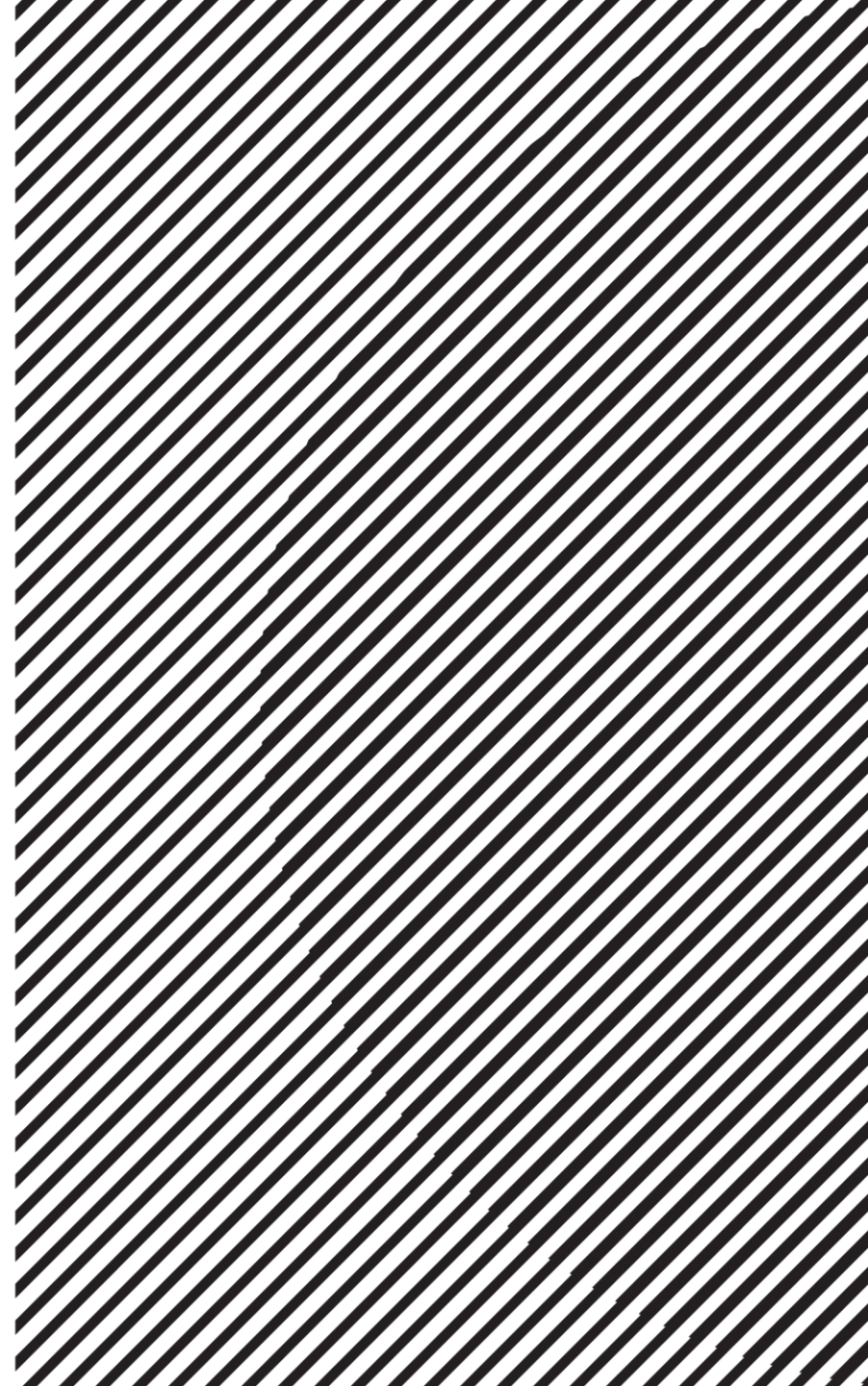

Linear Algebra

주재걸
고려대학교 컴퓨터학과



Orthogonal Projection Perspective

- Back to the case of invertible $C = A^T A$, consider the orthogonal projection of \mathbf{b} onto $\text{Col } A$ as

$$\hat{\mathbf{b}} = f(\mathbf{b}) = A\hat{\mathbf{x}} = A \underbrace{(A^T A)^{-1} A^T}_{\hat{\mathbf{X}}} \mathbf{b}$$



