Analytic Geometry 2

Jo Ciucă
Australian National University
comp36706670@anu.edu.au

Outline

- Updates
- Orthonormal basis & Orthogonal Complement
- Orthogonal Projections in 1D
- Orthogonal Projections in n-D
- Gram-Schmidt Orthogonalisation

Updates

- A big thank you to everyone who has applied to be a course rep.
- Please sign up for Ed using https://edstem.org/au/join/2C7AqM.
- Time for a little poll:

https://forms.office.com/r/GFMF5zHB6i

Updates

- Assignment 1 has been released with a deadline of the 28th of August.

(d) Let $\mathbf{A}$ be a square matrix and $f(\mathbf{X})$ and $g(\mathbf{X})$ be n-th order polynomials, defined by $\sum_{i=0}^{n} a_i \mathbf{X}^i$ where a_i are arbitrary real numbers. Show that the matrices $f(\mathbf{A})$ and $g(\mathbf{A})$ commute, i.e, $f(\mathbf{A})g(\mathbf{A}) = g(\mathbf{A})f(\mathbf{A})$ for arbitrary order n.

(e) Let $\mathbf{A}$ and $\mathbf{B}$ be rectangular matrices of orders $n \times k$ and $r \times s$, respectively. The matrix of order $nr \times ks$ represented in a block form as

$$\begin{bmatrix} a_{11}\mathbf{B} & a_{12}\mathbf{B} & \dots & a_{1k}\mathbf{B} \\ a_{21}\mathbf{B} & a_{22}\mathbf{B} & \dots & a_{2k}\mathbf{B} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1}\mathbf{B} & a_{n2}\mathbf{B} & \dots & a_{nk}\mathbf{B} \end{bmatrix}$$

is called the Kronecker product $\mathbf{A} \otimes \mathbf{B}$ of the matrices $\mathbf{A}$ and $\mathbf{B}$.

- Self-assessment released also - solutions to be discussed with your tutors this week.

Orthonormal Basis & Orthogonal Complement

3.5 Orthonormal Basis

- Consider an n-dimensional vector space V and a basis $\{\boldsymbol{b}_1, \ldots, \boldsymbol{b}_\mathrm{n}\}$ of V. For all $i, j = 1, \cdots, n$, if

$$\left\langle \boldsymbol{b}_i, \boldsymbol{b}_j \right\rangle = 0 \quad \text{for} \quad i \neq j \qquad (1)$$

$$\left\langle \boldsymbol{b}_i, \boldsymbol{b}_i \right\rangle = 1$$

then the basis is called an orthonormal basis (ONB).

If only (1) is satisfied, the basis is called an orthogonal basis.

Example (Orthonormal Basis)

- The canonical/standard basis for a Euclidean vector space $\mathbb{R}^n$ is an orthonormal basis, where the inner product is the dot product of vectors

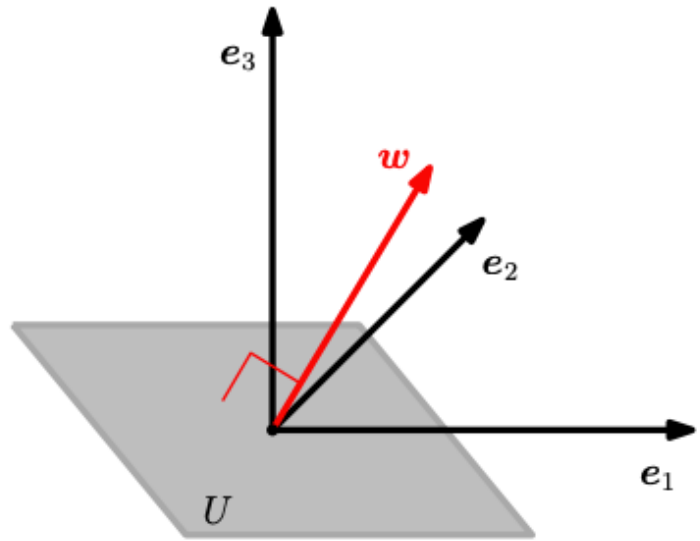
For $\mathbb{R}^3$: $$\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$$

- In $\mathbb{R}^2$, the vectors $\boldsymbol{b}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \boldsymbol{b}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$

  form an orthonormal basis since $b_1^\top b_2 = 0$ and $\|b_1\| = 1 = \|b_2\|$

3.6 Orthogonal Complement

- A 2-dimensional subspace U in a three-dimensional vector space can be described by its normal vector, which spans its orthogonal complement $U^{\perp}$.



