Linear Systems

Orthogonality

Sivakumar Balasubramanian

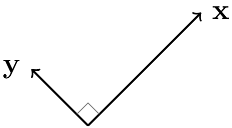
Department of Bioengineering
Christian Medical College, Bagayam
Vellore 632002

References

- S Boyd, Applied Linear Algebra: Chapters 5.
- G Strang, Linear Algebra: Chapters 3.

Orthogonality

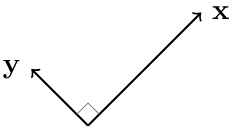
- Two vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ are orthognal if $\mathbf{x}^T\mathbf{y} = 0$.



- If we have a set of non-zero vectors $\mathcal{V} = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \ldots, \mathbf{v}_r\}$, we say this a set of mutually orthogonal vectors, if and only if, $\mathbf{v}_i^T\mathbf{v}_j = 0, \;\; 1 \leq i, j \leq r$ and $i \neq j$. $\mathcal{V}$ is also a linearly independent set of vectors.
- When the length of the vectors is 1, it is called an **orthonormal** set of vectors.
- A set of orthonormal vectors V also form an **orthonormal basis** of the subsape $span\left(\mathcal{V}\right)$.

Orthogonality

- Two vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ are orthognal if $\mathbf{x}^T\mathbf{y} = 0$.



- If we have a set of non-zero vectors $\mathcal{V} = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \ldots, \mathbf{v}_r\}$, we say this a set of mutually orthogonal vectors, if and only if, $\mathbf{v}_i^T\mathbf{v}_j = 0, \;\; 1 \leq i, j \leq r$ and $i \neq j$. $\mathcal{V}$ is also a linearly independent set of vectors.
- When the length of the vectors is 1, it is called an **orthonormal** set of vectors.
- A set of orthonormal vectors V also form an **orthonormal basis** of the subsapce $span\left(\mathcal{V}\right)$.

Is $\left\{\begin{bmatrix} 1 \\ -2 \\ 4 \end{bmatrix}, \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}\right\}$ an orthonormal set?. If no, how will you make it one?

Orthogonal Subspaces

- Two subspaces $\mathcal{V}, \mathcal{W}$ are orthogonal if every vector in one subspace is orthogonal to every vector in the other subspace.

$$\mathbf{v}^T \mathbf{w} = 0, \ \ \forall \mathbf{v} \in \mathcal{V} \text{ and } \forall \mathbf{w} \in \mathcal{W}$$

  Both subspaces $\mathcal{V}, \mathcal{W}$ are from the same space, e.g. $\mathbb{R}^n$

- Consider two subspaces $\mathcal{V}, \mathcal{W} \subset \mathbb{R}^n$, such that $\mathcal{V} + \mathcal{W} = \mathbb{R}^n$. If $\mathcal{V}$ and $\mathcal{W}$ are orthogonal subspaces, then $\mathcal{V}$ and $\mathcal{W}$ are **orthogonal complements** of each other.

$$\mathcal{W} \perp \mathcal{V} \ \ \rightarrow \ \ \mathcal{V}^\perp = \mathcal{W} \text{ or } \mathcal{W}^\perp = \mathcal{V}; \quad \left(\mathcal{V}^\perp\right)^\perp = \mathcal{V}$$

Orthogonal Subspaces

- Two subspaces $\mathcal{V}, \mathcal{W}$ are orthogonal if every vector in one subspace is orthogonal to every vector in the other subspace.

$$\mathbf{v}^T\mathbf{w} = 0, \ \forall \mathbf{v} \in \mathcal{V} \text{ and } \forall \mathbf{w} \in \mathcal{W}$$

Both subspaces $\mathcal{V}, \mathcal{W}$ are from the same space, e.g. $\mathbb{R}^n$

- Consider two subspaces $\mathcal{V}, \mathcal{W} \subset \mathbb{R}^n$, such that $\mathcal{V} + \mathcal{W} = \mathbb{R}^n$. If $\mathcal{V}$ and $\mathcal{W}$ are orthogonal subspaces, then $\mathcal{V}$ and $\mathcal{W}$ are **orthogonal complements** of each other.

$$\mathcal{W} \perp \mathcal{V} \ \rightarrow \ \mathcal{V}^\perp = \mathcal{W} \text{ or } \mathcal{W}^\perp = \mathcal{V}; \quad \left(\mathcal{V}^\perp\right)^\perp = \mathcal{V}$$

$\mathcal{V} = span\left\{\begin{bmatrix}1\\1\\1\end{bmatrix}^T, \begin{bmatrix}1\\0\\-1\end{bmatrix}^T\right\}$ and $\mathcal{W} = span\left\{\begin{bmatrix}1\\-2\\1\end{bmatrix}^T\right\}$. Is $\mathcal{V}^\perp = \mathcal{W}$? If we add $\begin{bmatrix}1\\-1\\0\end{bmatrix}^T$ to $\mathcal{W}$, is $\mathcal{V}^\perp = \mathcal{W}$ still true?