Orthogonal and Orthonormal Sets

- **Definition:** A set of vectors $\{\mathbf{u}_1, \dots, \mathbf{u}_p\}$ in \mathbb{R}^n is an **orthogonal set** if each pair of distinct vectors from the set is orthogonal. That is, if $\mathbf{u}_i \cdot \mathbf{u}_j = 0$ whenever $i \neq j$.
- **Definition:** A set of vectors $\{\mathbf{u}_1, \dots, \mathbf{u}_p\}$ in \mathbb{R}^n is an **orthonormal set** if it is an orthogonal set of **unit vectors**.
- Is an orthogonal (or orthonormal) set also a linearly independent set? What about its converse?

Yes

No



Orthogonal and Orthonormal Basis

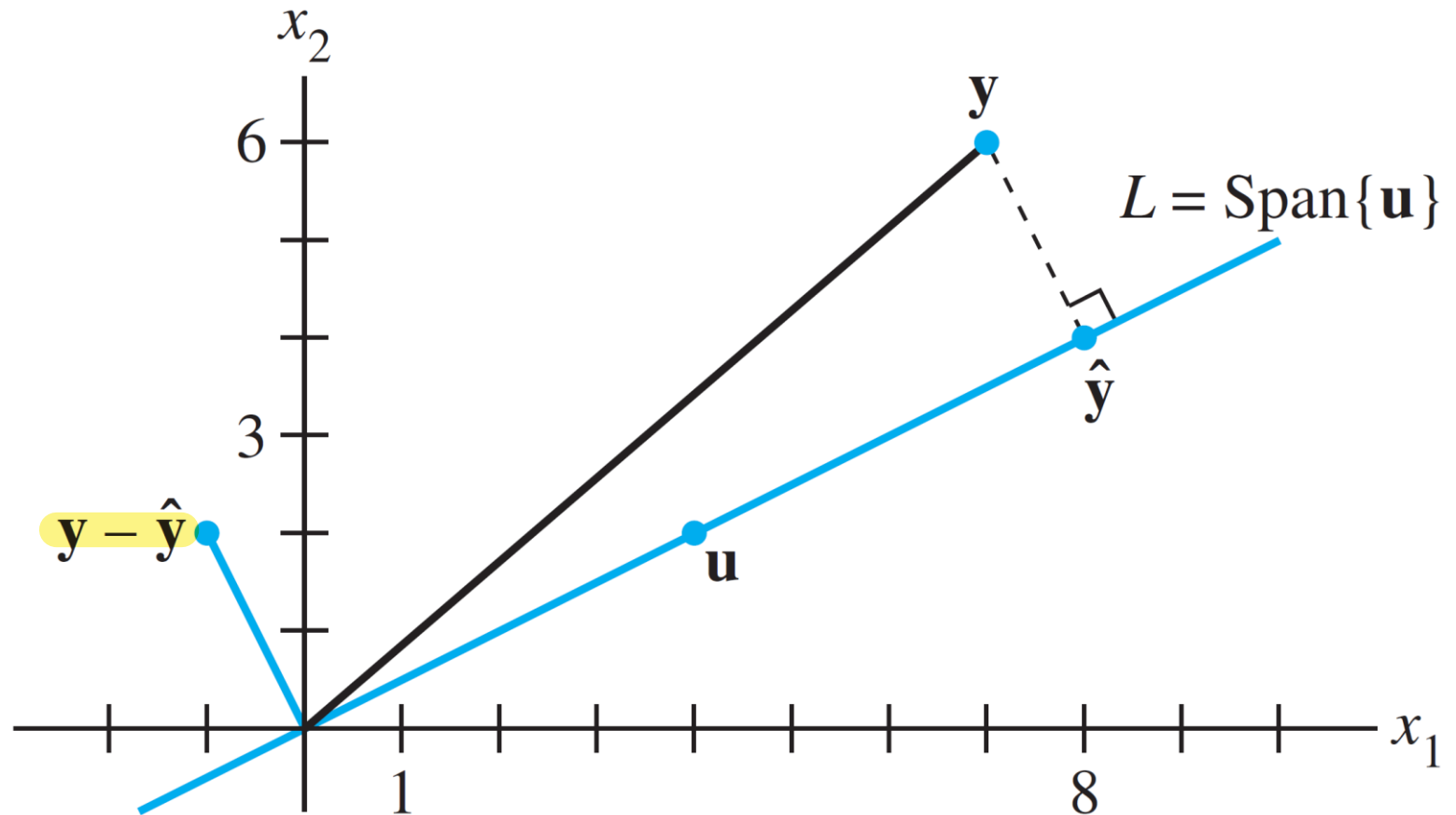
- Consider basis $\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ of a p -dimensional subspace W in \mathbb{R}^n .
- Can we make it as an orthogonal (or orthonormal) basis?
 - Yes, it can be done by Gram–Schmidt process. → QR factorization.
- Given the orthogonal basis $\{\mathbf{u}_1, \dots, \mathbf{u}_p\}$ of W ,
let's compute the orthogonal projection of $\mathbf{y} \in \mathbb{R}^n$ onto W .

