Analytic Geometry 2

Liang Zheng
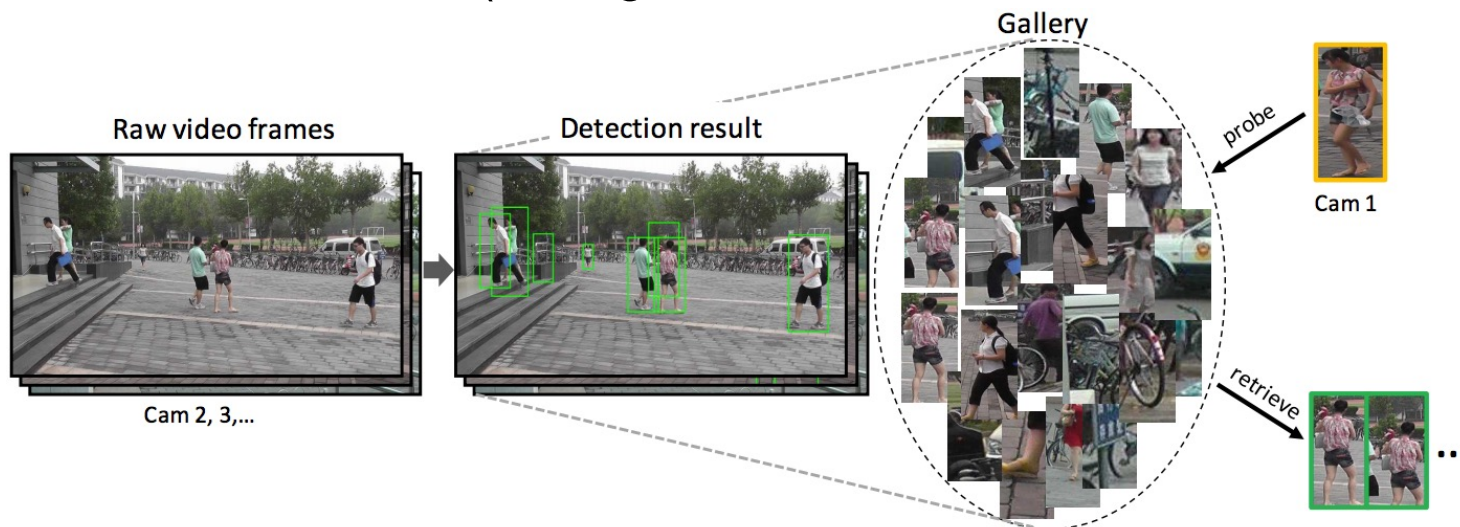
Australian National University

liang.zheng@anu.edu.au

Norm-Aware Embedding for Efficient Person Search (Chen et al., CVPR 2020)

Nanjing University of Science and Technology
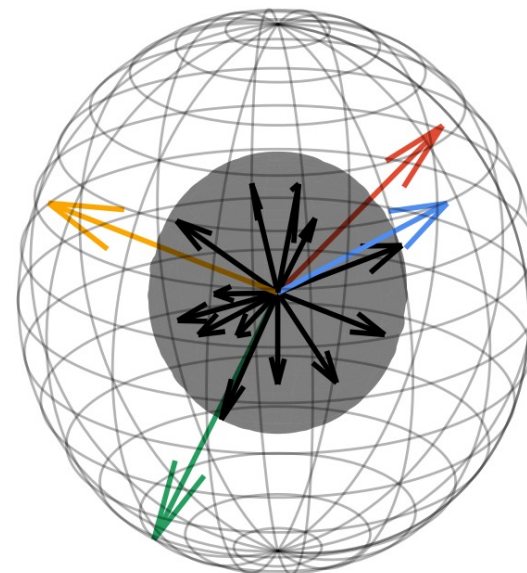Max Planck Institute for Informatics

- Person search (Zheng et al., CVPR 2017, Xiao et al. CVPR 2017)

What is the meaning of norm and angle?

