Lecture 3: Analytic Geometry

Yi, Yung (이융)

Mathematics for Machine Learning
https://yung-web.github.io/home/courses/mathml.html
KAIST EE

April 4, 2021

Roadmap

Roadmap

Norm

- A notion of the length of vectors
- Definition. A norm on a vector space V is a function $\|\cdot\| : V \mapsto \mathbb{R}$, such that for all $\lambda \in \mathbb{R}$ the following hold:
  - Absolutely homogeneous: $\|\lambda \boldsymbol{x}\| = |\lambda| \, \|\boldsymbol{x}\|$
  - Triangle inequality: $\|\boldsymbol{x} + \boldsymbol{y}\| \le \|\boldsymbol{x}\| + \|\boldsymbol{y}\|$
  - Positive definite: $\|\boldsymbol{x}\| \ge 0$ and $\|\boldsymbol{x}\| \iff \boldsymbol{x} = 0$

Example for $V \in \mathbb{R}^n$

- Manhattan Norm (also called ℓ_1 norm) For $\boldsymbol{x} = [x_1, \cdots, x_n] \in \mathbb{R}^n$,

$$\|\boldsymbol{x}\|_1 :== \sum_{i=1}^{n} |x_i|$$

- Euclidean Norm (also called ℓ_2 norm) For $\boldsymbol{x} \in \mathbb{R}^n$,

$$\|\boldsymbol{x}\|_2 :== \sqrt{\sum_{i=1}^{n} x_i^2} = \sqrt{\boldsymbol{x}^\mathsf{T}\boldsymbol{x}}$$

Roadmap

(1) Norms
(2) Inner Products
(3) Lengths and Distances
(4) Angles and Orthogonality
(5) Orthonormal Basis
(6) Orthogonal Complement
(7) Inner Product of Functions
(8) Orthogonal Projections
(9) Rotations

Motivation

- Need to talk about the length of a vector and the angle or distance between two vectors, where vectors are defined in abstract vector spaces
- To this end, we define the notion of inner product in an abstract manner.
- Dot product: A kind of inner product in vector space $\mathbb{R}^n$. $\boldsymbol{x}^\mathsf{T}\boldsymbol{y} = \sum_{i=1}^{n} x_i y_i$
- Question. How can we generalize this and do a similar thing in some other vector spaces?

Formal Definition

- An inner product is a mapping $\langle \cdot, \cdot \rangle : V \times V \mapsto \mathbb{R}$ that satisfies the following conditions for all vectors $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in V$ and all scalars $\lambda \in \mathbb{R}$:
  1. $\langle \boldsymbol{u} + \boldsymbol{v}, \boldsymbol{w} \rangle = \langle \boldsymbol{u}, \boldsymbol{w} \rangle + \langle \boldsymbol{v}, \boldsymbol{w} \rangle$
  2. $\langle \lambda \boldsymbol{v}, \boldsymbol{w} \rangle = \lambda \langle \boldsymbol{v}, \boldsymbol{w} \rangle$
  3. $\langle \boldsymbol{v}, \boldsymbol{w} \rangle = \langle \boldsymbol{w}, \boldsymbol{v} \rangle$
  4. $\langle \boldsymbol{v}, \boldsymbol{v} \rangle \geq 0$ and equal iff $\boldsymbol{v} = 0$
- The pair $(V, \langle \cdot, \cdot \rangle)$ is called an inner product space.

Example

- Example. $V = \mathbb{R}^n$ and the dot product $\langle \boldsymbol{x}, \boldsymbol{y} \rangle := \boldsymbol{x}^\mathsf{T}\boldsymbol{y}$
- Example. $V = \mathbb{R}^2$ and $\langle \boldsymbol{x}, \boldsymbol{y} \rangle := x_1y_1 - (x_1y_2 + x_2y_1) + 2x_2y_2$
- Example. $V = \{\text{continuous functions in } \mathbb{R} \text{ over } [a, b]\}$, $\langle u, v \rangle := \int_a^b u(x)v(x)dx$

Symmetric, Positive Definite Matrix

- A square matrix $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ that satisfies the following is called symmetric, positive definite (or just positive definite):

$$\forall \boldsymbol{x} \in V \setminus \{0\} : \boldsymbol{x}^\mathsf{T}\boldsymbol{A}\boldsymbol{x} > 0.$$

  If only $\geq$ in the above holds, then $\boldsymbol{A}$ is called symmetric, positive semidefinite.