Orthogonal Projection \hat{y} of y onto Line

- Consider the orthogonal projection \hat{y} of y onto one-dimensional subspace L .

- $\hat{y} = \text{proj}_L y = \frac{y \cdot u}{u \cdot u} u$

- If u is a unit vector,
 $\hat{y} = \text{proj}_L y = (y \cdot u) u$



Orthogonal Projection $\hat{\mathbf{y}}$ of \mathbf{y} onto Plane

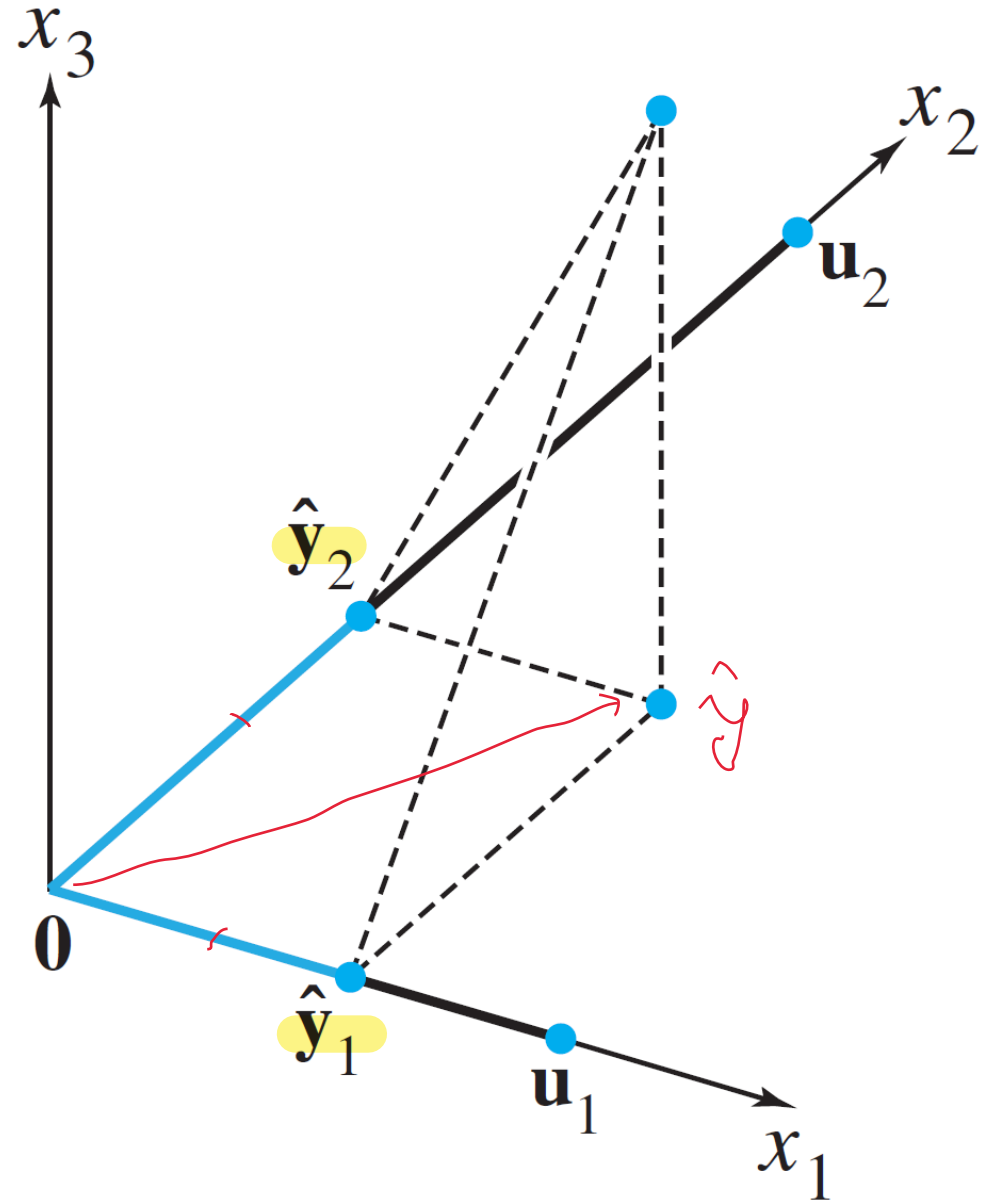
- Consider the orthogonal projection $\hat{\mathbf{y}}$ of \mathbf{y} onto two-dimensional subspace W