Relationship between the Four Fundamental Spaces

- $\mathcal{C}\left(\mathbf{A}\right), \mathcal{N}\left(\mathbf{A}^T\right) \subseteq \mathbb{R}^m$ are orthogonal complements.

$$\mathcal{C}\left(\mathbf{A}\right) \perp \mathcal{N}\left(\mathbf{A}^T\right)$$

- $\mathcal{C}\left(\mathbf{A}^T\right), \mathcal{N}\left(\mathbf{A}\right) \subseteq \mathbb{R}^n$ are orthogonal complements.

$$\mathcal{C}\left(\mathbf{A}^T\right) \perp \mathcal{N}\left(\mathbf{A}\right)$$

- $\dim \mathcal{C}\left(\mathbf{A}\right) + \dim \mathcal{N}\left(\mathbf{A}^T\right) = m \implies \mathcal{C}\left(\mathbf{A}\right) + \mathcal{N}\left(\mathbf{A}^T\right) = \mathbb{R}^m$
- $\dim \mathcal{C}\left(\mathbf{A}^T\right) + \dim \mathcal{N}\left(\mathbf{A}\right) = n \implies \mathcal{C}\left(\mathbf{A}^T\right) + \mathcal{N}\left(\mathbf{A}\right) = \mathbb{R}^n$

Relationship between the Four Fundamental Spaces

- $\mathcal{C}(\mathbf{A}), \mathcal{N}(\mathbf{A}^T) \subseteq \mathbb{R}^m$ are orthogonal complements.

$$\mathcal{C}(\mathbf{A}) \perp \mathcal{N}(\mathbf{A}^T)$$

- $\mathcal{C}(\mathbf{A}^T), \mathcal{N}(\mathbf{A}) \subseteq \mathbb{R}^n$ are orthogonal complements.

$$\mathcal{C}(\mathbf{A}^T) \perp \mathcal{N}(\mathbf{A})$$

- $\dim \mathcal{C}(\mathbf{A}) + \dim \mathcal{N}(\mathbf{A}^T) = m \implies \mathcal{C}(\mathbf{A}) + \mathcal{N}(\mathbf{A}^T) = \mathbb{R}^m$
- $\dim \mathcal{C}(\mathbf{A}^T) + \dim \mathcal{N}(\mathbf{A}) = n \implies \mathcal{C}(\mathbf{A}^T) + \mathcal{N}(\mathbf{A}) = \mathbb{R}^n$

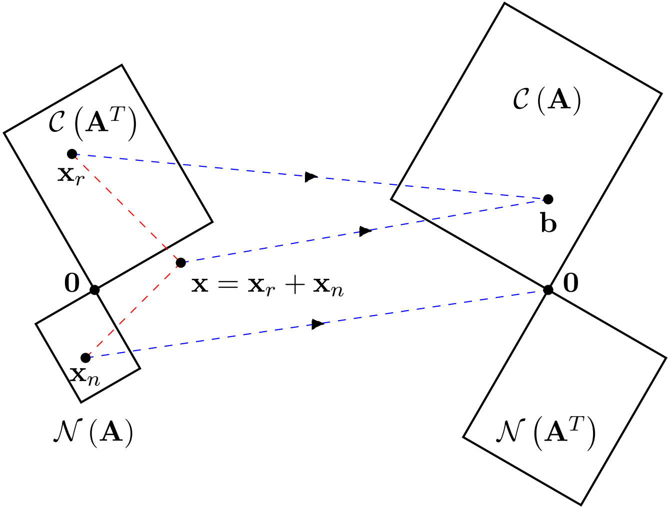
$$\mathbf{A} = \begin{bmatrix} 1 & -2 & 1 & 0 & 1 \\ 2 & -4 & 1 & -1 & -2 \\ -1 & 2 & 1 & 1 & 2 \\ 2 & -4 & -2 & -2 & -4 \end{bmatrix}$$

