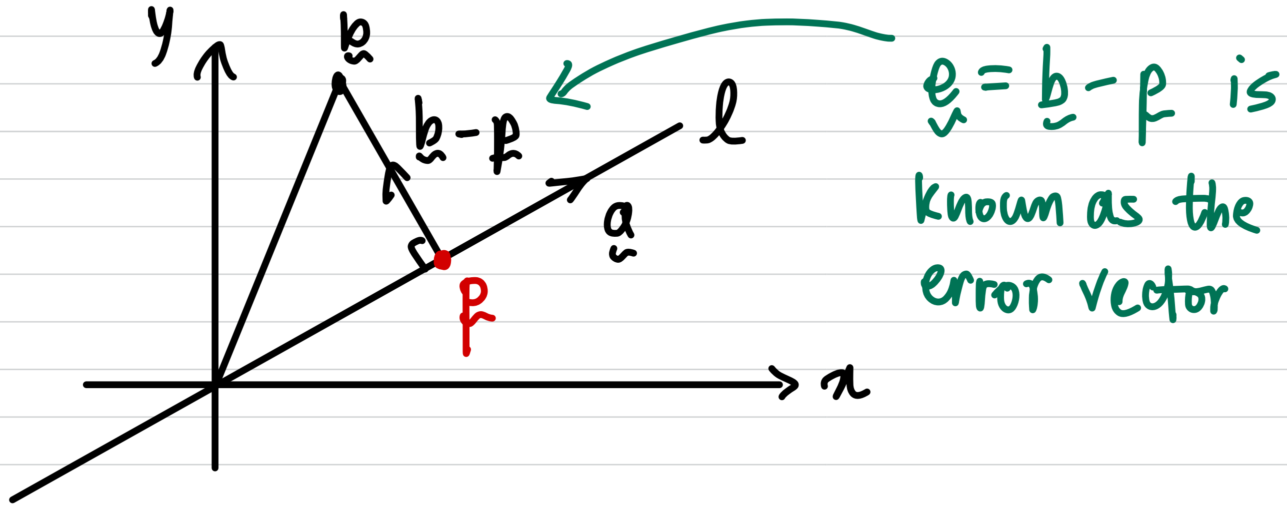
CSD2250/MAT250 Lecture 10 (Fri 10/6)

Topics covered: • Projections onto lines and subspaces

In our course, the main motivation for projections is the <u>least squares approximation</u>, which will be gone through after the recess week.

① Projection of a vector $\underset{\sim}{b}$ onto a line.

Suppose $\underset{\sim}{b} = (b_1, \dots, b_m)$ is a fixed point in $\mathbb{R}^m$ and ℓ is a line passing through the origin in the direction of $\underset{\sim}{a} = (a_1, \dots, a_m)$.



We are interested in finding the point $\underset{\sim}{p}$ on the line ℓ that is <u>closest</u> to $\underset{\sim}{b}$. To find this point $\underset{\sim}{p}$, we note that the line connecting $\underset{\sim}{b}$ and $\underset{\sim}{p}$ **must be <u>orthogonal</u> to the line ℓ!**

* The main observation that is key to finding the point $\underset{\sim}{p}$ is that it **must be parallel to $\underset{\sim}{a}$!**

Thus,

$$\underset{\sim}{p} = \hat{x}\,\underset{\sim}{a}$$

where $\hat{x}$ is a constant we want to find.

* Here, we note that the vector $\underset{\sim}{b} - \underset{\sim}{p}$ is **orthogonal** to $\underset{\sim}{a}$, hence

$$\underset{\sim}{a}^T(\underset{\sim}{b} - \underset{\sim}{p}) = 0$$

$$\Rightarrow \underset{\sim}{a}^T(\underset{\sim}{b} - \hat{x}\,\underset{\sim}{a}) = \underset{\sim}{a}^T\underset{\sim}{b} - \hat{x}\,\underset{\sim}{a}^T\underset{\sim}{a} = 0$$

$$\Rightarrow \hat{x} = \frac{\underset{\sim}{a}^T\underset{\sim}{b}}{\underset{\sim}{a}^T\underset{\sim}{a}}.$$

