MA-106 Linear Algebra

M.K. Keshari

Department of Mathematics
Indian Institute of Technology Bombay
Powai, Mumbai - 76

2nd February, 2015
D1 - Lecture 12

• If W is a subspace of $\mathbb{R}^n$, then its orthogonal complement $W^\perp = N(A)$, where A is the matrix whose rows are basis vectors of W.

• If $x \in \mathbb{R}^n$, then x can be written uniquely as $x = w + w'$, where $w \in W$ and $w' \in W^\perp$.

• If $\{w_1, w_2\}$ are basis of W, then an orthogonal basis of W is given by $\{w_1, w_2 - \text{proj}_{w_1} w_2\}$, where $\text{proj}_{w_1} w_2 = \dfrac{w_1^T w_2}{w_1^T w_1} w_1$.

• If the system $Ax = b$ is inconsistent, then least square solution $\hat{x}$ is given by $A^T A \hat{x} = A^T b$.

• If $(A^T A)$ is invertible, then $\hat{x} = (A^T A)^{-1} A^T b$ and $\hat{b} = A\hat{x} = A(A^T A)^{-1} A^T b$ is the projection of b on $C(A)$.

• The matrix P is called a **projection** matrix if P is symmetric and $P^2 = P$. Pb is the projection of b on to the column space of P.

If $P = A(A^T A)^{-1} A^T$, then P is a projection matrix.

- $\boxed{N(A^TA) = N(A)}$.

Clearly, $N(A) \subset N(A^TA)$. For converse, take $x \in N(A^TA)$.

$$\begin{aligned} \text{Now } A^TAx = 0 &\Rightarrow x^T(A^TAx) = (Ax)^T(Ax) = \|Ax\|^2 = 0 \\ &\Rightarrow Ax = 0 \\ &\Rightarrow x \in N(A). \end{aligned}$$

$\square$

- If A is $m \times n$, then A^TA is $n \times n$. Since $N(A) = N(A^TA)$. By rank nullity theorem, $\text{rank}(A) = \text{rank}(A^TA) = n - \dim N(A)$.

- A^TA is invertible $\iff$ $\text{rank}(A^TA) = \text{rank}(A) = n$.

- If columns are A are linearly independent, then least square solution of $Ax = b$ is given by $A^TA\hat{x} = A^Tb \Rightarrow$ $\boxed{\hat{x} = (A^TA)^{-1}A^Tb}$ and orthogonal projection of b on $C(A)$ is $\boxed{A\hat{x} = A(A^TA)^{-1}A^Tb}$.

Example - Best Straight-line fit

Question 1. We want to find the best straight line $b = C + Dt$ which fits the data and gives least square error.
$(t, b) = (-2, 4), (-1, 3), (0, 1), (2, 0)$.

We want to solve $C - 2D = 4,\ C - D = 3,\ C = 1,\ C + 2D = 0$.

The system $\begin{pmatrix} 1 & -2 \\ 1 & -1 \\ 1 & 0 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} C \\ D \end{pmatrix} = \begin{pmatrix} 4 \\ 3 \\ 1 \\ 0 \end{pmatrix}$, $\quad Ax = b$ is inconsistent.

Find the least square solution by solving $A^T A \hat{x} = A^T b$.

Question 2. Find the best quadratic curve $b = C + Dt + Et^2$ which fits the above data and gives least square error.

We want to solve $C - 2D + 4E = 4,\ C - D + E = 3,\ C = 1,$
$C + 2D + 4E = 0$.
Write it as $Ax = b$ and find the least square solution $\hat{x}$.

Orthogonal columns

If the columns of $m \times n$ matrix A are linearly independent and orthogonal, then computation in least square problem becomes easy.