$$\mathbf{E} = \begin{bmatrix} \mathbf{E}_1 \\ \mathbf{E}_2 \end{bmatrix} = \left[\begin{array}{cccc} 1 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ -3 & 2 & 1 & 0 \\ \hline 0 & 0 & 2 & 1 \end{array}\right]$$

$$\mathbf{R} = \begin{bmatrix} \mathbf{R}_1 \\ \mathbf{R}_2 \end{bmatrix} = \left[\begin{array}{ccccc} 1 & -2 & 1 & 0 & 1 \\ 0 & 0 & -1 & -1 & -4 \\ 0 & 0 & 0 & -1 & -5 \\ \hline 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

- Is $\mathcal{C}(\mathbf{A}) \perp \mathcal{N}(\mathbf{A}^T)$?
- Is $\mathcal{C}(\mathbf{A}^T) \perp \mathcal{N}(\mathbf{A})$?
- What is $\dim \mathcal{C}(\mathbf{A})$, $\dim \mathcal{N}(\mathbf{A}^T)$, $\dim \mathcal{C}(\mathbf{A}^T)$, $\dim \mathcal{N}(\mathbf{A})$?

Relationship between the Four Fundamental Spaces



- $\mathbf{x}_r$ and $\mathbf{x}_n$ are the components of $\mathbf{x} \in \mathbb{R}^n$ in the row space and nullspace of $\mathbf{A}$.
- **Nullspace** $\mathcal{N}\left(\mathbf{A}\right)$ is mapped to $\mathbf{0}$.

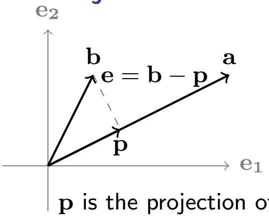
$$\mathbf{A}\mathbf{x}_n = \mathbf{0}$$

- **Row space** $\mathcal{C}\left(\mathbf{A}^T\right)$ is mapped to the **column space** $\mathcal{C}\left(\mathbf{A}\right)$.

$$\mathbf{A}\mathbf{x}_r = \mathbf{A}\left(\mathbf{x}_r + \mathbf{x}_n\right) = \mathbf{A}\mathbf{x} = \mathbf{b}$$

- The mapping from the **row space** to the **column space** is invertible, i.e. every $\mathbf{x}_r$ is mapped to a unique element in $\mathcal{C}\left(\mathbf{A}\right)$
- What sort of mapping does $\mathbf{A}^T$ do?

Orthogonal Projection onto Subspaces



$\mathbf{p}$ is the projection of $\mathbf{b}$ onto $\mathbf{a}$.

$\mathbf{P} = \frac{\mathbf{a}\mathbf{a}^T}{\mathbf{a}^T\mathbf{a}}$ is the projection matrix onto the line $\mathbf{a}$.

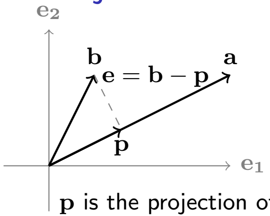
$\|\mathbf{e}\|$ is the distance of the point $\mathbf{b}$ from the line along $\mathbf{a}$. This distance is shortest when, $\mathbf{e} \perp \mathbf{a}$.

$$\mathbf{a}^T(\mathbf{b} - \mathbf{p}) = \mathbf{a}^T(\mathbf{b} - \alpha\mathbf{a}) = \mathbf{a}^T\mathbf{b} - \alpha\mathbf{a}^T\mathbf{a} = 0$$

$$\alpha = \frac{\mathbf{a}^T\mathbf{b}}{\mathbf{a}^T\mathbf{a}} \implies \mathbf{p} = \frac{\mathbf{a}^T\mathbf{b}}{\mathbf{a}^T\mathbf{a}}\mathbf{a}$$

$$\mathbf{p} = \frac{\mathbf{a}^T\mathbf{b}}{\mathbf{a}^T\mathbf{a}}\mathbf{a} = \mathbf{a}\frac{\mathbf{a}^T\mathbf{b}}{\mathbf{a}^T\mathbf{a}} = \frac{\mathbf{a}\mathbf{a}^T}{\mathbf{a}^T\mathbf{a}}\mathbf{b} = \mathbf{P}\mathbf{b}$$

Orthogonal Projection onto Subspaces



$\mathbf{p}$ is the projection of $\mathbf{b}$ onto $\mathbf{a}$.