- Generally, normal vectors can be used to describe $\boxed{(n-1)\text{ dimensional}}$ **hyperplanes** in n-dimensional vector and affine spaces.

3.6 Orthogonal Complement

- We now look at vector spaces that are orthogonal to each other
- Consider a D-dimensional vector space V and an M-dimensional subspace $U \subseteq V$. The orthogonal complement $U^{\perp}$ is a $(D-M)$-dimensional subspace of V and contains all vectors in V that are orthogonal to every vector in U.
- $U \cap U^{\perp} = \{\mathbf{0}\}$ so that any vector $\boldsymbol{x} \in V$ can be uniquely decomposed into

$$x = \sum_{m=1}^{M} \lambda_m \boldsymbol{b}_m + \sum_{j=1}^{D-M} \psi_j \boldsymbol{b}_j^{\perp}, \lambda_m, \psi_j \in \mathbb{R}$$

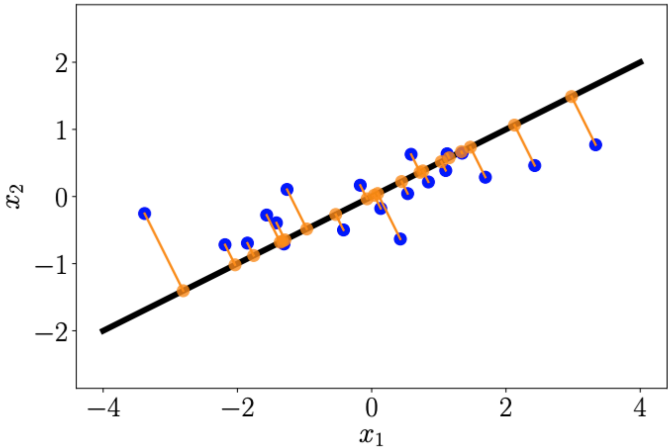
- Where $(\boldsymbol{b}_1, \ldots, \boldsymbol{b}_M)$ is a basis of U and $(\boldsymbol{b}_1^{\perp}, \ldots, \boldsymbol{b}_{D-M}^{\perp})$ is a basis of $U^{\perp}$.

Orthogonal Projections

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)

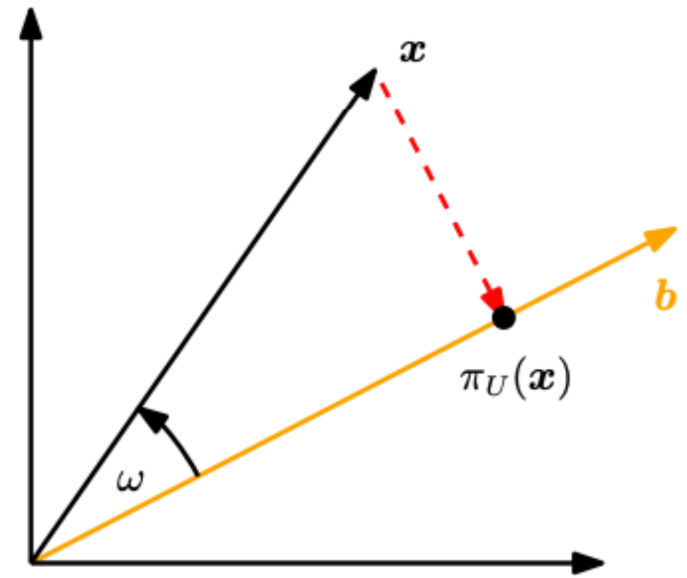


3.8 Orthogonal Projections

- Let V be a vector space and $U \subseteq V$ a subspace of V. A linear mapping $\pi: V \rightarrow V$ 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 $\boldsymbol{P}_\pi^2 = \boldsymbol{P}_\pi$.

