Orthogonality of the Four Subspaces

Rohit Budhiraja, IITK

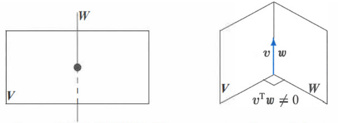
Applied Linear Algebra for Wireless Communications

Recap and agenda for today's class

- Discussed the following in last lecture
  - Linear independence, column and row spaces
- Discuss the concept of orthogonal subspaces,and projection today
  - Chapter 4.1 and 4.2 of the book

Orthogonal subspaces (1)

- Recall two vectors $\mathbf{v}$ and $\mathbf{w}$ are orthogonal when $\mathbf{v}^T\mathbf{w} = 0$
- Two subspaces $\mathbf{V}$ and $\mathbf{W}$ of a vector space are orthogonal if
  - $\mathbf{v}^T\mathbf{w} = 0$ for all $\mathbf{v}$ in $\mathbf{V}$ and $\mathbf{w}$ in $\mathbf{W}$



orthogonal plane V and line W non-orthogonal planes

- Orthogonality is impossible when dim $\mathbf{V}$ + dim $\mathbf{W}$ > dim (whole space)
- Every vector $\mathbf{x}$ in $N(A)$ is perpendicular to every row of A, because $A\mathbf{x} = \mathbf{0}$

$$Ax = \begin{bmatrix} \text{row } 1 \\ \vdots \\ \text{row } m \end{bmatrix} \begin{bmatrix} x \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} \quad \begin{matrix} \longleftarrow & (\text{row } 1) \cdot x \text{ is zero} \\ & \\ \longleftarrow & (\text{row } m) \cdot x \text{ is zero} \end{matrix}$$

Orthogonal subspaces (2)

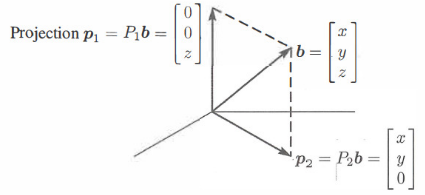
- Nullspace $N(A)$ and the row space $C(A^T)$ are orthogonal subspaces of $\mathbf{R}^n$
- Every vector $\mathbf{y}$ in $N(A^T)$ is perpendicular to every column of A

$$C(A) \perp N(A^{\mathrm{T}}) \qquad A^{\mathrm{T}}\boldsymbol{y} = \begin{bmatrix} (\textbf{column 1})^{\mathrm{T}} \\ \cdots \\ (\textbf{column } \boldsymbol{n})^{\mathrm{T}} \end{bmatrix} \begin{bmatrix} \boldsymbol{y} \end{bmatrix} = \begin{bmatrix} 0 \\ \cdot \\ 0 \end{bmatrix}$$

- Left nullspace $N(A^T)$ and the column space $C(A)$ are orthogonal in $\mathbf{R}^m$
- Orthogonal complement of a subspace $\mathbf{V}$
  - Contains every vector that is perpendicular to $\mathbf{V}$, and denoted as $V^\perp$
- Nullspace is the orthogonal complement of the row space
- Left nullspace is the orthogonal complement of the column space

Projections (1)

- What are the projections of $\mathbf{b} = (x, y, z)$ onto the z axis and the xy plane?



- What matrices P_1 and P_2 produce those projections onto a line and a plane

$$\text{Onto the } z \text{ axis} : P_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{Onto the } xy \text{ axis} : P_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Projections (2)

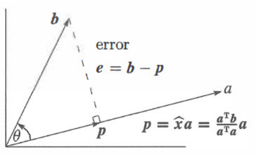
- Projections on z axis and xv plane are given respectively as

$$\mathbf{p}_1 = P_1\mathbf{b} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ z \end{bmatrix} \mathbf{p}_2 = P_2\mathbf{b} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x \\ y \\ 0 \end{bmatrix}$$

- The vectors give $\mathbf{p}_1 + \mathbf{p}_2 = \mathbf{b}$. Matrices give $P_1 + P_2 = \mathbf{I}$
- More than just orthogonal, line and plane are orthogonal complements
  - Their dimensions add to $1 + 2 = 3$
- Every vector $\mathbf{b}$ in the whole space is the sum of its parts in the two subspaces

Projection Onto a Line (1)

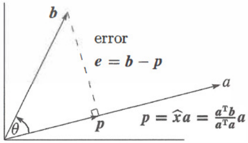
- To project any **b** onto a line



- A line goes through the origin in the direction of $\mathbf{a} = (a_1, \ldots, a_m)$
- Along that line, we want the point **p** closest to $\mathbf{b} = (b_1, \ldots, b_m)$
- Projection **p** is a multiple of **a** i.e., $\mathbf{p} = \hat{x}\mathbf{a}$, which we need to calculate
- Line from **b** to **p** is orthogonal to the vector **a**
  - This is the dotted line marked $\mathbf{e} = \mathbf{b} - \mathbf{p} = \mathbf{b} - \hat{x}\mathbf{a}$

Projection Onto a Line (2)