Norm can differentiate person from background

Angle can differentiate different persons

3.8 Orthogonal Projections

- High-dimensional data.

- only a few dimensions contain most information

- When we compress or visualize high-dimensional data, we will lose information.

- To minimize this compression loss, we want to find the most informative dimensions in the data.

- Orthogonal projections of high-dimensional data retain as much information as possible

Orthogonal projection (orange dots) of a two-dimensional dataset (blue dots) onto a one-dimensional subspace (straight line)
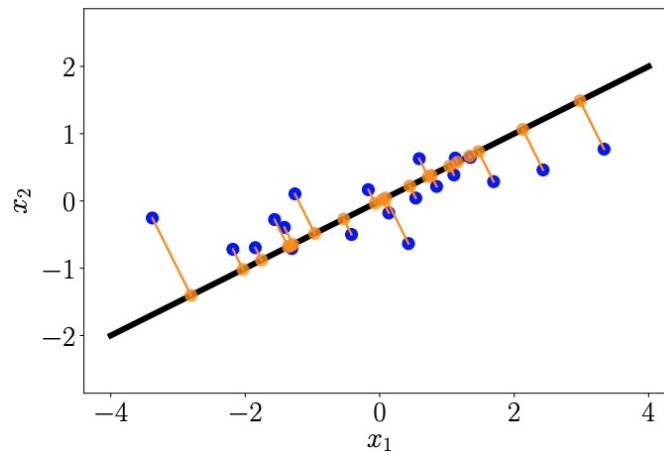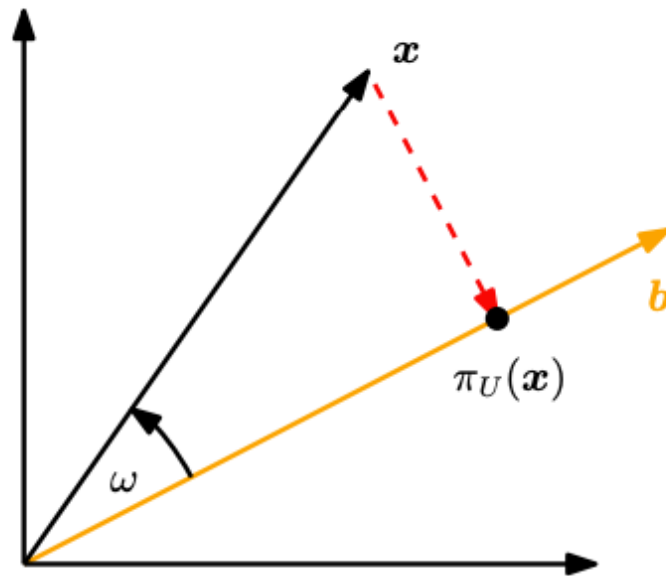
3.8 Orthogonal Projections

- Let V be a vector space and $U \subseteq V$ a subspace of V. A linear mapping $\pi: V \to V$ is called a projection if $\pi^2 = \pi \circ \pi = \pi$.

- Linear mappings can be expressed by transformation matrices.

- The projection matrices $\boldsymbol{P}_\pi$ has the property $\boldsymbol{P}_\pi^2 = \boldsymbol{P}_\pi$.

3.8.1 Projection onto One-Dimensional Subspaces (Lines)

- Assume we are given a line (one-dimensional subspace) through the origin with basis vector $b \in \mathbb{R}^n$.

- When we project $x \in \mathbb{R}^n$ onto U, we seek the vector $\pi_U(x)$ that is closest to x.