* Thus, the point $\underset{\sim}{p}$ on the line ℓ that is closest to the point $\underset{\sim}{b}$ is

$$\underset{\sim}{p} = \frac{\underset{\sim}{a}^T\underset{\sim}{b}}{\underset{\sim}{a}^T\underset{\sim}{a}}\,\underset{\sim}{a}.$$

$\underset{\sim}{p}$ is called the **projection of $\underset{\sim}{b}$ onto the line ℓ through $\underset{\sim}{a}$.**

Eg. Find the projection of $\underset{\sim}{b}$ onto the line through $\underset{\sim}{a}$, where

$$\underset{\sim}{b} = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} \quad \text{and} \quad \underset{\sim}{a} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

Also check that the error vector $\underset{\sim}{e} = \underset{\sim}{b} - \underset{\sim}{p}$ is orthogonal to $\underset{\sim}{a}$.

Solution

Solve for $$\hat{x} = \frac{\underset{\sim}{a}^T \underset{\sim}{b}}{\underset{\sim}{a}^T \underset{\sim}{a}} = \frac{1+2+2}{3} = \frac{5}{3}.$$

The projection of $\underset{\sim}{b}$ onto the line through $\underset{\sim}{a}$ is the point

$$\underset{\sim}{p} = \frac{5}{3}\underset{\sim}{a} = \frac{5}{3}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

The error vector $\underset{\sim}{e}$ is

$$\underset{\sim}{e} = \underset{\sim}{b} - \underset{\sim}{p} = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} - \frac{5}{3}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} -\frac{2}{3} \\ \frac{1}{3} \\ \frac{1}{3} \end{bmatrix} = \frac{1}{3}\begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}.$$

Check that $\underset{\sim}{e}$ is orthogonal to $\underset{\sim}{a}$:

$$\underset{\sim}{e}^T \underset{\sim}{a} = \begin{bmatrix} -\frac{2}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

$$= -\frac{2}{3} + \frac{1}{3} + \frac{1}{3} = 0. \quad \checkmark$$

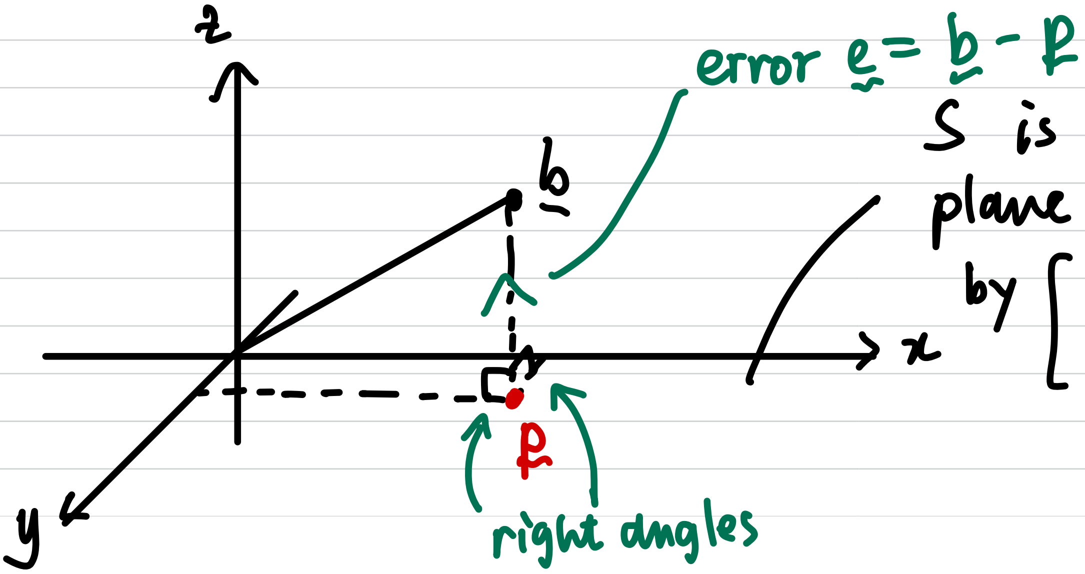
② Projection of a vector $\underset{\sim}{b}$ onto a subspace

let a subspace S be spanned by these vectors

$$\underset{\sim}{a}_1, \ldots, \underset{\sim}{a}_n \in \mathbb{R}^m.$$

* We assume that these vectors are linearly independent. We want to find a point 'on' the subspace S where it is closest to a fixed point $\underset{\sim}{b} \in \mathbb{R}^m$. Note that when $n=1$, then it is the projection of $\underset{\sim}{b}$ onto the line through $\underset{\sim}{a}_1$, thus we consider $n>1$.