- $\boldsymbol{A}_1 = \begin{pmatrix} 9 & 6 \\ 6 & 5 \end{pmatrix}$ is positive definite.
- $\boldsymbol{A}_2 = \begin{pmatrix} 9 & 6 \\ 6 & 3 \end{pmatrix}$ is not positive definite.

Inner Product and Positive Definite Matrix (1)

- Consider an n-dimensional vector space V with an inner product $\langle \cdot, \cdot \rangle$ and an ordered basis $B = (\boldsymbol{b}_1, \dots, \boldsymbol{b}_n)$ of V.
- Any $\boldsymbol{x}, \boldsymbol{y} \in V$ can be represented as: $\boldsymbol{x} = \sum_{i=1}^n \psi_i \boldsymbol{b}_i$ and $\boldsymbol{y} = \sum_{i=j}^n \lambda_j \boldsymbol{b}_j$ for some ψ_i and λ_j, $i, j = 1, \dots, n$.

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \left\langle \sum_{i=1}^n \psi_i \boldsymbol{b}_i, \sum_{i=j}^n \lambda_j \boldsymbol{b}_j \right\rangle = \sum_{i=1}^n \sum_{j=1}^n \psi_i \langle \boldsymbol{b}_i, \boldsymbol{b}_j \rangle \lambda_j = \hat{\boldsymbol{x}}^\mathsf{T}\boldsymbol{A}\hat{\boldsymbol{y}},$$

  where $\boldsymbol{A}_{ij} = \langle \boldsymbol{b}_i, \boldsymbol{b}_j \rangle$ and $\hat{\boldsymbol{x}}$ and $\hat{\boldsymbol{y}}$ are the coordinates w.r.t. B.

Inner Product and Positive Definite Matrix (2)

- Then, if $\forall \boldsymbol{x} \in V \setminus \{0\} : \boldsymbol{x}^\mathsf{T}\boldsymbol{A}\boldsymbol{x} > 0$ (i.e., $\boldsymbol{A}$ is symmetric, positive definite), $\hat{\boldsymbol{x}}^\mathsf{T}\boldsymbol{A}\hat{\boldsymbol{y}}$ legitimately defines an inner product (w.r.t. B)
- Properties
  - The kernel of $\boldsymbol{A}$ is only $\{0\}$, because $\boldsymbol{x}^\mathsf{T}\boldsymbol{A}\boldsymbol{x} > 0$ for all $\boldsymbol{x} \neq 0 \implies \boldsymbol{A}\boldsymbol{x} \neq 0$ if $\boldsymbol{x} \neq 0$.
  - The diagonal elements a_{ii} of $\boldsymbol{A}$ are all positive, because $a_{ii} = \boldsymbol{e}_i^\mathsf{T}\boldsymbol{A}\boldsymbol{e}_i > 0$.

Roadmap

(1) Norms

(2) Inner Products

(3) Lengths and Distances

(4) Angles and Orthogonality

(5) Orthonormal Basis

(6) Orthogonal Complement

(7) Inner Product of Functions

(8) Orthogonal Projections

(9) Rotations

Length

- Inner product naturally induces a norm by defining:
$$\|x\| := \sqrt{\langle \boldsymbol{x}, \boldsymbol{x} \rangle}$$
- Not every norm is induced by an inner product
- Cachy-Schwarz inequality. For the induced norm by the inner product,
$$|\langle \boldsymbol{x}, \boldsymbol{y} \rangle| \le \|\boldsymbol{x}\| \; \|\boldsymbol{y}\|$$