(a) Projection of $x \in \mathbb{R}^2$ onto a subspace U with basis vector b.
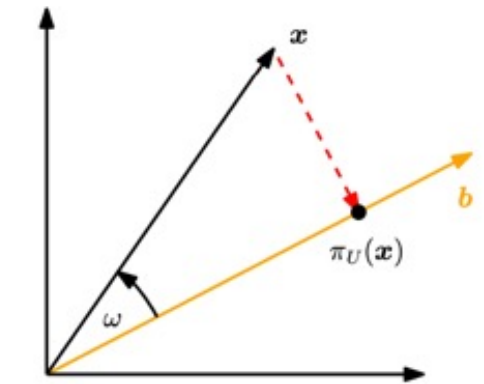
- The projection $\pi_U(x)$ should be closest to x.

  ➡ $\|x - \pi_U(x)\|$ is minimal.

  ➡ $\pi_U(x) - x$ is orthogonal to U, which is spanned by b.

  ➡ $\langle \pi_U(x) - x, b \rangle = 0$

- $\pi_U(x)$ is an element of U spanned by b.

  ➡ $\pi_U(x) = \lambda b$ for some $\lambda \in \mathbb{R}$.
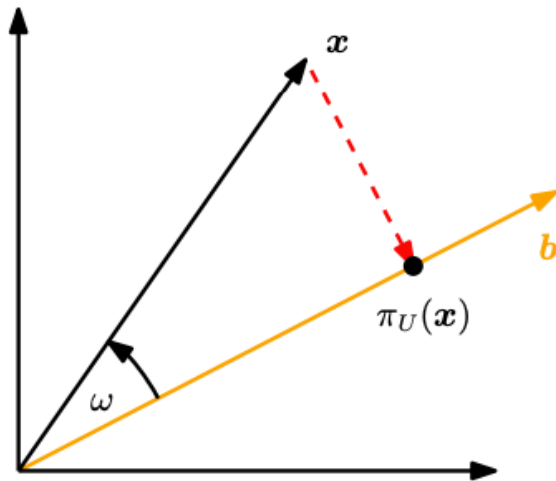
(a) Projection of $x \in \mathbb{R}^2$ onto a subspace U with basis vector b.

How to determine λ, $\pi_U(x)$ and the projection matrix P_π?

1. Finding the coordinate λ

(a) Projection of $x \in \mathbb{R}^2$ onto a subspace U with basis vector b.

- The orthogonality condition

$$\langle x - \pi_U(x), b \rangle = 0 \overset{\pi_U(x)=\lambda b}{\Longleftrightarrow} \langle x - \lambda b, b \rangle = 0$$

- We use the bilinearity of inner product

$$\langle x, b \rangle - \lambda \langle b, b \rangle = 0 \Longleftrightarrow \lambda = \frac{\langle x, b \rangle}{\langle b, b \rangle} = \frac{\langle b, x \rangle}{\|b\|^2}$$

inner products are symmetric

- If we choose $\langle \cdot, \cdot \rangle$ to be the dot product, we obtain

$$\lambda = \frac{b^\top x}{b^\top b} = \frac{b^\top x}{\|b\|^2}$$

- If $\|b\| = 1$ then λ is given by $b^\top x$.

2. Finding the projection point $\pi_U(x) \in U$

- Since $\pi_U(x) = \lambda b$, we immediately obtain

$$\pi_U(x) = \lambda b = \frac{<x, b>}{\|b\|^2} b = \frac{b^\top x}{\|b\|^2} b$$

Assuming dot product

We can also compute the length of $\pi_U(x)$ as

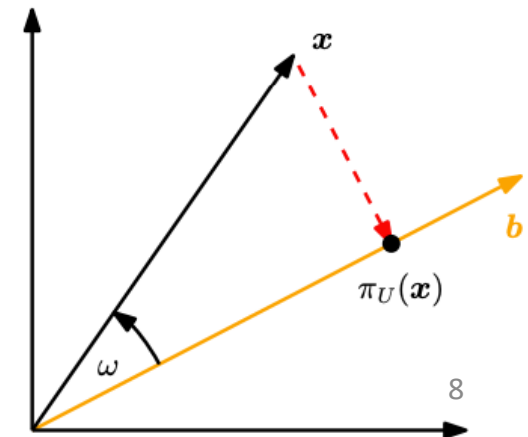$$\|\pi_U(x)\| = \|\lambda b\| = |\lambda|\, \|b\|$$

Hence, our projection is of length $|\lambda|$ times the length of b.

