Analytic Geometry

CS115 - Math for Computer Science

TS. Lương Ngọc Hoàng
TS. Dương Việt Hằng

September 9, 2023

Roadmap

- Norm
- Inner Products
- Lengths and Distances
- Angles and Orthogonality
- Orthonormal Basis
- Orthogonal Complement
- Inner Product of Functions
- Orthogonal Projections
- Rotations

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: $\|\boldsymbol{x} + \boldsymbol{y}\| \leq \|\boldsymbol{x}\| + \|\boldsymbol{y}\|$
  - Positive definite: $\|\boldsymbol{x}\| \geq 0$ and $\|\boldsymbol{x}\| \Longleftrightarrow \boldsymbol{x} = 0$

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

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

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

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

$$\|\mathbf{x}\|_2 :== \sqrt{\sum_{i=1}^{n} x_i^2} = \sqrt{\mathbf{x}^\top \mathbf{x}}$$

Formal Definitions

- 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.

Examples

- Example. $V = \mathbb{R}^n$ and the dot product $\langle \mathbf{x}, \mathbf{y} \rangle := \mathbf{x}^\mathsf{T}\mathbf{y}$

- Example. $V = \mathbb{R}^2$ and $\langle \mathbf{x}, \mathbf{y} \rangle := x_1 y_1 - (x_1 y_2 + x_2 y_1) + 2x_2 y_2$

- Example. $V = \{\text{continuous functions in } \mathbb{R} \text{ over } [a, b]\}$, $\langle u, v \rangle := \int_a^b u(x)v(x)dx$

Positive/Negative Definite Matrix

Definitions:

1) An *nxn* symmetric real matrix ***A*** is said to be **positive-definite if $\boldsymbol{x^T A x > 0}$** for all non-zero x$\in R^n$

2) An *nxn* symmetric real matrix ***A*** is said to be **positive-semidefinite if $\boldsymbol{x^T A x \geq 0}$** for all non-zero x$\in R^n$

3) An *nxn* symmetric real matrix ***A*** is said to be nega**tive-definite if $\boldsymbol{x^T A x < 0}$** for all non-zero x$\in R^n$

4) An *nxn* symmetric real matrix ***A*** is said to be nega**tive-semidefinite if** $\boldsymbol{x^T A x \leq 0}$ for all non-zero x$\in R^n$

Example 1: $A = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$

Example 2: $A = \begin{bmatrix} 9 & -15 \\ -15 & 25 \end{bmatrix}$

Positive/Negative Definite Matrix

Definitions:

1) An *nxn* symmetric real matrix ***A*** is said to be **positive-definite if $x^T Ax > 0$** for all non-zero x$\in R^n$

2) An *nxn* symmetric real matrix ***A*** is said to be **positive-semidefinite if $x^T Ax \geq 0$** for all non-zero x$\in R^n$

3) An *nxn* symmetric real matrix ***A*** is said to be nega**tive-definite if $x^T Ax < 0$** for all non-zero x$\in R^n$

4) An *nxn* symmetric real matrix ***A*** is said to be nega**tive-definite if $x^T Ax \leq 0$** for all non-zero x$\in R^n$

Theory:

Let A be *nxn* symmetric real matrix ***A.*** All eigenvalues of A are real.

1) **A** is positive definite if and only if all of its eigenvalues are positive
2) A is positive semi-definite if and only if all of its eigenvalues are non-negative.
3) A is negative definite if and only if all of its eigenvalues are negative
4) A is negative semi-definite if and only if all of its eigenvalues are non-positive.
5) A is indefinite if and only if it has both positive and negative eigenvalues.

Positive definiteness

- **Test 1:** A matrix **A** will be positive definite if all its eigenvalues are positive; that is, all the values of λ that satisfy the determinental equation

$$\left|\mathbf{A} - \lambda \mathrm{I}\right| = 0$$

should be positive.

Negative definiteness

- Equivalently, a matrix is **negative-definite** if all its **eigenvalues** are **negative**
- It is **positive-semidefinite** if all its **eigenvalues are all greater than or equal to zero**
- It is **negative-semidefinite** if all its **eigenvalues are all less than or equal to zero**

Example: $A = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$

Example: $A = \begin{pmatrix} 6 & 5 & 12 \\ 5 & 19 & 0 \\ 12 & 3 & 7 \end{pmatrix}$

Positive definiteness

- **Test 2:** Another test that can be used to find the positive definiteness of a matrix **A** of order n involves evaluation of the determinants

$$A = |a_{11}|$$

$$A_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$$

$$A_3 = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

$$A_3 = \begin{vmatrix} a_{11} & a_{12} & a_{13} \cdots a_{1n} \\ a_{21} & a_{22} & a_{23} \cdots a_{2n} \\ a_{31} & a_{32} & a_{33} \cdots a_{3n} \\ \vdots & & \\ a_{n1} & a_{n2} & a_{n3} \cdots a_{nn} \end{vmatrix}$$