Distance

- Now, we can introduce a notion of distance using a norm as:
  Distance. $d(\boldsymbol{x}, \boldsymbol{y}) := \|\boldsymbol{x} - \boldsymbol{y}\| = \sqrt{\langle \boldsymbol{x} - \boldsymbol{y}, \boldsymbol{x} - \boldsymbol{y} \rangle}$
- If the dot product is used as an inner product in $\mathbb{R}^n$, it is Euclidian distance.
- Note. The distance between two vectors does NOT necessarily require the notion of norm. Norm is just sufficient.
- Generally, if the following is satisfied, it is a suitable notion of distance, called metric.
  - *Positive definite*. $d(\boldsymbol{x}, \boldsymbol{y}) \ge 0$ for all $\boldsymbol{x}, \boldsymbol{y}$ and $d(\boldsymbol{x}, \boldsymbol{y}) = 0 \iff \boldsymbol{x} = \boldsymbol{y}$
  - *Symmetric*. $d(\boldsymbol{x}, \boldsymbol{y}) = d(\boldsymbol{y}, \boldsymbol{x})$
  - *Triangle inequality*. $d(\boldsymbol{x}, \boldsymbol{z}) \le d(\boldsymbol{x}, \boldsymbol{y}) + d(\boldsymbol{y}, \boldsymbol{z})$

Angle, Orthogonal, and Orthonormal

- Using C-S inequality,
$$-1 \le \frac{\langle \boldsymbol{x}, \boldsymbol{y} \rangle}{\|\boldsymbol{x}\| \; \|\boldsymbol{y}\|} \le 1$$
- Then, there exists a unique $\omega \in [0, \pi]$ with
$$\cos \omega = \frac{\langle \boldsymbol{x}, \boldsymbol{y} \rangle}{\|\boldsymbol{x}\| \; \|\boldsymbol{y}\|}$$
- We define ω as the angle between $\boldsymbol{x}$ and $\boldsymbol{y}$.
- Definition. If $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = 0$, in other words their angle is $\pi/2$, we say that they are orthogonal, denoted by $\boldsymbol{x} \perp \boldsymbol{y}$. Additionally, if $\|x\| = \|y\| = 1$, they are orthonormal.

Example

- Orthogonality is defined by a given inner product. Thus, different inner products may lead to different results about orthogonality.
- Example. Consider two vectors $\boldsymbol{x} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $\boldsymbol{y} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$
- Using the dot product as the inner product, they are orthogonal.
- However, using $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \boldsymbol{x}^\mathsf{T} \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} \boldsymbol{y}$, they are not orthogonal.

$$\cos\omega = \frac{\langle \boldsymbol{x}, \boldsymbol{y} \rangle}{\|\boldsymbol{x}\| \ \|\boldsymbol{y}\|} = -\frac{1}{3} \implies \omega \approx 1.91 \text{ rad } \approx 109.5^\circ$$

Orthogonal Matrix

- Definition. A square matrix $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ is an orthogonal matrix, iff its columns (or rows) are orthonormal so that

$$\boldsymbol{A}\boldsymbol{A}^\mathsf{T} = I = \boldsymbol{A}^\mathsf{T}\boldsymbol{A}, \text{ implying } \boldsymbol{A}^{-1} = \boldsymbol{A}^\mathsf{T}.$$