- Using the dot product as an inner product, we get

$$\pi_U(x) = \frac{b^\top x}{\|b\|^2} b = \frac{b^\top x}{\|b\|\|x\|} \frac{\|x\|}{\|b\|} b = \cos\omega \frac{\|x\|}{\|b\|} b$$

$$\implies \|\pi_U(x)\| = \left\| \cos\omega \frac{\|x\|}{\|b\|} b \right\| = |\cos\omega| \frac{\|x\|}{\|b\|} \|b\| = |\cos\omega|\|x\|$$

ω is the angle between x and b. This equation should look familiar from trigonometry.

3. Finding the projection matrix P_π
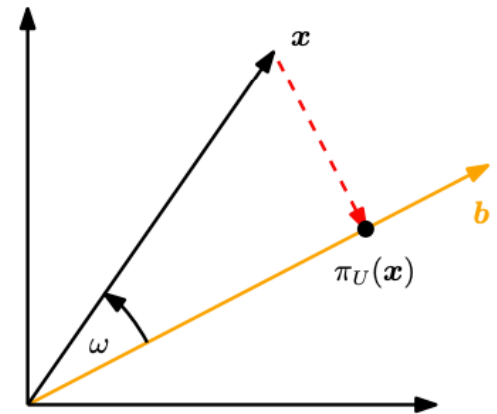
- A projection is a linear mapping

- There exists a projection matrix P_π such that $\pi_U(x) = P_\pi x$

- With the dot product as inner product and

$$\pi_U(x) = \lambda b = b\lambda = b\frac{b^\top x}{\|b\|^2} = \frac{bb^\top}{\|b\|^2}x$$

  we immediately see that

$$P_\pi = \frac{bb^\top}{\|b\|^2}$$

- Note that $b\,b^\top$ (and, consequently, P_π) is a symmetric matrix (of rank 1), and $\|b\|^2 = \langle b, b \rangle$ is a scalar.

Example (Projection onto a Line)

- Find the projection matrix $\boldsymbol{P}_\pi$ onto the line through the origin spanned by $\boldsymbol{b} = \begin{bmatrix} 1 & -1 \end{bmatrix}^\top$.

$$\boldsymbol{P}_\pi = \frac{\boldsymbol{b}\boldsymbol{b}^\top}{\|\boldsymbol{b}\|^2} = \frac{1}{2}\begin{bmatrix} 1 \\ -1 \end{bmatrix}\begin{bmatrix} 1 & -1 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$$

- We choose a particular $\boldsymbol{x}$ and see whether its projection lies in the subspace spanned by $\boldsymbol{b}$. For $\boldsymbol{x} = \begin{bmatrix} 3 & 5 \end{bmatrix}^\top$, the projection is

$$\pi_U(\boldsymbol{x}) = \boldsymbol{P}_\pi \boldsymbol{x} = \frac{1}{2}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} 3 \\ 5 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} -2 \\ 2 \end{bmatrix} \in span\begin{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} \end{bmatrix}$$

- Further application of $\boldsymbol{P}_\pi$ to $\pi_U(\boldsymbol{x})$ does not change anything, i.e., $\boldsymbol{P}_\pi \pi_U(\boldsymbol{x}) = \pi_U(\boldsymbol{x})$. This is expected because according to the definition of Projection, we know that a projection matrix $\boldsymbol{P}_\pi$ satisfies $\boldsymbol{P}_\pi^2 \boldsymbol{x} = \boldsymbol{P}_\pi \boldsymbol{x}$ for all $\boldsymbol{x}$.