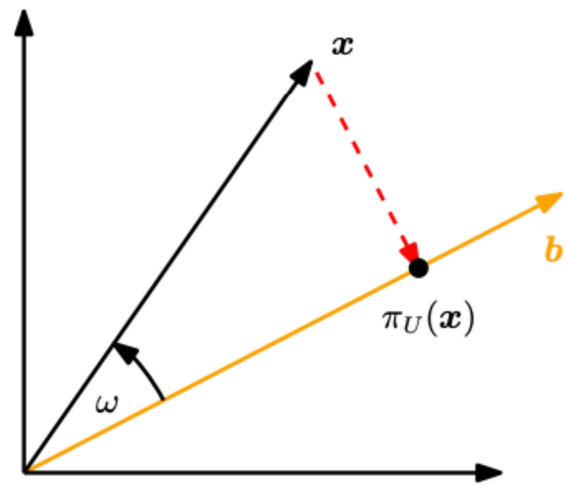
3.8.1 Projection onto One-Dimensional Subspaces (Lines)

- Assume we are given a line (one-dimensional subspace) through the origin with basis vector $b \in \mathbb{R}^n$.
- When we project $\boldsymbol{x} \in \mathbb{R}^n$ onto U, we seek the vector $\pi_U(\boldsymbol{x})$ that is closest to $\boldsymbol{x}$.



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

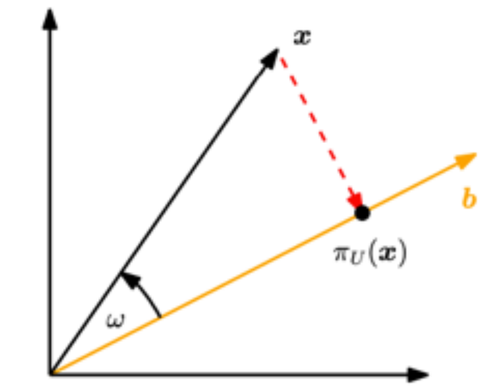
- The projection $\pi_U(\boldsymbol{x})$ should be closest to $\boldsymbol{x}$.
  - ➡ $\left\|\boldsymbol{x} - \pi_U(\boldsymbol{x})\right\|$ is minimal.
  - ➡ $\pi_U(\boldsymbol{x}) - \boldsymbol{x}$ is orthogonal to U, which is spanned by $\boldsymbol{b}$.
  - ➡ $\left\langle \pi_U(\boldsymbol{x}) - \boldsymbol{x},\ \boldsymbol{b} \right\rangle = 0$
- $\pi_U(\boldsymbol{x})$ is an element of U spanned by $\boldsymbol{b}$.
  - ➡ $\pi_U(\boldsymbol{x}) = \lambda \boldsymbol{b}$ for some $\lambda \in \mathbb{R}$.



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

How to determine λ, $\pi_U(\boldsymbol{x})$ and the projection matrix $\boldsymbol{P}_\pi$?

1. Finding the coordinate λ



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

- The orthogonality condition

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

- We use the bilinearity of inner product

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

inner products are symmetric

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

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

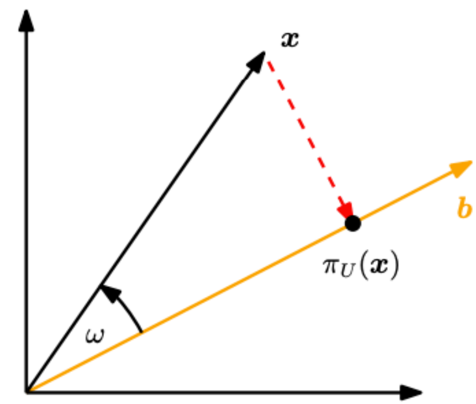
- If $\|b\| = 1$ then λ is given by $b^\top x$.