- The matrix **A** will be **positive definite** if and only if all the values $A_1, A_2, A_3, \ldots A_n$ are positive
- The matrix **A** will be **negative definite** if and only if the sign of A_j is $(-1)^j$ for $j=1,2,\ldots,n$
- If some of the A_j are positive and the remaining A_j are zero, the matrix **A** will be **positive semidefinite**

Example:

$$A = \begin{pmatrix} 6 & 5 & 12 \\ 5 & 19 & 0 \\ 12 & 3 & 7 \end{pmatrix}$$

A???

Inner Product and Positive Definite Matrix

- Consider an n-dimensional vector space V with an inner product $\langle \cdot, \cdot \rangle$ and an ordered basis $B = (\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n)$ 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, $i, j = 1, \ldots, n$.

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

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.

- Then, if $\forall \boldsymbol{x} \in V \setminus \{0\} : \boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{x} > 0$ (i.e., $\boldsymbol{A}$ is symmetric, positive definite), $\hat{x}^\top \boldsymbol{A}\hat{y}$ legitimately defines an inner product (w.r.t. B)

- Properties
  - The kernel of $\boldsymbol{A}$ is only $\{0\}$, because $\boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{x} > 0$ 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 $a_{ii} = \boldsymbol{e}_i{}^\top \boldsymbol{A}\boldsymbol{e}_i > 0$.

Length

- Inner product naturally induces a norm by defining:

$$\|x\| := \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$$

- Not every norm is induced by an inner product

- Cachy-Schwarz inequality. For the induced norm by the inner product,

$$|\langle \mathbf{x}, \mathbf{y} \rangle| \leq \|\mathbf{x}\| \; \|\mathbf{y}\|$$

Distance

- Now, we can introduce a notion of distance using a norm as:

  Distance. $d(\mathbf{x}, \mathbf{y}) := \|\mathbf{x} - \mathbf{y}\| = \sqrt{\langle \mathbf{x} - \mathbf{y}, \mathbf{x} - \mathbf{y} \rangle}$

- 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*. $d(\mathbf{x}, \mathbf{y}) \geq 0$ for all $\mathbf{x}, \mathbf{y}$ and $d(\mathbf{x}, \mathbf{y}) = 0 \iff \mathbf{x} = \mathbf{y}$
  - *Symmetric*. $d(\mathbf{x}, \mathbf{y}) = d(\mathbf{y}, \mathbf{x})$
  - *Triangle inequality*. $d(\mathbf{x}, \mathbf{z}) \leq d(\mathbf{x}, \mathbf{y}) + d(\mathbf{y}, \mathbf{z})$

Angle, Orthogonal, and Orthonormal

- Using C-S inequality,

$$-1 \leq \frac{\langle \boldsymbol{x}, \boldsymbol{y} \rangle}{\|\boldsymbol{x}\| \ \|\boldsymbol{y}\|} \leq 1$$

- 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 $\boldsymbol{y} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \boldsymbol{x}^{\mathsf{T}} \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} \boldsymbol{y},$$ 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

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

  - We can use $\boldsymbol{A}^{-1} = \boldsymbol{A}^{\top}$ for the definition of orthogonal matrices.
  - Fact 1. $\boldsymbol{A}, \boldsymbol{B}$: orthogonal $\implies$ $\boldsymbol{A}\boldsymbol{B}$: orthogonal
  - Fact 2. $\boldsymbol{A}$: orthogonal $\implies$ $\det(\boldsymbol{A}) = \pm 1$

Orthogonal Matrix

- The linear mapping Φ by orthogonal matrices preserve length and angle (for the dot product)

$$\|\Phi(\boldsymbol{A})\| = \|\boldsymbol{A}\boldsymbol{x}\|^2 = (\boldsymbol{A}\boldsymbol{x})^\top(\boldsymbol{A}\boldsymbol{x}) = \boldsymbol{x}^\top\boldsymbol{A}^\top\boldsymbol{A}\boldsymbol{x} = \boldsymbol{x}^\top\boldsymbol{x} = \|\boldsymbol{x}\|^2$$

$$\cos\omega = \frac{(\boldsymbol{A}\boldsymbol{x})^\top(\boldsymbol{A}\boldsymbol{y})}{\|\boldsymbol{A}\boldsymbol{x}\|\ \|\boldsymbol{A}\boldsymbol{y}\|} = \frac{\boldsymbol{x}^\top\boldsymbol{A}^\top\boldsymbol{A}\boldsymbol{y}}{\sqrt{\boldsymbol{x}^\top\boldsymbol{A}^\top\boldsymbol{A}\boldsymbol{x}\boldsymbol{y}^\top\boldsymbol{A}^\top\boldsymbol{A}\boldsymbol{y}}} = \frac{\boldsymbol{x}^\top\boldsymbol{y}}{\|\boldsymbol{x}\|\ \|\boldsymbol{y}\|}$$