$\|\mathbf{e}\|$ is the distance of the point $\mathbf{b}$ from the line along $\mathbf{a}$. This distance is shortest when, $\mathbf{e} \perp \mathbf{a}$.

$$\mathbf{a}^T(\mathbf{b} - \mathbf{p}) = \mathbf{a}^T(\mathbf{b} - \alpha\mathbf{a}) = \mathbf{a}^T\mathbf{b} - \alpha\mathbf{a}^T\mathbf{a} = 0$$

$$\alpha = \frac{\mathbf{a}^T\mathbf{b}}{\mathbf{a}^T\mathbf{a}} \implies \mathbf{p} = \frac{\mathbf{a}^T\mathbf{b}}{\mathbf{a}^T\mathbf{a}}\mathbf{a}$$

$$\mathbf{p} = \frac{\mathbf{a}^T\mathbf{b}}{\mathbf{a}^T\mathbf{a}}\mathbf{a} = \mathbf{a}\frac{\mathbf{a}^T\mathbf{b}}{\mathbf{a}^T\mathbf{a}} = \frac{\mathbf{a}\mathbf{a}^T}{\mathbf{a}^T\mathbf{a}}\mathbf{b} = \mathbf{P}\mathbf{b}$$

$\mathbf{P} = \frac{\mathbf{a}\mathbf{a}^T}{\mathbf{a}^T\mathbf{a}}$ is the projection matrix onto the line $\mathbf{a}$.

Find the orthogonal projection matrix associated $\mathbf{a}$, and find the projection of $\mathbf{b}$ on to $span(\{\mathbf{a}\})$.

- $\mathbf{a} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$; $\mathbf{b} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$
- $\mathbf{a} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$; $\mathbf{b} = \begin{bmatrix} 6 \\ 3 \end{bmatrix}$
- $\mathbf{a} = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$; $\mathbf{b} = \begin{bmatrix} -2 \\ -4 \\ -2 \end{bmatrix}$

Orthogonal Projection onto Subspaces

- We can also project vectors onto other subspaces, which is the generalization of the projection to a 1 dimensional subspace, i.e. the line.
- Consider a vector $\mathbf{b} \in \mathbb{R}^n$ and a subspace $\mathcal{S} \subseteq \mathbb{R}^n$ spanned by the orthonormal basis $\{\mathbf{u}_1, \mathbf{u}_2, \ldots \mathbf{u}_r\}$.
  $\mathbf{b}_{\mathcal{S}}$ – the orthogonal projection of $\mathbf{b}$ onto $\mathcal{S}$ is given by the following,
  $$\mathbf{b}_{\mathcal{S}} = \mathbf{U}\mathbf{U}^T\mathbf{b}; \;\; \mathbf{U} = \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \ldots & \mathbf{u}_r \end{bmatrix}$$
  $$\text{Projection matrix } \mathbf{P}_{\mathcal{S}} = \mathbf{U}\mathbf{U}^T$$
- A projection matrix is **idempotent**, i.e. $\mathbf{P}^2 = \mathbf{P}$. What does this mean in terms of projecting a vector on to a subspace?

Orthogonal Projection onto Subspaces

- We can also project vectors onto other subspaces, which is the generalization of the projection to a 1 dimensional subspace, i.e. the line.
- Consider a vector $\mathbf{b} \in \mathbb{R}^n$ and a subspace $\mathcal{S} \subseteq \mathbb{R}^n$ spanned by the orthonormal basis $\{\mathbf{u}_1, \mathbf{u}_2, \ldots \mathbf{u}_r\}$.
  $\mathbf{b}_{\mathcal{S}}$ – the orthogonal projection of $\mathbf{b}$ onto $\mathcal{S}$ is given by the following,
  $$\mathbf{b}_{\mathcal{S}} = \mathbf{U}\mathbf{U}^T\mathbf{b}; \;\; \mathbf{U} = \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \ldots & \mathbf{u}_r \end{bmatrix}$$
  $$\text{Projection matrix } \mathbf{P}_{\mathcal{S}} = \mathbf{U}\mathbf{U}^T$$
- A projection matrix is **idempotent**, i.e. $\mathbf{P}^2 = \mathbf{P}$. What does this mean in terms of projecting a vector on to a subspace?