  - We can use $\boldsymbol{A}^{-1} = \boldsymbol{A}^\mathsf{T}$ for the definition of orthogonal matrices.
  - Fact 1. $\boldsymbol{A}, \boldsymbol{B}$: orthogonal $\implies$ $\boldsymbol{AB}$: orthogonal
  - Fact 2. $\boldsymbol{A}$: orthogonal $\implies$ $\det(\boldsymbol{A}) = \pm 1$
- The linear mapping Φ by orthogonal matrices preserve length and angle (for the dot product)

$$\|\Phi(\boldsymbol{A})\| = \|\boldsymbol{Ax}\|^2 = (\boldsymbol{Ax})^\mathsf{T}(\boldsymbol{Ax}) = \boldsymbol{x}^\mathsf{T}\boldsymbol{A}^\mathsf{T}\boldsymbol{Ax} = \boldsymbol{x}^\mathsf{T}\boldsymbol{x} = \|\boldsymbol{x}\|^2$$

$$\cos\omega = \frac{(\boldsymbol{Ax})^\mathsf{T}(\boldsymbol{Ay})}{\|\boldsymbol{Ax}\| \ \|\boldsymbol{Ay}\|} = \frac{\boldsymbol{x}^\mathsf{T}\boldsymbol{A}^\mathsf{T}\boldsymbol{Ay}}{\sqrt{\boldsymbol{x}^\mathsf{T}\boldsymbol{A}^\mathsf{T}\boldsymbol{Axy}^\mathsf{T}\boldsymbol{A}^\mathsf{T}\boldsymbol{Ay}}} = \frac{\boldsymbol{x}^\mathsf{T}\boldsymbol{y}}{\|\boldsymbol{x}\| \ \|\boldsymbol{y}\|}$$

Roadmap

Orthonormal Basis

- Basis that is orthonormal, i.e., they are all orthogonal to each other and their lengths are 1.
- Standard basis in $\mathbb{R}^n$, $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$, is orthonormal.
- Question. How to obtain an orthonormal basis?

1. Use Gaussian elimination to find a basis for a vector space spanned by a set of vectors.
   - Given a set $\{\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n\}$ of unorthogonal and unnormalized basis vectors. Apply Gaussian elimination to the augmented matrix $(\boldsymbol{B}\boldsymbol{B}^\mathsf{T} | \boldsymbol{B})$
2. Constructive way: Gram-Schmidt process (we will cover this later)

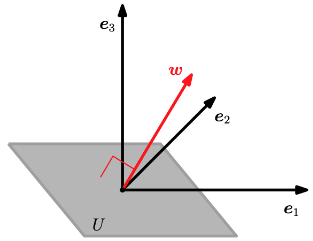
Orthogonal Complement (1)

- Consider D-dimensional vector space V and M-dimensional subspace $W \subset V$. 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.
- $U \cap U^\perp = 0$
- Any vector $x \in V$ can be uniquely decomposed into:

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

where $(\boldsymbol{b}_1 \ldots, \boldsymbol{b}_M)$ and $(\boldsymbol{b}_1^\perp, \ldots, \boldsymbol{b}_{D-M}^\perp)$ are the bases of U and $U^\perp$, respectively.

Orthogonal Complement (2)



- The vector $\boldsymbol{w}$ with $\|\boldsymbol{w}\| = 1$, which is orthogonal to U, is the basis of $U^\perp$.
- Such $\boldsymbol{w}$ is called normal vector to U.
- For a linear mapping represented by a matrix $\boldsymbol{A} \in \mathbb{R}^{m \times n}$, the solution space of $\boldsymbol{Ax} = 0$ is $\text{row}(\boldsymbol{A})^\perp$, where $\text{row}(\boldsymbol{A})$ is the row space of $\boldsymbol{A}$ (i.e., span of row vectors). In other words, $\text{row}(\boldsymbol{A})^\perp = \ker(\boldsymbol{A})$

Inner Product of Functions

- Remind: $V = \{\text{continuous functions in } \mathbb{R} \text{ over } [a, b]\}$, the following is a proper inner product.

$$\langle u, v \rangle := \int_a^b u(x)v(x)dx$$

- Example. Choose $u(x) = \sin(x)$ and $v(x) = \cos(x)$, where we select $a = -\pi$ and $b = \pi$. Then, since $f(x) = u(x)v(x)$ is odd (i.e., $f(-x) = -f(x)$),

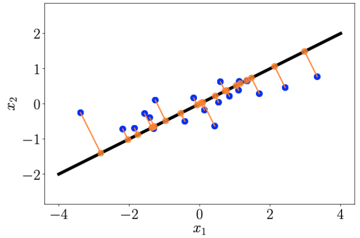
$$\int_{-\pi}^{\pi} u(x)v(x)dx = 0.$$

- Thus, u and v are orthogonal.
- Similarly, $\{1, \cos(x), \cos(2x), \cos(3x), \ldots,\}$ is orthogonal over $[-\pi, \pi]$.