Orthonormal Basis

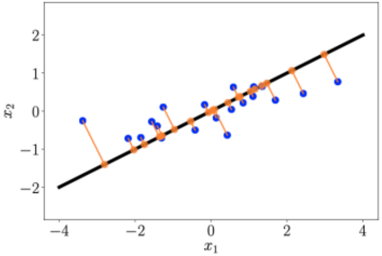
- Basis that is orthonormal, i.e., they are all orthogonal to each other and their lengths are 1.
- Standard basis in $\mathbb{R}^n$, $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$, is orthonormal.
- Question. How to obtain an orthonormal basis?

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 $(\boldsymbol{B}\boldsymbol{B}^\top | \boldsymbol{B})$

2. Constructive way: Gram-Schmidt process

Orthogonal Projections

• 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
• Example. Projection of 2D dataset onto 1D subspace

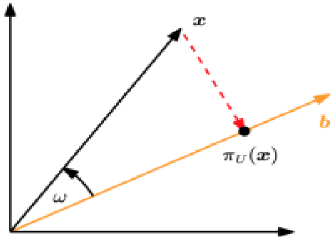


Projection onto Lines (1D Subspaces)

- 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)?

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

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



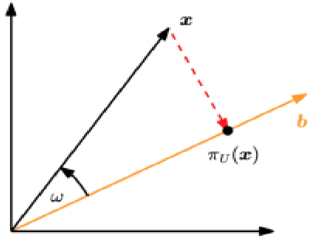
(a) Projection of $\boldsymbol{x} \in \mathbb{R}^2$ onto a subspace U with basis vector $\boldsymbol{b}$.

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?

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

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



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

Example

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

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

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

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

Projection onto General Subspace

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

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

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

- $\mathbb{R}^n \to m$-Dim, $(m < n)$
- A basis matrix

$$B = (\boldsymbol{b}_1, \cdots, \boldsymbol{b}_m) \in \mathbb{R}^{n \times m}$$

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

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

- $\lambda \in \mathbb{R}^1$ and $\boldsymbol{\lambda} \in \mathbb{R}^m$ are the coordinates in the projected spaces, respectively.
- $(\boldsymbol{B}^\top \boldsymbol{B})^{-1} \boldsymbol{B}^\top$ is called pseudo-inverse.

Example

- $U = \text{span}[\begin{pmatrix}1\\1\\1\end{pmatrix}, \begin{pmatrix}0\\1\\2\end{pmatrix}] \subset \mathbb{R}^3$ and $\boldsymbol{x} = \begin{pmatrix}6\\0\\0\end{pmatrix}$. Check that $\{(1\ \ 1\ \ 1)^\top, (0\ \ 1\ \ 2)^\top\}$ is a basis.

- Let $\boldsymbol{B} = \begin{pmatrix}1 & 0\\1 & 1\\1 & 2\end{pmatrix}$. Then, $\boldsymbol{B}^\top \boldsymbol{B} = \begin{pmatrix}1 & 1 & 2\\0 & 1 & 2\end{pmatrix}\begin{pmatrix}1 & 0\\1 & 1\\1 & 2\end{pmatrix} = \begin{pmatrix}3 & 3\\3 & 5\end{pmatrix}$

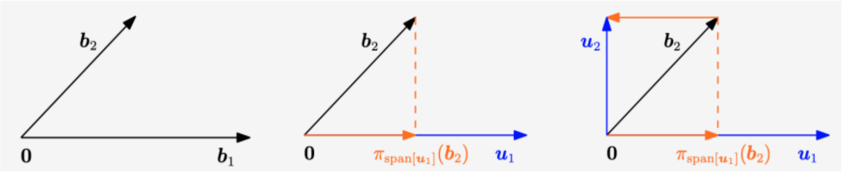
- Can see that $\boldsymbol{P}_\pi = \boldsymbol{B}(\boldsymbol{B}^\top \boldsymbol{B})^{-1}\boldsymbol{B}^\top = \frac{1}{6}\begin{pmatrix}5 & 2 & -1\\2 & 2 & 2\\-1 & 2 & 5\end{pmatrix}$, and

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

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

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



Example

- 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.