Linear Control and Estimation

Orthogonality

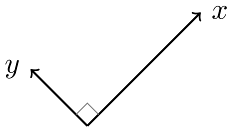
Sivakumar Balasubramanian

Department of Bioengineering
Christian Medical College, Bagayam
Vellore 632002

References

Orthogonality

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



- If we have a set of non-zero vectors $V = \{v_1, v_2, v_3, \ldots, v_r\}$, we say this a set of mutually orthogonal vectors, if and only if,

$$v_i^T v_j = 0, \;\; 1 \leq i, j \leq r \text{ and } i \neq j$$

  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 $span\,(V)$.

Orthogonal Subspaces

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

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

  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.

- $\left(V^\perp\right)^\perp = V$

Relationship between the Four Fundamental Spaces

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

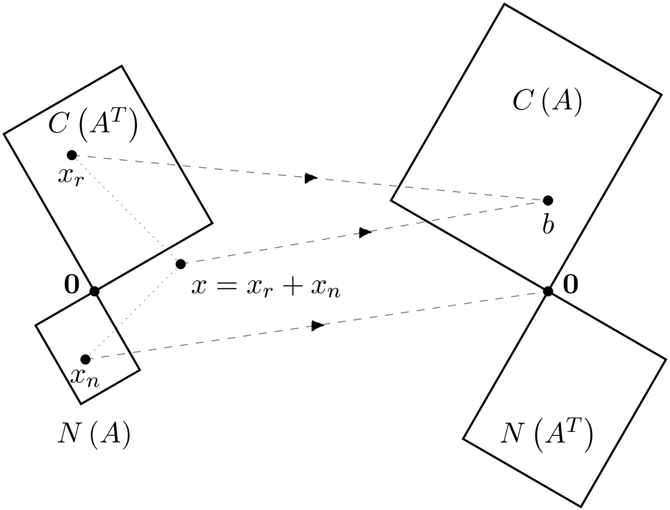
$$C(A), N\left(A^T\right) \subseteq \mathbb{R}^m \quad \text{and} \quad C(A) \perp N\left(A^T\right)$$

- Row space and nullspace of A are orthogonal complements.

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

- $\dim C(A) + \dim N\left(A^T\right) = m$
- $\dim C\left(A^T\right) + \dim N(A) = n$

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 $B = \{x_1, x_2, \ldots x_n\}$, where $x_i \in \mathbb{R}^m, \;\; \forall i \in \{1, 2, \ldots n\}$, how can we find a orthonormal basis $\{u_1, u_2, \ldots u_n\}$ for the $span\,(B)$? $\longrightarrow$ **Gram-Schmidt Algorithm**
- Its an iterative procedure that can also detect if a given set B is linearly dependent.

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

end
return $\{u_i\}_{i=1}^n$;

Gram-Schmidt Orthogonalization

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

$$B = B = \{a_1, a_2, \dots a_n\}$$

$$\text{Let } U_1 = 0_{m\times 1} \quad \text{and} \quad U_i = \begin{bmatrix} u_1 & u_2 & \dots & u_{i-1} \end{bmatrix} \in \mathbb{R}^{m\times(i-1)}$$

$$U_i^T x_i = \begin{bmatrix} u_1^T x_i \\ u_2^T x_i \\ \vdots \\ u_{i-1}^T x_i \end{bmatrix} \quad \text{and} \quad U_i U_i^T x_i = \sum_{j=1}^{i-1} \left(u_j^T x_i\right) u_j$$

$$u_i = \frac{\left(I - U_i U_i^T\right) x_i}{\left\| \left(I - U_i U_i^T\right) x_i \right\|}$$

QR Decomposition

- Gram-Schmidt procedure leads us to another form of matrix decompostion – **QR decompostion**.
- Given a matrix $A = \begin{bmatrix} a_1 & a_2 & \ldots & a_n \end{bmatrix} \in \mathbb{R}^{n \times n}$, whose columns for a linearly independent set. Gramm-Schmidt algorithm produces a orthonormal basis $\{q_1, q_2, \ldots q_n\}$ for $C(A)$.

$$q_1 = \frac{a_1}{r_1} \quad \text{and} \quad q_i = \frac{a_i - \sum_{j=1}^{i-1} (q_j^T a_i) q_j}{r_k}$$

where, $r_1 = \|a_1\|$ and $r_k = \left\| a_i - \sum_{j=1}^{i-1} (q_j^T a_i) q_j \right\|$.

$$a_1 = r_1 q_1 \quad \text{and} \quad a_i = \sum_{j=1}^{i-1} \left(q_j^T a_i \right) + r_i q_i$$

$$A = \begin{bmatrix} a_1 & a_2 \ldots & a_n \end{bmatrix} = \begin{bmatrix} q_1 & q_2 \ldots & q_n \end{bmatrix} \begin{bmatrix} r_1 & q_1^T a_1 & q_1^T a_2 & \ldots & q_1^T a_n \\ 0 & r_2 & q_2^T a_2 & \ldots & q_2^T a_n \\ 0 & 0 & r_2 & \ldots & q_3^T a_n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & r_n \end{bmatrix} = QR$$

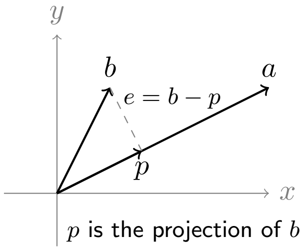
QR Decomposition

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

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

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

Orthogonal Projection onto Subspaces



p is the projection of b onto a.

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

$$a^T(b-p) = a^T(b-\alpha a) = a^Tb - \alpha a^Ta = 0$$

$$\alpha = \frac{a^Tb}{a^Ta} \implies p = \frac{a^Tb}{a^Ta}a$$

$$p = \frac{a^Tb}{a^Ta}a = a\frac{a^Tb}{a^Ta} = \frac{aa^T}{a^Ta}b = Pb$$

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

- We can also project vectors onto other subspaces, which is the generalization of the project to a 1 dimensional subspace, i.e. the line.
- Consider a vector $b \in \mathbb{R}^n$ and a subspace $S \subseteq \mathbb{R}^n$ spanned by the orthonormal basis $\{u_1, u_2, \dots u_r\}$. b_S – the orthogonal projection of b onto S is given by the following,

$$b_S = UU^Tb; \; U = \begin{bmatrix} u_1 & u_2 & \dots & u_r \end{bmatrix}$$

Projection matrix $P_S = UU^T$

- A projection matrix is **idempotent**, i.e. $P^2 = P$. What does this mean in terms of projecting a vector on to a subspace?