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

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

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

Assuming dot product



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

$$\|\pi_U(\boldsymbol{x})\| = \|\lambda \boldsymbol{b}\| = |\lambda| \, \|\boldsymbol{b}\|$$

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

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

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

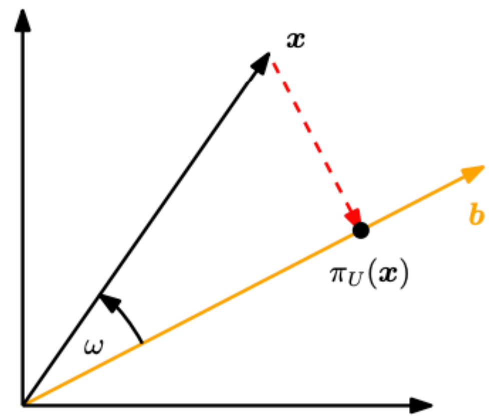
$$\|\pi_U(\boldsymbol{x})\| = \left\| \cos\omega \frac{\|\boldsymbol{x}\|}{\|\boldsymbol{b}\|}\boldsymbol{b} \right\| = |\cos\omega| \frac{\|\boldsymbol{x}\|}{\|\boldsymbol{b}\|}\|\boldsymbol{b}\| = |\cos\omega| \, \|\boldsymbol{x}\|$$

ω is the angle between $\boldsymbol{x}$ and $\boldsymbol{b}$. 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 $\boldsymbol{b} = [1 \quad -1]^\top$.

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

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

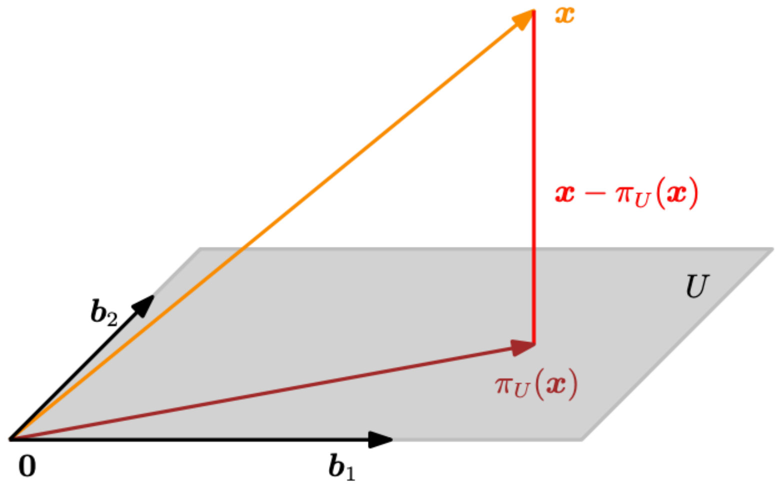
$$\pi_U(\boldsymbol{x}) = \boldsymbol{P}_\pi \boldsymbol{x} = \frac{1}{2}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} 3 \\ 5 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} -2 \\ 2 \end{bmatrix} \in span\left[\begin{bmatrix} 1 \\ -1 \end{bmatrix}\right]$$

- 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 $\boldsymbol{x}$.

Projection on General Subspaces

3.8.2 Projection onto General Subspaces

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



Projecting $\boldsymbol{x} \in \mathbb{R}^3$ onto a two-dimensional subspace

- Assume $(\boldsymbol{b}_1, \cdots, \boldsymbol{b}_m)$ is a basis of U.
- The projection $\pi_U(\boldsymbol{x})$ is a component of U.

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

- How to determine λ_i, $\pi_U(\boldsymbol{x})$ and $\boldsymbol{P}_\pi$?

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

- The linear combination

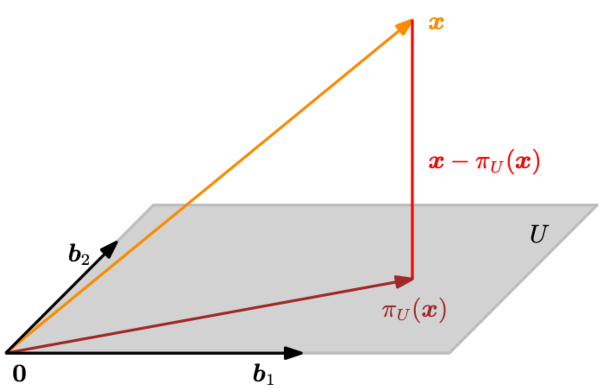
$$\pi_U(\boldsymbol{x}) = \sum_{i=1}^{m} \lambda_i \boldsymbol{b}_i = \boldsymbol{B}\boldsymbol{\lambda} \qquad \boldsymbol{B} = [\boldsymbol{b}_1, \ \ldots, \ \boldsymbol{b}_m] \in \mathbb{R}^{n\times m}, \boldsymbol{\lambda} = [\lambda_1, \ldots, \lambda_m]^{\mathrm{T}} \in \mathbb{R}^m$$