Eg.



S is the xy-plane, spanned by $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$.

We want to find the point $\underset{\sim}{p}$ on the subspace S that is closest to $\underset{\sim}{b}$. Since $\underset{\sim}{p}$ is on S, then $\underset{\sim}{p}$ can be written as a linear combination of $\underset{\sim}{a}_1, \dots, \underset{\sim}{a}_n$:

$$\underset{\sim}{p} = \hat{x}_1 \underset{\sim}{a}_1 + \dots + \hat{x}_n \underset{\sim}{a}_n$$

where $\hat{x}_1, \dots, \hat{x}_n$ are to be determined.

Now, $\underset{\sim}{b} - \underset{\sim}{p}$ must be orthogonal to the subspace S, and hence must be orthogonal to all of $\underset{\sim}{a}_1, \dots, \underset{\sim}{a}_n$. Hence

$$\underset{\sim}{a}_1^T (\underset{\sim}{b} - \underset{\sim}{p}) = 0,$$

$$\vdots$$

$$\underset{\sim}{a}_n^T (\underset{\sim}{b} - \underset{\sim}{p}) = 0.$$

* If we let A be the matrix whose columns are $\underset{\sim}{a}_1, \dots, \underset{\sim}{a}_n$ and $\hat{x} = (\hat{x}_1, \dots, \hat{x}_n)$, then the equations above can be written as

$$A^T (\underset{\sim}{b} - A\hat{x}) = \underset{\sim}{0}$$

$$\Leftrightarrow \quad A^T A \hat{x} = A^T \underset{\sim}{b}.$$

★ **A^TA is invertible if and only if A has linearly independent columns.**

<u>Proof</u> We show that $N(A^TA) = N(A)$.

To show this, we show <u>two</u> things:

① If $\underset{\sim}{x} \in N(A^TA)$, then $\underset{\sim}{x} \in N(A)$.

② If $\underset{\sim}{x} \in N(A)$, then $\underset{\sim}{x} \in N(A^TA)$.

① Let $\underset{\sim}{x} \in N(A^TA)$. Then $A^TA\underset{\sim}{x} = \underset{\sim}{0}$.

Multiplying this equation by $\underset{\sim}{x}^T$ gives

$$\underset{\sim}{x}^T A^T A \underset{\sim}{x} = \underset{\sim}{0}$$

$$\Rightarrow \quad (A\underset{\sim}{x})^T(A\underset{\sim}{x}) = 0 \Rightarrow \|Ax\|^2 = 0$$

$$\Rightarrow \|A\underset{\sim}{x}\|^2 = 0$$

Since the length of $A\underset{\sim}{x}$ is 0, then $A\underset{\sim}{x}$ must be the zero vector, and hence

$$A\underset{\sim}{x} = \underset{\sim}{0} \Rightarrow \underset{\sim}{x} \in N(A).$$

② Let $\underset{\sim}{x} \in N(A)$, then $A\underset{\sim}{x} = \underset{\sim}{0}$. Multiplying this equation by A^T gives $A^TA\underset{\sim}{x} = A^T\underset{\sim}{0} = \underset{\sim}{0}$.

Hence $A^TA\underset{\sim}{x} = \underset{\sim}{0} \Rightarrow \underset{\sim}{x} \in N(A^TA)$.

Therefore, the nullspace of A is the same as the nullspace of A^TA. Hence, if the columns of A are linearly independent, then the only solution to $A\underset{\sim}{x}=\underset{\sim}{0}$ is $\underset{\sim}{x}=\underset{\sim}{0}$ and by the statement above, the only solution to $A^TA\underset{\sim}{x}=\underset{\sim}{0}$ is $\underset{\sim}{x}=\underset{\sim}{0}$. Since A^TA is square, it must mean that A^TA is invertible!