3.8.2 Projection onto General Subspaces

- We look at orthogonal projections of vectors $x \in \mathbb{R}^n$ onto lower-dimensional subspaces $U \subseteq \mathbb{R}^n$ with $\dim(U) = m \geq 1$.
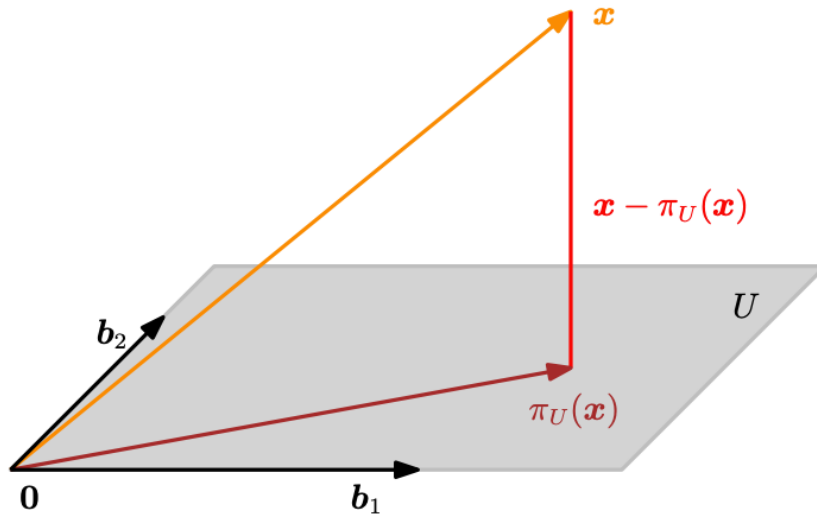
Projecting $x \in \mathbb{R}^3$ onto a two-dimensional subspace

- Assume $(b_1, \cdots, b_m)$ is a basis of U.

- The projection $\pi_U(x)$ is a component of U.

$$\Longrightarrow \quad \pi_U(x) = \sum_{i=1}^{m} \lambda_i b_i$$

- How to determine λ_i, $\pi_U(x)$ and P_π?

1. Find the coordinates $\lambda_i, \cdots, \lambda_m$

- The linear combination

$$\pi_U(x) = \sum_{i=1}^{m} \lambda_i b_i = B\lambda \qquad B = [b_1, \ldots, b_m] \in \mathbb{R}^{n \times m}, \lambda = [\lambda_1, \ldots, \lambda_m]^\top \in \mathbb{R}^m$$

should be closest to $x \in \mathbb{R}^n$,

➡ the vector connecting $\pi_U(x) \in U$ and $x \in \mathbb{R}^n$ must be orthogonal to all basis vectors of U.

➡ We obtain m simultaneous conditions (using the dot product)

$$\langle b_1, x - \pi_U(x) \rangle = b_1^\top (x - \pi_U(x)) = 0$$
$$\vdots$$
$$\langle b_m, x - \pi_U(x) \rangle = b_m^\top (x - \pi_U(x)) = 0$$

with $\pi_U(x) = B\lambda$, we re-write the above as

$$b_1^\top (x - B\lambda) = 0$$
$$\vdots$$
$$b_m^\top (x - B\lambda) = 0$$

such that we obtain a homogeneous linear equation system

$$\begin{bmatrix} b_1^\top \\ \vdots \\ b_m^\top \end{bmatrix} [x - B\lambda] = 0 \iff B^\top (x - B\lambda) = 0 \iff B^\top B\lambda = B^\top x.$$