Assume v_1, v_2 are orthogonal and $A = (v_1 \quad v_2)$. Then

$$A^T A = \begin{pmatrix} v_1^T \\ v_2^T \end{pmatrix} \begin{pmatrix} v_1 & v_2 \end{pmatrix} = \begin{pmatrix} v_1^T v_1 & v_1^T v_2 \\ v_2^T v_1 & v_2^T v_2 \end{pmatrix} = \begin{pmatrix} \|v_1\|^2 & 0 \\ 0 & \|v_2\|^2 \end{pmatrix}.$$

Example: $\begin{pmatrix} -1 & 2 \\ 1 & 0 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$

This is consistent only when $a = c - 2b$.
Since the column vectors are orthogonal, $\hat{x} = (A^T A)^{-1} A^T b$

$$\begin{pmatrix} \hat{x_1} \\ \hat{x_2} \end{pmatrix} = \begin{pmatrix} 3 & 0 \\ 0 & 8 \end{pmatrix}^{-1} \begin{pmatrix} -1 & 1 & 1 \\ 2 & 0 & 2 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} \frac{1}{3}(-a+b+c) \\ \frac{1}{4}(a+c) \end{pmatrix}$$

Orthogonal Matrix

If A is $m \times n$ matrix whose column vectors ($\in \mathbb{R}^m$) form an *orthonormal set*, then

$$A^T A = \begin{pmatrix} v_1^T \\ \vdots \\ v_n^T \end{pmatrix} \begin{pmatrix} v_1 & \dots & v_n \end{pmatrix} = \begin{pmatrix} v_1^T v_1 & \dots & v_1^T v_n \\ \vdots & & \vdots \\ v_n^T v_1 & \dots & v_n^T v_n \end{pmatrix} = I_m.$$

Definition. A square matrix A whose column vectors form an orthonormal set is called an **orthogonal** matrix.

- If Q is an orthogonal matrix, then
  - $Q^T Q = I = QQ^T$. (since Q^T is the inverse of Q)
  - $\|Qx\| = \sqrt{(Qx)^T(Qx)} = \sqrt{x^T Q^T Qx} = \sqrt{x^T x} = \|x\|$.
  - Row vectors of Q are orthonormal.
    It follows from $QQ^T = I$.

Examples of Orthogonal matrices:

1. $$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}.$$
2. $$\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$
3. $$\frac{1}{3}\begin{pmatrix} 2 & 1 & -2 \\ 2 & -2 & 1 \\ 1 & 2 & 2 \end{pmatrix}.$$
4. $$\frac{1}{2}\begin{pmatrix} 1 & -1 & -1 & 1 \\ 1 & 1 & 1 & 1 \\ -1 & -1 & 1 & 1 \\ -1 & 1 & -1 & 1 \end{pmatrix}.$$

Gram-Schmidt Process

If the set of vectors $v_1, \dots, v_r$ in $\mathbb{R}^n$ are linearly independent, then we can find an orthonormal set of vectors $q_1, \dots, q_r$ such that $\text{Span}\{v_1, \dots, v_r\} = \text{Span}\{q_1, \dots, q_r\}$.

First find an othogonal set. Let $\boxed{w_1 = v_1}$, $\boxed{w_2 := v_2 - \text{proj}_{w_1} v_2}$. Then $w_1 \perp w_2$ and $\text{Span}\{v_1, v_2\} = \text{Span}\{w_1, w_2\}$.

Let $c_1 w_1 + c_2 w_2$ be the projection of v_3 on $\text{Span}\{w_1, w_2\}$. Then $v_3 - c_1 w_1 - c_2 w_2$ is orthogonal to $\text{Span}\{w_1, w_2\}$. Therefore $w_1^T(v_3 - c_1 w_1 - c_2 w_2) = 0$ gives $c_1 = \dfrac{w_1^T v_3}{\|w_1\|^2}$.

Similarly $c_2 = \dfrac{w_2^T v_3}{\|w_2\|^2}$. Therefore projection of v_3 on the space $\text{Span}\{w_1, w_2\}$ is $\text{proj}_{w_1} v_3 + \text{proj}_{w_2} v_3$. Define

$$w_3 = v_3 - \text{proj}_{\text{Span}\{w_1, w_2\}} v_3 = v_3 - \frac{w_1^T v_3}{\|w_1\|^2} w_1 - \frac{w_2^T v_3}{\|w_2\|^2} w_2.$$

$\{w_1, w_2, w_3\}$: orthogonal set, $\text{Span}\{w_1, w_2, w_3\} = \text{Span}\{v_1, v_2, v_3\}$.