Find the orthogonal projection matrix associated $\mathcal{U} = \left\{ \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} \right\}$, and find the projection of $\mathbf{b} = \begin{bmatrix} 2 \\ 2 \\ 3 \end{bmatrix}$ on to $span\left(\mathcal{U}\right)$.

Orthogonal Projection onto Subspaces

- Consider two matrices $\mathbf{U}_1, \mathbf{U}_2$ whose columns form an orthonormal basis of the subspace $\mathcal{S} \subseteq \mathbb{R}^m$, $\mathcal{C}(\mathbf{U}_1) = \mathcal{C}(\mathbf{U}_2)$.
- The projection matrix onto the subspace $\mathcal{S}$, $\mathbf{U}_1\mathbf{U}_1^T = \mathbf{U}_2\mathbf{U}_2^T$. We get the same projection matrix irrespective of which orthonormal basis one uses.

Orthogonal Projection onto Subspaces

- Consider two matrices $\mathbf{U}_1, \mathbf{U}_2$ whose columns form an orthonormal basis of the subspace $\mathcal{S} \subseteq \mathbb{R}^m$, $\mathcal{C}\left(\mathbf{U}_1\right) = \mathcal{C}\left(\mathbf{U}_2\right)$.
- The projection matrix onto the subspace $\mathcal{S}$, $\mathbf{U}_1\mathbf{U}_1^T = \mathbf{U}_2\mathbf{U}_2^T$. We get the same projection matrix irrespective of which orthonormal basis one uses.

Let $\mathbf{U}_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$ and $\mathbf{U}_1 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 0 & 0 \end{bmatrix}$. Find the corresponding projection matrices.

Orthogonal Projection onto Subspaces

- Two subspaces $\mathcal{V}, \mathcal{W} \subseteq \mathcal{V}$ are said to be **complementary subspaces** of $\mathcal{V}$, when

$$\mathcal{X} + \mathcal{Y} = \mathcal{V} \quad \text{and} \quad \mathcal{X} \cup \mathcal{Y} = \mathbf{0}$$

- When to subspaces $\mathcal{V}, \mathcal{W} \subseteq \mathbb{R}^m$ are complementary, then any vector $\mathbf{x} \in \mathbb{R}^m$ can be uniquely represented as $\mathbf{x} = \mathbf{v} + \mathbf{w}$, where $\mathbf{v} \in \mathcal{V}$ and $\mathbf{w} \in \mathcal{W}$ and $\mathbf{v}, \mathbf{w}$ are the components of $\mathbf{x}$ in $\mathcal{V}$ and $\mathcal{W}$ respectively.
- When $\mathcal{V} \perp \mathcal{W}$, then $\mathbf{v}^T\mathbf{w} = 0$; $\mathbf{v}, \mathbf{w}$ are orthogonal components.
- If $\mathbf{P}_{\mathcal{S}}$ is the orthogonal projection matrix onto $\mathcal{S}$, then what is the projection matrix onto $\mathcal{S}^{\perp}$?

Orthogonal Projection onto Subspaces

- Two subspaces $\mathcal{V}, \mathcal{W} \subseteq \mathcal{V}$ are said to be **complementary subspaces** of $\mathcal{V}$, when

$$\mathcal{X} + \mathcal{Y} = \mathcal{V} \quad \text{and} \quad \mathcal{X} \cup \mathcal{Y} = \mathbf{0}$$

- When to subspaces $\mathcal{V}, \mathcal{W} \subseteq \mathbb{R}^m$ are complementary, then any vector $\mathbf{x} \in \mathbb{R}^m$ can be uniquely represented as $\mathbf{x} = \mathbf{v} + \mathbf{w}$, where $\mathbf{v} \in \mathcal{V}$ and $\mathbf{w} \in \mathcal{W}$ and $\mathbf{v}, \mathbf{w}$ are the components of $\mathbf{x}$ in $\mathcal{V}$ and $\mathcal{W}$ respectively.
- When $\mathcal{V} \perp \mathcal{W}$, then $\mathbf{v}^T\mathbf{w} = 0$; $\mathbf{v}, \mathbf{w}$ are orthogonal components.
- If $\mathbf{P}_{\mathcal{S}}$ is the orthogonal projection matrix onto $\mathcal{S}$, then what is the projection matrix onto $\mathcal{S}^{\perp}$?

