

Linear Control and Estimation

Orthogonality

Sivakumar Balasubramanian

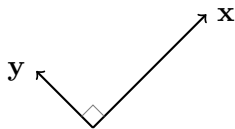
Department of Bioengineering
Christian Medical College, Bagayam
Vellore 632002

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 orthogonal if $\mathbf{x}^T \mathbf{y} = 0$.



- ▶ If we have a set of non-zero vectors $V = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \dots, \mathbf{v}_r\}$, we say this a set of mutually orthogonal vectors, if and only if,

$$\mathbf{v}_i^T \mathbf{v}_j = 0, \quad 1 \leq i, j \leq r \text{ and } i \neq j$$

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 subspace $\text{span}(V)$.

Orthogonal Subspaces

- ▶ Two subspaces V, W are orthogonal if every vector in one subspace is orthogonal to every vector in the other subspace.

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

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

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

$$W \perp V \rightarrow V^\perp = W \text{ or } W^\perp = V$$

W is set of all vector orthogonal.

- ▶ $(V^\perp)^\perp = V$

Relationship between the Four Fundamental Spaces

- ▶ Consider a matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$. $C(\mathbf{A})$, $C(\mathbf{A}^T)$, $N(\mathbf{A})$, and $N(\mathbf{A}^T)$ are the four fundamental subspaces.
- ▶ Column space and left nullspace of \mathbf{A} are orthogonal complements.

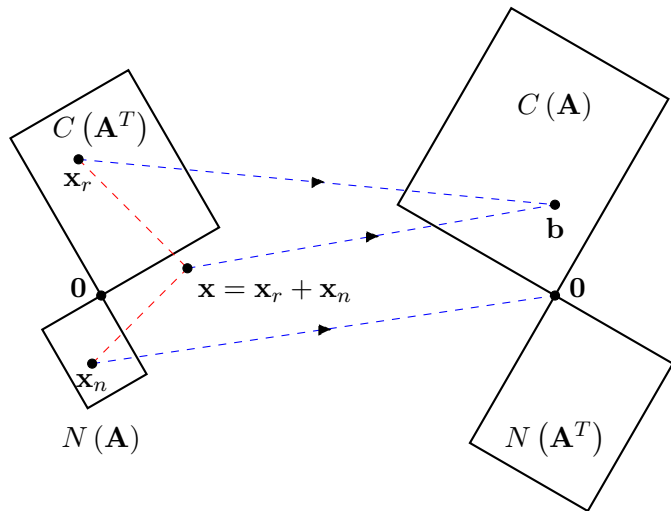
$$C(\mathbf{A}), N(\mathbf{A}^T) \subseteq \mathbb{R}^m \quad \text{and} \quad C(\mathbf{A}) \perp N(\mathbf{A}^T)$$

- ▶ Row space and nullspace of \mathbf{A} are orthogonal complements.

$$C(\mathbf{A}^T), N(\mathbf{A}) \subseteq \mathbb{R}^n \quad \text{and} \quad C(\mathbf{A}^T) \perp N(\mathbf{A})$$

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

Relationship between the Four Fundamental Spaces



- ▶ x_r and x_n are the components of $x \in \mathbb{R}^n$ in the row and nullspaces of A .

- ▶ **Nullspace** $N(A)$ is mapped to 0 .

$$Ax_n = 0$$

- ▶ **Row space** $C(A^T)$ is mapped to the **column space** $C(A)$.

$$Ax_r = A(x_r + x_n) = Ax = b$$

- ▶ The mapping from the **row space** to the **column space** is invertible, i.e. every x_r is mapped to a unique element in $C(A)$
- ▶ What sort of mapping does A^T do?

Gram-Schmidt Orthogonalization

- ▶ Given a linearly independent set of vectors $B = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$, where $\mathbf{x}_i \in \mathbb{R}^m$, $\forall i \in \{1, 2, \dots, n\}$, how can we find an orthonormal basis $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n\}$ for $\text{span}(B)$? \rightarrow **Gram-Schmidt Algorithm**
- ▶ Its an iterative procedure that can also detect if a given set 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 B is linearly independent, else return nothing.

for $i = 1, 2, \dots, n$ **do**

1. $\tilde{\mathbf{q}}_i = \mathbf{x}_i - \sum_{j=1}^{i-1} (\mathbf{u}_j^T \mathbf{x}_i) \mathbf{u}_j \rightarrow$ **(Orthogonalization step);**

2. **If** $\tilde{\mathbf{q}}_i = 0$ **then return;**

3. $\mathbf{u}_i = \tilde{\mathbf{q}}_i / \|\tilde{\mathbf{q}}_i\| \rightarrow$ **(Normalization step);**

end

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

Gram-Schmidt Orthogonalization

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

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

$$\text{Let } \mathbf{U}_1 = 0_{m \times 1} \quad \text{and} \quad \mathbf{U}_i = [\mathbf{u}_1 \quad \mathbf{u}_2 \quad \dots \quad \mathbf{u}_{i-1}] \in \mathbb{R}^{m \times (i-1)}$$

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

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

QR Decomposition

- ▶ Gram-Schmidt procedure leads us to another form of matrix decomposition – **QR decomposition**.
- ▶ Given a matrix $\mathbf{A} = [\mathbf{a}_1 \quad \mathbf{a}_2 \quad \dots \quad \mathbf{a}_n] \in \mathbb{R}^{n \times n}$, whose columns form a linearly independent set. Gram-Schmidt algorithm produces an orthonormal basis $\{\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_n\}$ for $C(\mathbf{A})$.

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

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

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

$$\mathbf{A} = [\mathbf{a}_1 \quad \mathbf{a}_2 \quad \dots \quad \mathbf{a}_n] = [\mathbf{q}_1 \quad \mathbf{q}_2 \quad \dots \quad \mathbf{q}_n] \begin{bmatrix} r_1 & \mathbf{q}_1^T \mathbf{a}_2 & \mathbf{q}_1^T \mathbf{a}_3 & \dots & \mathbf{q}_1^T \mathbf{a}_n \\ 0 & r_2 & \mathbf{q}_2^T \mathbf{a}_3 & \dots & \mathbf{q}_2^T \mathbf{a}_n \\ 0 & 0 & r_3 & \dots & \mathbf{q}_3^T \mathbf{a}_n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & r_n \end{bmatrix} = \mathbf{Q}\mathbf{R}$$

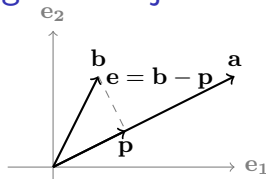
QR Decomposition

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

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

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

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} .

- ▶ 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 $S \subseteq \mathbb{R}^n$ spanned by the orthonormal basis $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_r\}$.

\mathbf{b}_S – the orthogonal projection of \mathbf{b} onto S is given by the following,

$$\mathbf{b}_S = \mathbf{U} \mathbf{U}^T \mathbf{b}; \quad \mathbf{U} = [\mathbf{u}_1 \quad \mathbf{u}_2 \quad \dots \quad \mathbf{u}_r]$$

Projection matrix $\mathbf{P}_S = \mathbf{U} \mathbf{U}^T$

- ▶ 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?