By induction,

$$w_r := v_r - \frac{w_1^T v_r}{\|w_1\|^2} w_1 - \frac{w_2^T v_r}{\|w_2\|^2} w_2 - \cdots - \frac{w_{r-1}^T v_r}{\|w_{r-1}\|^2} w_{r-1}$$

Then $\{w_1, \ldots, w_r\}$ is an orthogonal set and $\mathrm{Span}\{w_1, \ldots, w_r\} = \mathrm{Span}\{v_1, \ldots, v_r\}$.

Now take $q_1 = \dfrac{w_1}{\|w_1\|}$, $q_2 = \dfrac{w_2}{\|w_2\|}$, ..., $q_r = \dfrac{w_r}{\|w_r\|}$.

Then $\{q_1, \ldots, q_r\}$ is an orthonormal set and

$\mathrm{Span}\{v_1, v_2, \ldots, v_r\} = \mathrm{Span}\{q_1, q_2, \ldots, q_r\}$.

Example

Let $S = \left\{ v_1 = \begin{pmatrix} 3 \\ 1 \\ -1 \\ 3 \end{pmatrix}, v_2 = \begin{pmatrix} -5 \\ 1 \\ 5 \\ -7 \end{pmatrix}, v_3 = \begin{pmatrix} 1 \\ 1 \\ -2 \\ 8 \end{pmatrix} \right\}$

and $W = \text{Span}(S)$. Find an orthonormal basis for W.

Verify that $\{v_1, v_2, v_3\}$ are linearly independent.
Hence S is a basis of W.
For linear independence, check that rank of the matrix $\begin{pmatrix} v_1 & v_2 & v_3 \end{pmatrix}$ is 3.

Use Gram-Schmidt method, $w_1 = v_1$, $w_2 = v_2 - \dfrac{w_1^T v_2}{\|w_1\|^2} w_1$

$$= v_2 - \frac{-15+1-5-21}{9+1+1+9} w_1 = v_2 - \frac{-40}{20} w_1 = v_2 + 2w_1. = \begin{pmatrix} 1 \\ 3 \\ 3 \\ -1 \end{pmatrix}$$

Recall $w_1 = \begin{pmatrix} 3 \\ 1 \\ -1 \\ 3 \end{pmatrix}$, $w_2 = \begin{pmatrix} 1 \\ 3 \\ 3 \\ -1 \end{pmatrix}$, $v_3 = \begin{pmatrix} 1 \\ 1 \\ -2 \\ 8 \end{pmatrix}$.

(check $w_1^T w_2 = 0$).

Now $w_3 = v_3 - \dfrac{w_1^T v_3}{\|w_1\|^2} w_1 - \dfrac{w_2^T v_3}{\|w_2\|^2} w_2$.

$$w_3 = v_3 - \frac{3+1+2+24}{20} w_1 - \frac{1+3-6-8}{20} w_2$$

$$= \begin{pmatrix} 1 \\ 1 \\ -2 \\ 8 \end{pmatrix} - \frac{3}{2} \begin{pmatrix} 3 \\ 1 \\ -1 \\ 3 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} 1 \\ 3 \\ 3 \\ -1 \end{pmatrix} = \begin{pmatrix} -3 \\ 1 \\ 1 \\ 3 \end{pmatrix}$$

Check $w_1^T w_3 = 0$, $w_2^T w_3 = 0$.

Hence $\{w_1, w_2, w_3\}$ is an orthogonal basis of S. An orthonormal basis for S is $\{\frac{1}{\sqrt{20}} w_1, \frac{1}{\sqrt{20}} w_2, \frac{1}{\sqrt{20}} w_3\}$. $\square$