Let $\mathbf{u} = \begin{bmatrix} 1 & 1 \end{bmatrix}^T$. Find out the projection matrices $\mathbf{P}_{\mathbf{u}}$ and $\mathbf{P}_{\mathbf{u}^{\perp}}$? Verify that $\mathbf{P}_{\mathbf{u}^{\perp}} = \frac{\mathbf{u}^{\perp}\left(\mathbf{u}^{\perp}\right)^T}{\left(\mathbf{u}^{\perp}\right)^T\mathbf{u}^{\perp}}$.

Orthogonal Projection onto Subspaces

- An orthogonal projection matrix $\mathbf{P}_{\mathcal{S}}$ onto a subspace $\mathcal{S}$ represents a linear mapping, $\mathbf{P}_{\mathcal{S}} : \mathbb{R}^m \to \mathbb{R}^m$. What are the four fundamental subspaces of $\mathbf{P}_{\mathcal{S}}$?

$$\mathcal{C}\left(\mathbf{P}_{\mathcal{S}}\right) =$$

Orthogonal Projection onto Subspaces

- An orthogonal projection matrix $\mathbf{P}_{\mathcal{S}}$ onto a subspace $\mathcal{S}$ represents a linear mapping, $\mathbf{P}_{\mathcal{S}} : \mathbb{R}^m \to \mathbb{R}^m$. What are the four fundamental subspaces of $\mathbf{P}_{\mathcal{S}}$?

$$\mathcal{C}\left(\mathbf{P}_{\mathcal{S}}\right) = \ \mathcal{S}; \ \ \mathcal{N}\left(\mathbf{P}_{\mathcal{S}}\right) =$$

Orthogonal Projection onto Subspaces

- An orthogonal projection matrix $\mathbf{P}_{\mathcal{S}}$ onto a subspace $\mathcal{S}$ represents a linear mapping, $\mathbf{P}_{\mathcal{S}} : \mathbb{R}^m \to \mathbb{R}^m$. What are the four fundamental subspaces of $\mathbf{P}_{\mathcal{S}}$?

$$\mathcal{C}\left(\mathbf{P}_{\mathcal{S}}\right) = \mathcal{S};\ \mathcal{N}\left(\mathbf{P}_{\mathcal{S}}\right) = \mathcal{S}^{\perp}$$

$$\mathcal{N}\left(\mathbf{P}_{\mathcal{S}}^{T}\right) = \mathcal{S}^{\perp};\ \mathcal{C}\left(\mathbf{P}_{\mathcal{S}}^{T}\right) = \mathcal{S}$$

Orthogonal Projection onto Subspaces

- An orthogonal projection matrix $\mathbf{P}_{\mathcal{S}}$ onto a subspace $\mathcal{S}$ represents a linear mapping, $\mathbf{P}_{\mathcal{S}} : \mathbb{R}^m \to \mathbb{R}^m$. What are the four fundamental subspaces of $\mathbf{P}_{\mathcal{S}}$?

$$\mathcal{C}\left(\mathbf{P}_{\mathcal{S}}\right) = \mathcal{S};\ \mathcal{N}\left(\mathbf{P}_{\mathcal{S}}\right) = \mathcal{S}^{\perp}$$

$$\mathcal{N}\left(\mathbf{P}_{\mathcal{S}}^T\right) = \mathcal{S}^{\perp};\ \mathcal{C}\left(\mathbf{P}_{\mathcal{S}}^T\right) = \mathcal{S}$$

Let $\mathbf{U} = \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{-2}{\sqrt{6}} \end{bmatrix}$. Find the orthogonal projection matrix $\mathbf{P}_{\mathbf{U}}$ onto $\mathcal{C}\left(\mathbf{U}\right)$. Describe the four fundamental subspaces of $\mathbf{P}_{\mathbf{U}}$.

Orthogonal Projection onto Subspaces

- An orthogonal projection matrix $\mathbf{P}_{\mathcal{S}}$ onto a subspace $\mathcal{S}$ represents a linear mapping, $\mathbf{P}_{\mathcal{S}} : \mathbb{R}^m \to \mathbb{R}^m$. What are the four fundamental subspaces of $\mathbf{P}_{\mathcal{S}}$?

$$\mathcal{C}\left(\mathbf{P}_{\mathcal{S}}\right) = \mathcal{S};\ \mathcal{N}\left(\mathbf{P}_{\mathcal{S}}\right) = \mathcal{S}^{\perp}$$

$$\mathcal{N}\left(\mathbf{P}_{\mathcal{S}}^T\right) = \mathcal{S}^{\perp};\ \mathcal{C}\left(\mathbf{P}_{\mathcal{S}}^T\right) = \mathcal{S}$$

Let $\mathbf{U} = \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{-2}{\sqrt{6}} \end{bmatrix}$. Find the orthogonal projection matrix $\mathbf{P}_{\mathbf{U}}$ onto $\mathcal{C}\left(\mathbf{U}\right)$. Describe the four fundamental subspaces of $\mathbf{P}_{\mathbf{U}}$.