---

*Previously, we assumed that

$$\underset{\sim}{a}_1, \dots, \underset{\sim}{a}_n$$

are linearly independent. Hence the columns of A are linearly independent, and thus, A^TA is invertible. We can then solve for $\hat{x}$:

$$\boxed{\hat{x} = (A^TA)^{-1}A^T\underset{\sim}{b}},$$

and the point $\underset{\sim}{p}$ on the subspace S that is closest to $\underset{\sim}{b}$ is

$$* \quad \boxed{\underset{\sim}{p} = A(A^TA)^{-1}A^T\underset{\sim}{b}.}$$

*Note that $(A^TA)^{-1} \neq A^{-1}(A^T)^{-1}$ because A is not necessarily a square matrix!

E.g Let S be the subspace spanned by the vectors

$$\underset{\sim}{a_1} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \text{ and } \underset{\sim}{a_2} = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}.$$

Find the shortest distance from the point $\underset{\sim}{b} = \begin{bmatrix} 6 \\ 0 \\ 0 \end{bmatrix}$ to S.

Solution

(1) Check that the columns of A are linearly independent, where

$$A = \begin{bmatrix} | & | \\ \underset{\sim}{a_1} & \underset{\sim}{a_2} \\ | & | \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix}.$$

The rref of A is $\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$. The no. of pivots $= 2 =$ no. of columns, hence the columns of A are linearly independent.

(2) Find the projection $\underset{\sim}{p}$ of $\underset{\sim}{b}$ onto S.

We can find $\hat{x}$ by either computing $(A^T A)^{-1}$ or solving $A^T A \hat{x} = A^T \underset{\sim}{b}$. We do the latter.

$$A^TA = \begin{bmatrix} 3 & 3 \\ 3 & 5 \end{bmatrix}$$

$$\underbrace{\begin{bmatrix} 3 & 3 \\ 3 & 5 \end{bmatrix}}_{A^TA} \underbrace{\begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \end{bmatrix}}_{\hat{x}} = \underbrace{\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \end{bmatrix}}_{A^T} \underbrace{\begin{bmatrix} 6 \\ 0 \\ 0 \end{bmatrix}}_{\underset{\sim}{b}}$$

$$\Rightarrow \begin{bmatrix} 3 & 3 \\ 3 & 5 \end{bmatrix} \begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \end{bmatrix} = \begin{bmatrix} 6 \\ 0 \end{bmatrix}$$

$R_2 - R_1$

$$\Rightarrow \begin{bmatrix} 3 & 3 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \end{bmatrix} = \begin{bmatrix} 6 \\ -6 \end{bmatrix}$$

$$\Rightarrow \hat{x}_2 = -3, \quad \hat{x}_1 = 5$$

$$\therefore \hat{x} = \begin{bmatrix} 5 \\ -3 \end{bmatrix}$$

The projection $\underset{\sim}{p} = A\hat{x} = 5\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - 3\begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}$

$$= \begin{bmatrix} 5 \\ 2 \\ -1 \end{bmatrix}.$$

(3) The shortest distance from $\underset{\sim}{b}$ to the subspace S is the distance btw $\underset{\sim}{b}$ and $\underset{\sim}{p}$.

$$\|\underset{\sim}{e}\| = \|\underset{\sim}{b} - \underset{\sim}{p}\| = \sqrt{1^2 + (-2)^2 + 1^2} = \sqrt{6}.$$

Exercises

(1) Project $\underset{\sim}{b} = \begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix}$ onto the line through $\underset{\sim}{a} = \begin{bmatrix} -1 \\ -3 \\ -1 \end{bmatrix}$. Check that $\underset{\sim}{e}$ is orthogonal to $\underset{\sim}{a}$.

(2) Project $\underset{\sim}{b}$ onto the columnspace of A by solving $A^T A \hat{x} = A^T \underset{\sim}{b}$ and $\underset{\sim}{p} = A\hat{x}$:

(a) $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$, $\underset{\sim}{b} = \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix}$,

(b) $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 0 & 1 \end{bmatrix}$, $\underset{\sim}{b} = \begin{bmatrix} 4 \\ 4 \\ 6 \end{bmatrix}$.

Check that $\underset{\sim}{e} = \underset{\sim}{b} - \underset{\sim}{p}$ is orthogonal to $\underset{\sim}{a}$.