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

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

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

$$\begin{gathered}\langle \boldsymbol{b}_1, \boldsymbol{x} - \pi_U(\boldsymbol{x}) \rangle = \boldsymbol{b}_1^{\mathrm{T}}(\boldsymbol{x} - \pi_U(\boldsymbol{x})) = 0 \\ \vdots \\ \langle \boldsymbol{b}_m, \boldsymbol{x} - \pi_U(\boldsymbol{x}) \rangle = \boldsymbol{b}_m^{\mathrm{T}}(\boldsymbol{x} - \pi_U(\boldsymbol{x})) = 0\end{gathered}$$



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

$$\begin{gathered}\boldsymbol{b}_1^{\mathrm{T}}(\boldsymbol{x} - \boldsymbol{B}\boldsymbol{\lambda}) = 0 \\ \vdots \\ \boldsymbol{b}_m^{\mathrm{T}}(\boldsymbol{x} - \boldsymbol{B}\boldsymbol{\lambda}) = 0\end{gathered}$$

such that we obtain a homogeneous linear equation system

$$\begin{bmatrix} \boldsymbol{b}_1^{\mathrm{T}} \\ \vdots \\ \boldsymbol{b}_m^{\mathrm{T}} \end{bmatrix} [\boldsymbol{x} - \boldsymbol{B}\boldsymbol{\lambda}] = 0 \Leftrightarrow \boldsymbol{B}^{\mathrm{T}}(\boldsymbol{x} - \boldsymbol{B}\boldsymbol{\lambda}) = 0 \Leftrightarrow \boldsymbol{B}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{\lambda} = \boldsymbol{B}^{\mathrm{T}}\boldsymbol{x}.$$

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

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

- $\boldsymbol{b}_1, \ . \ . \ . \ , \boldsymbol{b}_m$ are a basis of U, so they are linearly independent.

$$\Rightarrow r\left(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{B}\right) = r(\boldsymbol{B}) = m$$

This allows us to solve λ

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

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

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

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

3. Find the projection matrix $\boldsymbol{P}_\pi$

- We have $\boldsymbol{P}_\pi\, \boldsymbol{x} = \pi_U(\boldsymbol{x})$
- From step 2, we have

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

- We can immediately see that

$$\boldsymbol{P}_\pi = \boldsymbol{B}(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{B})^{-1}\boldsymbol{B}^{\mathrm{T}}$$

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

$$\boldsymbol{P}_\pi = \boldsymbol{B}(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{B})^{-1}\boldsymbol{B}^{\mathrm{T}}$$

into

$$\boldsymbol{P}_\pi = \frac{\boldsymbol{b}\boldsymbol{b}^\top}{\|\boldsymbol{b}\|^2}$$

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

Example - Projection onto a Two-dimensional Subspace

- For a subspace $U = \text{span}[\begin{bmatrix}1\\1\\1\end{bmatrix}, \begin{bmatrix}0\\1\\2\end{bmatrix}] \subseteq \mathbb{R}^3$, and $\boldsymbol{x} = \begin{bmatrix}6\\0\\0\end{bmatrix} \in \mathbb{R}^3$, find the coordinates λ 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 $\boldsymbol{B}^{\mathrm{T}}\boldsymbol{B}\lambda = \boldsymbol{B}^{\mathrm{T}}\boldsymbol{x}$ to find λ:

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

Example - Projection onto a Two-dimensional Subspace

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

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

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

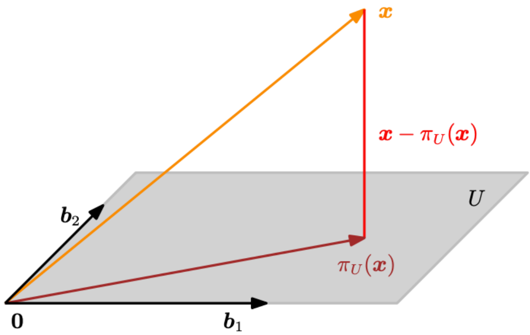
$$\left\| \boldsymbol{x} - \pi_U(\boldsymbol{x}) \right\| = \left\| \begin{bmatrix} 1 & -2 & 1 \end{bmatrix}^{\mathrm{T}} \right\| = \sqrt{6}$$