- $\hat{\mathbf{y}} = \text{proj}_L \mathbf{y} = \frac{\mathbf{y} \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1} \mathbf{u}_1 + \frac{\mathbf{y} \cdot \mathbf{u}_2}{\mathbf{u}_2 \cdot \mathbf{u}_2} \mathbf{u}_2$

- If \mathbf{u}_1 and \mathbf{u}_2 are unit vectors,

$$\hat{\mathbf{y}} = \text{proj}_L \mathbf{y} = (\mathbf{y} \cdot \mathbf{u}_1) \mathbf{u}_1 + (\mathbf{y} \cdot \mathbf{u}_2) \mathbf{u}_2$$

- Projection is done independently on each orthogonal basis vector.



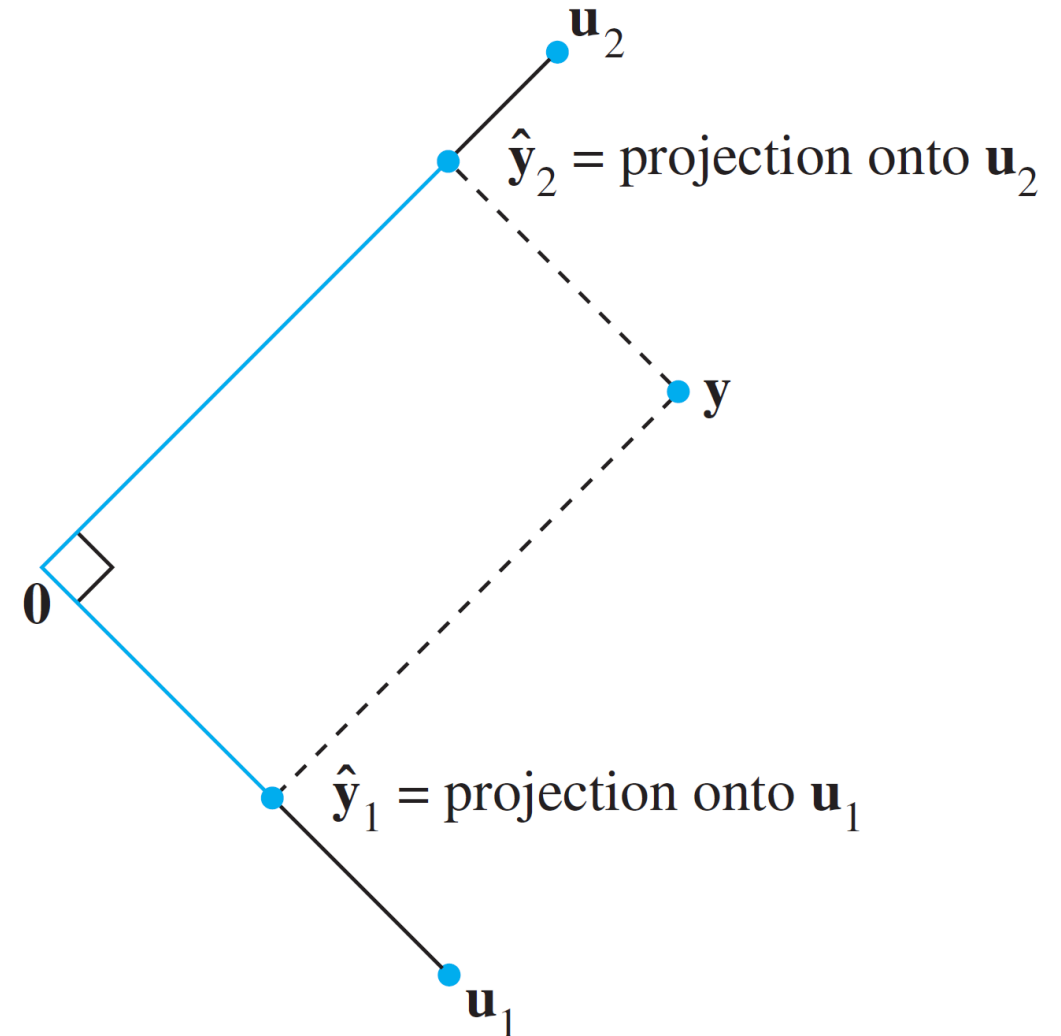
Orthogonal Projection when $\mathbf{y} \in W$