12
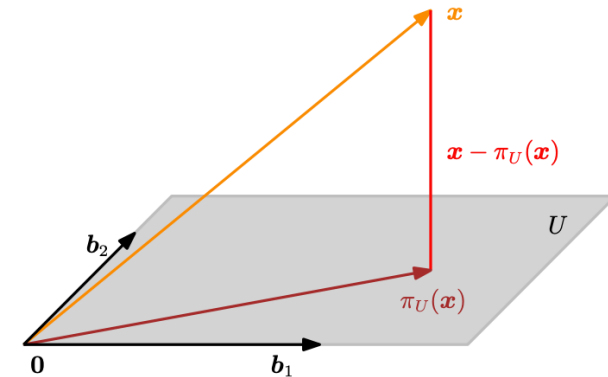
1. Find the coordinates $\lambda_i, \cdots, \lambda_m$

$$B^{\mathrm{T}}B\lambda = B^{\mathrm{T}}x.$$

- $b_1, \ldots, b_m$ are a basis of U, so they are linearly independent.

$$\longrightarrow \quad r(B^{\mathrm{T}}B) = r(B) = m$$

This allows us to solve λ

$$\lambda = (B^{\mathrm{T}}B)^{-1}B^{\mathrm{T}}x$$

- Recall: the matrix $(B^{\mathrm{T}}B)^{-1}B^{\mathrm{T}}$ is also called the pseudo-inverse of B.

2. Find the projection $\pi_U(x) \in U$. We already established that $\pi_U(x) = B\lambda$. Therefore, we calculate $\pi_U(x)$ as

$$\pi_U(x) = B(B^{\mathrm{T}}B)^{-1}B^{\mathrm{T}}x$$

3. Find the projection matrix P_π

- We have $P_\pi x = \pi_U(x)$

- From step 2, we have

$$\pi_U(x) = B(B^\mathrm{T}B)^{-1}B^\mathrm{T}x$$

- We can immediately see that

$$P_\pi = B(B^\mathrm{T}B)^{-1}B^\mathrm{T}$$

- If $\dim(U) = 1$, i.e., projecting onto a 1-dim subspace, we have $B^\mathrm{T}B$ is a scalar. We can reduce

$$P_\pi = B(B^\mathrm{T}B)^{-1}B^\mathrm{T}$$

into

$$P_\pi = \frac{bb^\top}{\|b\|^2}$$

which is exactly the projection matrix in the 1-D case.

Example - Projection onto a Two-dimensional Subspace

- For a subspace $U = \mathrm{span}[\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}] \subseteq \mathbb{R}^3$, and $\boldsymbol{x} = \begin{bmatrix} 6 \\ 0 \\ 0 \end{bmatrix} \in \mathbb{R}^3$, find the coordinates $\boldsymbol{\lambda}$ of $\pi_U(\boldsymbol{x})$ in terms of U, the projection point $\pi_U(\boldsymbol{x})$ and the projection matrix $\boldsymbol{P}_\pi$.

- Solution

- First, the generating set of U is a basis (linear independence) and write the basis vectors of U into a matrix $\boldsymbol{B} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix}$.

- Second, we compute the matrix $\boldsymbol{B}^\mathrm{T}\boldsymbol{B}$ and the vector $\boldsymbol{B}^\mathrm{T}\boldsymbol{x}$ as

$$\boldsymbol{B}^\mathrm{T}\boldsymbol{B} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 3 & 3 \\ 3 & 5 \end{bmatrix}, \boldsymbol{B}^\mathrm{T}\boldsymbol{x} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} 6 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 6 \\ 0 \end{bmatrix}$$

- Third, we solve the normal equation $\boldsymbol{B}^\mathrm{T}\boldsymbol{B}\boldsymbol{\lambda} = \boldsymbol{B}^\mathrm{T}\boldsymbol{x}$ to find $\boldsymbol{\lambda}$:

$$\begin{bmatrix} 3 & 3 \\ 3 & 5 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix} = \begin{bmatrix} 6 \\ 0 \end{bmatrix} \Leftrightarrow \boldsymbol{\lambda} = \begin{bmatrix} 5 \\ -3 \end{bmatrix}$$