- To project any **b** onto a line



$$
\begin{aligned}
\mathbf{a}^T(\mathbf{b} - \hat{x}\mathbf{a}) &= 0 \Rightarrow \mathbf{a}^T\mathbf{b} - \hat{x}\mathbf{a}^T\mathbf{a} = 0 \\
\hat{x} &= \frac{\mathbf{a}^T\mathbf{b}}{\mathbf{a}^T\mathbf{a}} \\
\mathbf{p} &= \hat{x}\mathbf{a} = \mathbf{a}\hat{x} = \frac{\mathbf{a}\mathbf{a}^T\mathbf{b}}{\mathbf{a}^T\mathbf{a}} = P\mathbf{b}
\end{aligned}
$$

- Here $P = \frac{\mathbf{a}\mathbf{a}^T}{\mathbf{a}^T\mathbf{a}}$ is the projection matrix

Projection Onto a Subspace (1)

- Project $\mathbf{b}$ to a space spanned by n line. ind. vectors $\mathbf{a}_1, \ldots, \mathbf{a}_n$ in $\mathbf{R}^m$
  - Find the combination $\mathbf{p} = \hat{x}_1\mathbf{a}_1 + \cdots + \hat{x}_n\mathbf{a}_n$ closest to a given vector $\mathbf{b}$
- If $A = [\mathbf{a}_1 \ldots, \mathbf{a}_n]$ then $\mathbf{p} = A\hat{\mathbf{x}}$ where $\hat{\mathbf{x}} = (\hat{x}_1, \ldots, \hat{x}_n)$
- Error vector $(\mathbf{b} - A\hat{\mathbf{x}})$now should be orthogonal to each vector $\mathbf{a}_n$

$$\begin{array}{c} \mathbf{a}_1^T(\mathbf{b} - A\widehat{\mathbf{x}}) = 0 \\ \vdots \\ \mathbf{a}_n^T(\mathbf{b} - A\widehat{\mathbf{x}}) = 0 \end{array} \quad \text{or} \quad \begin{bmatrix} -\mathbf{a}_1^T- \\ \vdots \\ -\mathbf{a}_n^T- \end{bmatrix} \begin{bmatrix} \mathbf{b} - A\widehat{\mathbf{x}} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \end{bmatrix}$$

- We accordingly have

$$\begin{aligned} A^T(\mathbf{b} - A\widehat{\mathbf{x}}) &= \mathbf{0} \Rightarrow A^TA\widehat{\mathbf{x}} = A^T\mathbf{b} \Rightarrow \widehat{\mathbf{x}} = (A^TA)^{-1}A^T\mathbf{b} \\ \mathbf{p} = A\widehat{\mathbf{x}} &= A(A^TA)^{-1}A^T\mathbf{b} \Rightarrow \mathbf{p} = P\mathbf{b} \text{ where } P = A(A^TA)^{-1}A^T \end{aligned}$$

Example of projection calculation (1)

- If $A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix}$ and $\mathbf{b} = \begin{bmatrix} 6 \\ 0 \\ 0 \end{bmatrix}$ then calculate $\hat{\mathbf{x}}, \mathbf{p}$ and P
- We have

$$A^T A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 3 & 3 \\ 3 & 5 \end{bmatrix} \text{ and } A^T \mathbf{b} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} 6 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 6 \\ 0 \end{bmatrix}$$

- Recall that we have $A^T A \widehat{\mathbf{x}} = A^T \mathbf{b}$

$$\begin{bmatrix} 3 & 3 \\ 3 & 5 \end{bmatrix} \begin{bmatrix} \widehat{x_1} \\ \widehat{x_2} \end{bmatrix} = \begin{bmatrix} 6 \\ 0 \end{bmatrix} \text{ gives } \widehat{\mathbf{x}} = \begin{bmatrix} \widehat{x_1} \\ \widehat{x_2} \end{bmatrix} = \begin{bmatrix} 5 \\ -3 \end{bmatrix}$$

Example of projection calculation (2)

- We have $p = \hat{x}_1\mathbf{a}_1 + \hat{x}_2\mathbf{a}_2$

$$\mathbf{p} = 5\begin{bmatrix}1\\1\\1\end{bmatrix} - 3\begin{bmatrix}0\\1\\2\end{bmatrix} = \begin{bmatrix}5\\2\\-1\end{bmatrix} \text{ The error is } \mathbf{e} = \mathbf{b} - \mathbf{p} = \begin{bmatrix}1\\-2\\1\end{bmatrix}$$

- We can calculate

$$(A^TA)^{-1} = \frac{1}{6}\begin{bmatrix}5 & -3\\-5 & 3\end{bmatrix} \text{ which gives } P = \frac{1}{6}\begin{bmatrix}5 & 2 & -1\\2 & 2 & 2\\-1 & 2 & 5\end{bmatrix}$$

Some facts about projection matrix P

- Matrix $P = A(A^T A)^{-1} A^T$ is decepetive
- If we try to split $(A^T A)^{-1}$ into A^{-1} times $(A^T)^{-1}$
- If you make that mistake, and then $P = AA^{-1}(A^T)^{-1}A^T = I$
  - This is wrong because matrix A is rectangular, and it has no inverse
- $A^T A$ is invertible if and only if A has linearly independent columns