- Consider the orthogonal projection $\hat{\mathbf{y}}$ of \mathbf{y} onto two-dimensional subspace W , where $\mathbf{y} \in W$

- $\hat{\mathbf{y}} = \text{proj}_L \mathbf{y} = \mathbf{y} = \frac{\mathbf{y} \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1} \mathbf{u}_1 + \frac{\mathbf{y} \cdot \mathbf{u}_2}{\mathbf{u}_2 \cdot \mathbf{u}_2} \mathbf{u}_2$

- If \mathbf{u}_1 and \mathbf{u}_2 are unit vectors,
 $\hat{\mathbf{y}} = \mathbf{y} = (\mathbf{y} \cdot \mathbf{u}_1) \mathbf{u}_1 + (\mathbf{y} \cdot \mathbf{u}_2) \mathbf{u}_2$

- The solution is the same as before.
Why?



Transformation: Orthogonal Projection

- Consider a transformation of orthogonal projection $\hat{\mathbf{b}}$ of \mathbf{b} , given **orthonormal** basis $\{\mathbf{u}_1, \mathbf{u}_2\}$ of a subspace W :

$$\hat{\mathbf{b}} = f(\mathbf{b}) = (\mathbf{b} \cdot \mathbf{u}_1)\mathbf{u}_1 + (\mathbf{b} \cdot \mathbf{u}_2)\mathbf{u}_2$$

$$= (\mathbf{u}_1^T \mathbf{b})\mathbf{u}_1 + (\mathbf{u}_2^T \mathbf{b})\mathbf{u}_2$$

$$= \mathbf{u}_1(\mathbf{u}_1^T \mathbf{b}) + \mathbf{u}_2(\mathbf{u}_2^T \mathbf{b})$$

$$= (\mathbf{u}_1 \mathbf{u}_1^T) \mathbf{b} + (\mathbf{u}_2 \mathbf{u}_2^T) \mathbf{b}$$

$$= (\mathbf{u}_1 \mathbf{u}_1^T + \mathbf{u}_2 \mathbf{u}_2^T) \mathbf{b}$$

$$= [\mathbf{u}_1 \quad \mathbf{u}_2] \begin{bmatrix} \mathbf{u}_1^T \\ \mathbf{u}_2^T \end{bmatrix} \mathbf{b} = U U^T \mathbf{b} \Rightarrow \text{linear transformation!}$$



Orthogonal Projection Perspective

- Let's verify the following, when $A = U = [\mathbf{u}_1 \quad \mathbf{u}_2]$ has orthonormal columns:

Back to the case of invertible $C = A^T A$, consider the orthogonal projection of \mathbf{b} onto $\text{Col } A$ as

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

- $C = A^T A = \begin{bmatrix} \mathbf{u}_1^T \\ \mathbf{u}_2^T \end{bmatrix} [\mathbf{u}_1 \quad \mathbf{u}_2] = I$. Thus,

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