Example - Projection onto a Two-dimensional Subspace

- Fourth, the projection point $\pi_U(x)$ of x onto U, i.e., into the column space of B, can be directly computed via

$$\pi_U(x) = B\lambda = \begin{bmatrix} 5 \\ 2 \\ -1 \end{bmatrix}$$

- The corresponding projection error is the norm of the difference between the original vector and its projection onto U, i.e.,

$$\|x - \pi_U(x)\| = \left\|[1 \quad -2 \quad 1]^{\mathrm{T}}\right\| = \sqrt{6}$$

- Fifth, the projection matrix (for any $x \in \mathbb{R}^3$) is given by

$$P_\pi = B(B^{\mathrm{T}}B)^{-1}B^{\mathrm{T}} = \frac{1}{6}\begin{bmatrix} 5 & 2 & -1 \\ 2 & 2 & 2 \\ -1 & 2 & 5 \end{bmatrix}$$

Things to note

- $\pi_U(x)$ is still in $\mathbb{R}^3$, although it lies in a 2-dim subspace $U \subseteq \mathbb{R}^3$
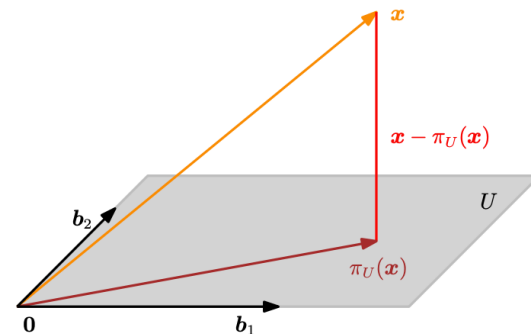
- We can find approximate solutions to unsolvable linear equation systems $Ax = b$ using projections.

- The idea is to find the vector in the subspace spanned by the columns of A that is closest to b, i.e., we compute the orthogonal projection of b onto the subspace spanned by the columns of A. --- least-squares solution