Now find $\mathbf{P}_{\mathbf{U}^{\perp}}$ and describe its four fundamental subspaces.

Gram-Schmidt Orthogonalization

- Given a linearly independent set of vectors $\mathcal{B} = \{\mathbf{x}_1, \mathbf{x}_2, \ldots \mathbf{x}_n\}$, where $\mathbf{x}_i \in \mathbb{R}^m, \;\; \forall i \in \{1, 2, \ldots n\}$, how can we find a orthonormal basis $\{\mathbf{u}_1, \mathbf{u}_2, \ldots \mathbf{u}_n\}$ for $span\,(\mathcal{B})$? $\longrightarrow$ **Gram-Schmidt Algorithm**
- Its an iterative procedure that can also detect if a given set $\mathcal{B}$ is linearly dependent.

Data: $\{\mathbf{x}_i\}_{i=1}^n$
Result: Return an orthonormal basis $\{\mathbf{u}_i\}_{i=1}^n$ if the set $\mathcal{B}$ is linearly independent, else return nothing.
for $i = 1, 2, \ldots n$ **do**
1. $\tilde{\mathbf{q}}_i = \mathbf{x}_i - \sum_{j=1}^{i-1} \left(\mathbf{u}_j^T \mathbf{x}_i\right) \mathbf{u}_i$ $\longrightarrow$(**Orthogonalization step**);
2. **If** $\tilde{\mathbf{q}}_i = 0$ **then return**;
3. $\mathbf{u}_i = \tilde{\mathbf{q}}_i / \left\|\tilde{\mathbf{q}}_i\right\|$ $\longrightarrow$(**Normalization step**);

end
return $\{\mathbf{u}_i\}_{i=1}^n$;

Gram-Schmidt Orthogonalization

- The algorithm can also be conveniently represented in a matrix form.

$$\mathcal{B} = \{\mathbf{a}_1, \mathbf{a}_2, \ldots \mathbf{a}_n\}$$

$$\text{Let } \mathbf{U}_1 = 0_{m \times 1} \quad \text{and} \quad \mathbf{U}_i = \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \ldots & \mathbf{u}_{i-1} \end{bmatrix} \in \mathbb{R}^{m \times (i-1)}$$

$$\mathbf{U}_i^T \mathbf{x}_i = \begin{bmatrix} \mathbf{u}_1^T \mathbf{x}_i \\ \mathbf{u}_2^T \mathbf{x}_i \\ \vdots \\ \mathbf{u}_{i-1}^T \mathbf{x}_i \end{bmatrix} \quad \text{and} \quad \mathbf{U}_i \mathbf{U}_i^T \mathbf{x}_i = \sum_{j=1}^{i-1} \left(\mathbf{u}_j^T \mathbf{x}_i\right) \mathbf{u}_j$$

$$\mathbf{u}_i = \frac{\left(\mathbf{I} - \mathbf{U}_i \mathbf{U}_i^T\right) \mathbf{x}_i}{\left\| \left(\mathbf{I} - \mathbf{U}_i \mathbf{U}_i^T\right) \mathbf{x}_i \right\|}$$

QR Decomposition

- Gram-Schmidt procedure leads us to another form of matrix decomposition – **QR decomposition**.
- Given a matrix $\mathbf{A} = \begin{bmatrix}\mathbf{a}_1 & \mathbf{a}_2 & \ldots & \mathbf{a}_n\end{bmatrix} \in \mathbb{R}^{n \times n}$, whose columns form a linearly independent set. Gramm-Schmidt algorithm produces a orthonormal basis $\{\mathbf{q}_1, \mathbf{q}_2, \ldots \mathbf{q}_n\}$ for $\mathcal{C}(\mathbf{A})$.