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

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

Things to note

- $\pi_U(\boldsymbol{x})$ is still in $\mathbb{R}^3$, although it lies in a 2-dim subspace $U \subseteq \mathbb{R}^3$



- For our case, if $\boldsymbol{B}$ columns are an orthonomal basis (ONB), i.e., $\boldsymbol{B}^{\mathrm{T}}\boldsymbol{B} = \boldsymbol{I}$, we have

$$\lambda = \boldsymbol{B}^{\mathrm{T}}\boldsymbol{x}$$
$$\pi_U(\boldsymbol{x}) = \boldsymbol{B}\boldsymbol{B}^{\mathrm{T}}\boldsymbol{x}$$

$$\left\langle \boldsymbol{b}_i, \boldsymbol{b}_j \right\rangle = 0 \quad \text{for} \quad i \neq j$$
$$\left\langle \boldsymbol{b}_i, \boldsymbol{b}_i \right\rangle = 1$$

- We can find approximate solutions to unsolvable linear equation systems $\boldsymbol{A}u = v$ using projections.
- The idea is to find the vector in the subspace spanned by the columns of $\boldsymbol{A}$ that is closest to v, i.e., we compute the orthogonal projection of v onto the subspace spanned by the columns of $\boldsymbol{A}$. --- least-squares solution

3.8.3 Gram-Schmidt Orthogonalization

3.8.3 Gram-Schmidt Orthogonalization

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

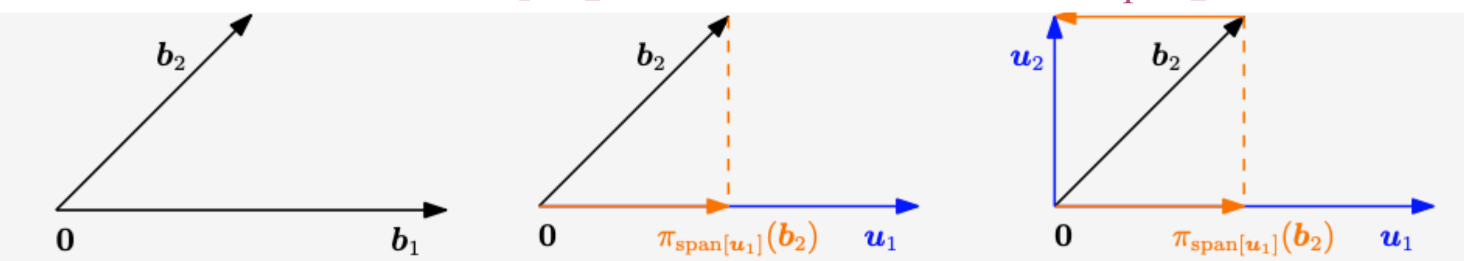
$$b_1 = \begin{bmatrix} 2 \\ 0 \end{bmatrix} \qquad b_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

- Using the Gram-Schmidt method, we construct an orthogonal basis (u_1, u_2) of $\mathbb{R}^2$ as follows (using dot product).

$$u_1 := b_1 = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$$

$$u_2 := b_2 - \pi_{\text{span}[u_1]}(b_2) = b_2 - \frac{u_1 u_1^{\mathrm{T}}}{\|u_1\|^2} b_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

- We immediately see that u_1, u_2 are orthogonal, i.e., $u_1^{\mathrm{T}} u_2 = 0$



(a) Original non-orthogonal basis vectors b_1, b_2.

(b) First new basis vector $u_1 = b_1$ and projection of b_2 onto the subspace spanned by u_1.

(c) Orthogonal basis vectors u_1 and $u_2 = b_2 - \pi_{\text{span}[u_1]}(b_2)$.

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

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

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

- 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 result will no longer change.

(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 visualize better and understand high-dimensional data.