- If B is an orthonomal basis (ONB), i.e., $B^{\mathrm{T}}B = I$, we have

$$\pi_U(x) = BB^{\mathrm{T}}x$$

$$\lambda = B^{\mathrm{T}}x$$

$$\langle b_i, b_j \rangle = 0 \quad \text{for} \quad i \neq j$$

$$\langle b_i, b_i \rangle = 1$$

3.8.3 Gram-Schmidt Orthogonalization

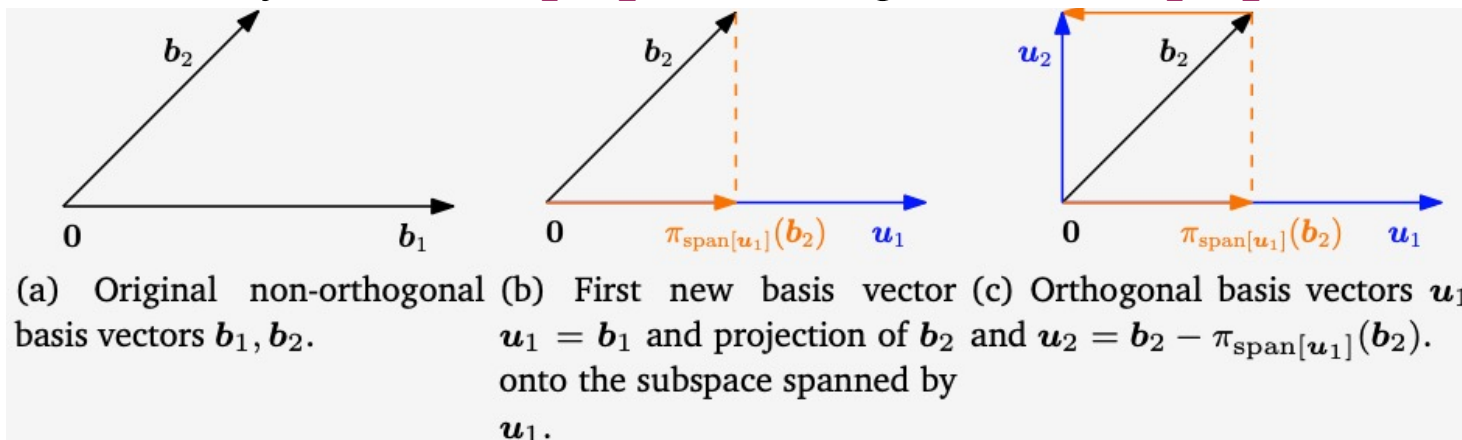- Consider a basis of $\mathbb{R}^2$

$$\boldsymbol{b}_1 = \begin{bmatrix} 2 \\ 0 \end{bmatrix} \qquad \boldsymbol{b}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

- Using the Gram-Schmidt method, we construct an orthogonal basis $(\boldsymbol{u}_1, \boldsymbol{u}_2)$ of $\mathbb{R}^2$ as follows (using dot product).

$$\boldsymbol{u}_1 := \boldsymbol{b}_1 = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$$

$$\boldsymbol{u}_2 := \boldsymbol{b}_2 - \pi_{\mathrm{span}[\boldsymbol{u}_1]}(\boldsymbol{b}_2) = \boldsymbol{b}_2 - \frac{\boldsymbol{u}_1 \boldsymbol{u}_1^{\mathrm{T}}}{\|\boldsymbol{u}_1\|^2} \boldsymbol{b}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

- We immediately see that $\boldsymbol{u}_1, \boldsymbol{u}_2$ are orthogonal, i.e., $\boldsymbol{u}_1^{\mathrm{T}} \boldsymbol{u}_2 = 0$

(a) Original non-orthogonal basis vectors $\boldsymbol{b}_1, \boldsymbol{b}_2$.

(b) First new basis vector $\boldsymbol{u}_1 = \boldsymbol{b}_1$ and projection of $\boldsymbol{b}_2$ onto the subspace spanned by $\boldsymbol{u}_1$.

(c) Orthogonal basis vectors $\boldsymbol{u}_1$ and $\boldsymbol{u}_2 = \boldsymbol{b}_2 - \pi_{\mathrm{span}[\boldsymbol{u}_1]}(\boldsymbol{b}_2)$.

3.8.3 Gram-Schmidt Orthogonalization

- Constructively transform basis $(\boldsymbol{b}_1, \cdots, \boldsymbol{b}_n)$ of an n-dim vector space V into an orthogonal/orthonormal basis $(\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n)$ of V.
$$\text{span}[\boldsymbol{b}_1, \cdots, \boldsymbol{b}_n] = \text{span}[\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n]$$

- The process iterates as follows

$$\boldsymbol{u}_1 := \boldsymbol{b}_1$$

$$\boldsymbol{u}_k := \boldsymbol{b}_k - \pi_{\text{span}[\boldsymbol{u}_1, \cdots, \boldsymbol{u}_{k-1}]}(\boldsymbol{b}_k), \qquad k = 2, \cdots, n$$

- The kth basis vector $\boldsymbol{b}_k$ is projected onto the subspace spanned by the first $k-1$ constructed orthogonal vectors $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_{k-1}$.

- This projection is then subtracted from $\boldsymbol{b}_k$ and yields a vector $\boldsymbol{u}_k$ that is orthogonal to the $(k-1)$-dim subspace spanned by $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_{k-1}$

- If we normalize $\boldsymbol{u}_k$, we obtain an ONB where $\|\boldsymbol{u}_k\| = 1$ for $k = 1, \cdots, n$.

Check your understanding

(A) Orthogonal projections are linear projections

(B) When applying orthogonal projection multiple times (>1), the projection result will not change anymore.

(C) Given a subspace to project on, orthogonal projection gives the minimum information loss (l_2).

(D) Gram-Schmidt Orthogonalization outputs the same number of basis vectors as the input.

(E) Projections allow us to better visualize and understand high-dimensional data.