Roadmap

Projection: Motivation

- Big data: high dimensional
- However, most information is contained in a few dimensions
- Projection: A process of reducing the dimensions (hopefully) without loss of much information[1]
- Example. Projection of 2D dataset onto 1D subspace



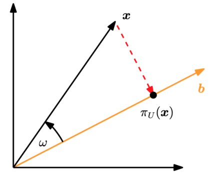
[1] In L10, we will formally study this with the topic of PCA (Principal Component Analysis).

Projection onto Lines (1D Subspaces)

- Consider a 1D subspace $U \subset \mathbb{R}^n$ spanned by the basis $\boldsymbol{b}$.
- For $\boldsymbol{x} \in \mathbb{R}^n$, what is its projection $\pi_U(\boldsymbol{x})$ onto U (assume the dot product)?

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

$$\Longrightarrow \lambda = \frac{\langle \boldsymbol{b}, \boldsymbol{x} \rangle}{\|\boldsymbol{b}\|^2} = \frac{\boldsymbol{b}^\mathsf{T} \boldsymbol{x}}{\|\boldsymbol{b}\|^2}, \text{ and } \pi_U(\boldsymbol{x}) = \lambda \boldsymbol{b} = \frac{\boldsymbol{b}^\mathsf{T} \boldsymbol{x}}{\|\boldsymbol{b}\|^2} \boldsymbol{b}$$



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

- Projection matrix $\boldsymbol{P}_\pi \in \mathbb{R}^{n \times n}$ in $\pi_U(\boldsymbol{x}) = \boldsymbol{P}_\pi \boldsymbol{x}$

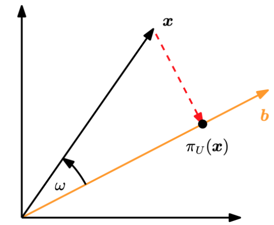
$$\pi_U(\boldsymbol{x}) = \lambda \boldsymbol{b} = \boldsymbol{b} \lambda = \frac{\boldsymbol{b}\boldsymbol{b}^\mathsf{T}}{\|\boldsymbol{b}\|^2} \boldsymbol{x}, \quad \boldsymbol{P}_\pi = \frac{\boldsymbol{b}\boldsymbol{b}^\mathsf{T}}{\|\boldsymbol{b}\|^2}$$

Inner Product and Projection

- We project $\boldsymbol{x}$ onto $\boldsymbol{b}$, and let $\pi_{\boldsymbol{b}}(\boldsymbol{x})$ be the projected vector.
- Question. Understanding the inner project $\langle \boldsymbol{x}, \boldsymbol{b} \rangle$ from the projection perspective?

$$\boxed{\langle \boldsymbol{x}, \boldsymbol{b} \rangle = \|\pi_{\boldsymbol{b}}(\boldsymbol{x})\| \times \|\boldsymbol{b}\|}$$

- In other words, the inner product of $\boldsymbol{x}$ and $\boldsymbol{b}$ is the product of (length of the projection of $\boldsymbol{x}$ onto $\boldsymbol{b}$) $\times$ (length of $\boldsymbol{b}$)



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

Example

- $\boldsymbol{b} = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}$

$$\boldsymbol{P}_\pi = \frac{\boldsymbol{b}\boldsymbol{b}^\mathsf{T}}{\|\boldsymbol{b}\|^2} = \frac{1}{9} \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix} \begin{pmatrix} 1 & 2 & 2 \end{pmatrix} = \frac{1}{9} \begin{pmatrix} 1 & 2 & 2 \\ 2 & 4 & 4 \\ 2 & 4 & 4 \end{pmatrix}$$