$$\mathbf{q}_1 = \frac{\mathbf{a}_1}{r_1} \quad \text{and} \quad \mathbf{q}_i = \frac{\mathbf{a}_i - \sum_{j=1}^{i-1}(\mathbf{q}_j^T\mathbf{a}_i)\mathbf{q}_j}{r_k}$$

where, $r_1 = \|\mathbf{a}_1\|$ and $r_k = \left\|\mathbf{a}_i - \sum_{j=1}^{i-1}(\mathbf{q}_j^T\mathbf{a}_i)\mathbf{q}_j\right\|$.

$$\mathbf{a}_1 = r_1\mathbf{q}_1 \quad \text{and} \quad \mathbf{a}_i = \sum_{j=1}^{i-1}\left(\mathbf{q}_j^T\mathbf{a}_i\right) + r_i\mathbf{q}_i$$

$$\mathbf{A} = \begin{bmatrix}\mathbf{a}_1 & \mathbf{a}_2 \ldots & \mathbf{a}_n\end{bmatrix} = \begin{bmatrix}\mathbf{q}_1 & \mathbf{q}_2 \ldots & \mathbf{q}_n\end{bmatrix}\begin{bmatrix} r_1 & \mathbf{q}_1^T\mathbf{a}_2 & \mathbf{q}_1^T\mathbf{a}_3 & \ldots & \mathbf{q}_1^T\mathbf{a}_n \\ 0 & r_2 & \mathbf{q}_2^T\mathbf{a}_3 & \ldots & \mathbf{q}_2^T\mathbf{a}_n \\ 0 & 0 & r_2 & \ldots & \mathbf{q}_3^T\mathbf{a}_n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & r_n \end{bmatrix} = \mathbf{QR}$$

QR Decomposition

Find the **QR** factorization for the following, if possible.

$$\mathbf{A} = \begin{bmatrix} -1 & 2 & 1 \\ 0 & 0 & 2 \\ -4 & 1 & 3 \end{bmatrix}$$

$$\mathbf{B} = \begin{bmatrix} 1 & 2 & -1 & -7 \\ 1 & 2 & 0 & -5 \\ -4 & 1 & 0 & -16 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

$$\mathbf{C} = \begin{bmatrix} 0 & 2 & -1 \\ 1 & 3 & -1 \\ -1 & 1 & 0 \\ 1 & 1 & 1 \\ 2 & 1 & 0 \end{bmatrix}$$

QR Decomposition

$$\mathbf{A} = \mathbf{QR}; \;\; \mathbf{A}, \mathbf{Q} \in \mathbb{R}^{m \times n}, \;\; \mathbf{R} \in \mathbb{R}^{n \times n}$$

- The columns of $\mathbf{Q}$ form an orthonormal basis for $\mathcal{C}\left(\mathbf{A}\right)$, and $\mathbf{R}$ is upper-triangular.
- Similar to $\mathbf{A} = \mathbf{LU}$, $\mathbf{A} = \mathbf{QR}$ can be used for used to solve $\mathbf{Ax} = \mathbf{b}$.

$$\mathbf{Ax} = \mathbf{QRx} = \mathbf{b} \implies \mathbf{Rx} = \mathbf{Q}^{-1}\mathbf{b} = \mathbf{Q}^{T}\mathbf{b}$$

QR Decomposition

$$\mathbf{A} = \mathbf{QR}; \;\; \mathbf{A}, \mathbf{Q} \in \mathbb{R}^{m \times n}, \;\; \mathbf{R} \in \mathbb{R}^{n \times n}$$

- The columns of $\mathbf{Q}$ form an orthonormal basis for $\mathcal{C}(\mathbf{A})$, and $\mathbf{R}$ is upper-triangular.
- Similar to $\mathbf{A} = \mathbf{LU}$, $\mathbf{A} = \mathbf{QR}$ can be used for used to solve $\mathbf{Ax} = \mathbf{b}$.

$$\mathbf{Ax} = \mathbf{QRx} = \mathbf{b} \implies \mathbf{Rx} = \mathbf{Q}^{-1}\mathbf{b} = \mathbf{Q}^T\mathbf{b}$$

Solve the following through $\mathbf{LU}$ and $\mathbf{QR}$ factorization.

$$\mathbf{Ax} = \begin{bmatrix} -1 & 2 & 1 \\ 0 & 0 & 2 \\ -4 & 1 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -1 \\ -2 \\ -2 \end{bmatrix} = \mathbf{b}$$