For $\boldsymbol{x} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$,

$$\pi_U(\boldsymbol{x}) = \boldsymbol{P}_\pi \boldsymbol{x} = \frac{1}{9} \begin{pmatrix} 1 & 2 & 2 \\ 2 & 4 & 4 \\ 2 & 4 & 4 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \frac{1}{9} \begin{pmatrix} 5 \\ 10 \\ 10 \end{pmatrix} \in \text{span}[\begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}]$$

Projection onto General Subspaces

- $\mathbb{R}^n \to$ 1-Dim
- A basis vector $\boldsymbol{b}$ in 1D subspace

$$\pi_U(\boldsymbol{x}) = \frac{\boldsymbol{b}\boldsymbol{b}^\mathsf{T}\boldsymbol{x}}{\boldsymbol{b}^\mathsf{T}\boldsymbol{b}}, \quad \lambda = \frac{\boldsymbol{b}^\mathsf{T}\boldsymbol{x}}{\boldsymbol{b}^\mathsf{T}\boldsymbol{b}}$$

$$\boldsymbol{P}_\pi = \frac{\boldsymbol{b}\boldsymbol{b}^\mathsf{T}}{\boldsymbol{b}^\mathsf{T}\boldsymbol{b}}$$

- $\mathbb{R}^n \to m$-Dim, $(m < n)$
- A basis matrix $\boldsymbol{B} = (\boldsymbol{b}_1, \cdots, \boldsymbol{b}_m) \in \mathbb{R}^{n \times m}$

$$\pi_U(\boldsymbol{x}) = \boldsymbol{B}(\boldsymbol{B}^\mathsf{T}\boldsymbol{B})^{-1}\boldsymbol{B}^\mathsf{T}\boldsymbol{x}, \quad \boldsymbol{\lambda} = (\boldsymbol{B}^\mathsf{T}\boldsymbol{B})^{-1}\boldsymbol{B}^\mathsf{T}\boldsymbol{x}$$

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

- $\lambda \in \mathbb{R}^1$ and $\boldsymbol{\lambda} \in \mathbb{R}^m$ are the coordinates in the projected spaces, respectively.
- $(\boldsymbol{B}^\mathsf{T}\boldsymbol{B})^{-1}\boldsymbol{B}^\mathsf{T}$ is called pseudo-inverse.
- How to derive is analogous to the case of 1-D lines (see pp. 71).

Example: Projection onto 2D Subspace

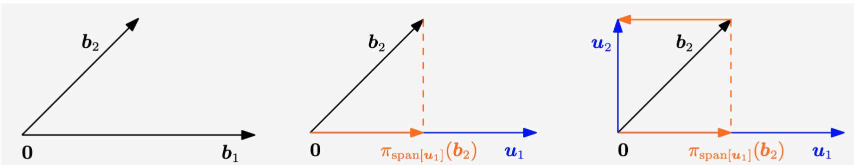
- $U = \text{span}[\begin{pmatrix}1\\1\\1\end{pmatrix}, \begin{pmatrix}0\\1\\2\end{pmatrix}] \subset \mathbb{R}^3$ and $\boldsymbol{x} = \begin{pmatrix}6\\0\\0\end{pmatrix}$. Check that $\{(1\ 1\ 1)^\mathsf{T}, (0\ 1\ 2)^\mathsf{T}\}$ is a basis.
- Let $\boldsymbol{B} = \begin{pmatrix}1 & 0\\1 & 1\\1 & 2\end{pmatrix}$. Then, $\boldsymbol{B}^\mathsf{T}\boldsymbol{B} = \begin{pmatrix}1 & 1 & 2\\0 & 1 & 2\end{pmatrix}\begin{pmatrix}1 & 0\\1 & 1\\1 & 2\end{pmatrix} = \begin{pmatrix}3 & 3\\3 & 5\end{pmatrix}$
- Can see that $\boldsymbol{P}_\pi = \boldsymbol{B}(\boldsymbol{B}^\mathsf{T}\boldsymbol{B})^{-1}\boldsymbol{B}^\mathsf{T} = \frac{1}{6}\begin{pmatrix}5 & 2 & -1\\2 & 2 & 2\\-1 & 2 & 5\end{pmatrix}$, and

$$\pi_U(\boldsymbol{x}) = \frac{1}{6}\begin{pmatrix}5 & 2 & -1\\2 & 2 & 2\\-1 & 2 & 5\end{pmatrix}\begin{pmatrix}6\\0\\0\end{pmatrix} = \begin{pmatrix}5\\2\\-1\end{pmatrix}$$

Gram-Schmidt Orthogonalization Method (G-S method)

- Constructively transform any basis $(\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n)$ of n-dimensional vector space V into an orthogonal/orthonormal basis $(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n)$ of V
- Iteratively construct as follows

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



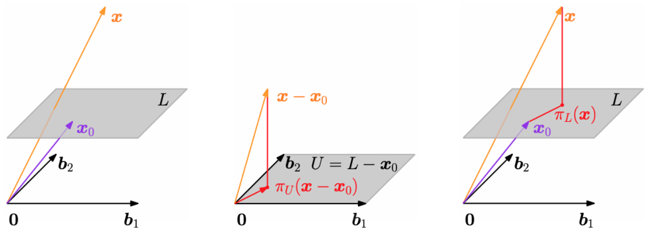
Example: G-S method

- A basis $(\boldsymbol{b}_1, \boldsymbol{b}_2) \in \mathbb{R}^2$, $\boldsymbol{b}_1 = \begin{pmatrix}2\\0\end{pmatrix}$ and $\boldsymbol{b}_2 = \begin{pmatrix}1\\1\end{pmatrix}$
- $\boldsymbol{u}_1 = \boldsymbol{b}_1 = \begin{pmatrix}2\\0\end{pmatrix}$ and

$$\boldsymbol{u}_2 = \boldsymbol{b}_2 - \pi_{\text{span}[\boldsymbol{u}_1]}(\boldsymbol{b}_2) = \frac{\boldsymbol{u}_1\boldsymbol{u}_2^\mathsf{T}}{\|\boldsymbol{u}_1\|}\boldsymbol{b}_2 = \begin{pmatrix}1\\1\end{pmatrix} - \begin{pmatrix}1 & 0\\0 & 0\end{pmatrix}\begin{pmatrix}1\\1\end{pmatrix} = \begin{pmatrix}0\\1\end{pmatrix}$$

- $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$ are orthogonal. If we want them to be orthonormal, then just normaliation would do the job.

Projection onto Affine Subspaces



- Affine space: $L = \boldsymbol{x}_0 + U$
- Affine subspaces are not vector spaces
- Idea: (i) move $\boldsymbol{x}$ to a point in U, (ii) do the projection, (iii) move back to L

$$\pi_L(\boldsymbol{x}) = \boldsymbol{x}_0 + \pi_U(\boldsymbol{x} - \boldsymbol{x}_0)$$

Roadmap

Rotation

- Length and angle preservation: two properties of linear mappings with orthogonal matrices. Let's look at some of their special cases.
- A linear mapping that rotates the given coordinate system by an angle θ.
- Basis change
- $\boldsymbol{e}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \to \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}$ and $\boldsymbol{e}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \to \begin{pmatrix} -\sin\theta \\ \cos\theta \end{pmatrix}$
- Rotation matrix $\boldsymbol{R}(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$
- Properties
  - Preserves distance: $\|\boldsymbol{x} - \boldsymbol{y}\| = \|\boldsymbol{R}_\theta(\boldsymbol{x}) - \boldsymbol{R}_\theta(\boldsymbol{y})\|$
  - Preserves angle

Questions?

1)