MATH 2111: Tutorial 13

T1A&T1B QUAN Xueyang
T1C&T2A SHEN Yinan
T2B&T2C ZHANG Fa

Department of Mathematics, HKUST

Review

- The Gram-Schmidt Process
- Least-Squares Problem
- Applications to Linear Models
- Diagonalization of Symmetric Matrices

Example 1

Let W be a subspace of $\mathbb{R}^n$, and let $\{\mathbf{u_1}, \mathbf{u_2}\}$ be an orthonormal basis for W. So for each $\mathbf{y} \in \mathbb{R}^n$, we have,

$$\mathrm{proj}_W(\mathbf{y}) = (\mathbf{y} \cdot \mathbf{u}_1)\mathbf{u}_1 + (\mathbf{y} \cdot \mathbf{u}_2)\mathbf{u}_2.$$

Based on this formula proving the following:
(1) $\mathrm{proj}_W : \mathbb{R}^n \to \mathbb{R}^n$ is a linear transformation.
(2) The kernel of proj_W is $W^\perp$.
(Hint: show that $\mathbf{y}$ is in the kernel of proj_W if and only if $\mathbf{y}$ is in $W^\perp$)
(3) What is proj_W^2? (Hint: if $\mathbf{y}$ is in W, $\mathrm{proj}_W(\mathbf{y}) = \mathbf{y}$)

(1) Prove by def:

① $\forall \vec{y}_1, \vec{y}_2 \in \mathbb{R}^n$,

$$proj_W(\vec{y}_1+\vec{y}_2) = ((\vec{y}_1+\vec{y}_2)\cdot\vec{u}_1)\vec{u}_1 + ((\vec{y}_1+\vec{y}_2)\cdot\vec{u}_2)\vec{u}_2$$
$$= \underline{(\vec{y}_1\cdot\vec{u}_1)\vec{u}_1} + \underline{(\vec{y}_2\cdot\vec{u}_1)\vec{u}_1} + \underline{(\vec{y}_1\cdot\vec{u}_2)}\vec{u}_2 + \underline{(\vec{y}_2\cdot\vec{u}_2)\vec{u}_2}$$
$$= \underline{proj_W(\vec{y}_1)} + \underline{proj_W(\vec{y}_2)}$$

② $\forall \vec{y} \in \mathbb{R}^n$, $c \in \mathbb{R}$.

$$proj_W(c\vec{y}) = (c\vec{y}\cdot\vec{u}_1)\vec{u}_1 + (c\vec{y}\cdot\vec{u}_2)\vec{u}_2$$
$$= c[(\vec{y}\cdot\vec{u}_1)\vec{u}_1 + (\vec{y}\cdot\vec{u}_2)\vec{u}_2]$$
$$= c\, proj_W(\vec{y})$$

Thus, $proj_W : \mathbb{R}^n \to \mathbb{R}^n$ is a linear transformation.

(2) ① If $\vec{y} \in W^\perp$, then

$\vec{y}\cdot\vec{u}_1 = 0$, $\vec{y}\cdot\vec{u}_2 = 0$

So, $proj_W(\vec{y}) = 0\vec{u}_1 + 0\vec{u}_2 = \vec{0}$

$\therefore \vec{y} \in kernel(proj_W)$.

② If $\vec{y} \in kernel(proj_W)$, i.e. $proj_W(\vec{y}) = \vec{0}$,

we have:

$$\vec{0} = (\vec{y}\cdot\vec{u}_1)\vec{u}_1 + (\vec{y}\cdot\vec{u}_2)\vec{u}_2,$$

Since $\{\vec{u}_1, \vec{u}_2\}$ is a basis for W, it is linearly independent,

$\therefore \vec{y}\cdot\vec{u}_1 = 0$, and $\vec{y}\cdot\vec{u}_2 = 0$.

$\therefore \vec{y} \in W^\perp$.

(3) For any $\vec{v} \in \mathbb{R}^n$, $proj_W^2(\vec{v}) = proj_W(\underbrace{proj_W(\vec{v})}_{\in W}) \underset{\uparrow}{=} proj_W(\vec{v})$.

for any $\vec{y} \in W$, $proj_W(\vec{y}) = \vec{y}$

Example 2

Find a QR factorization of the matrix.

$$\begin{bmatrix} 1 & 2 & 5 \\ -1 & 1 & -4 \\ -1 & 4 & -3 \\ 1 & -4 & 7 \\ 1 & 2 & 1 \end{bmatrix}$$

1. Find a basis of Col A.

Apply row operations (at least to REF(A)),

$$\begin{bmatrix} 1 & 2 & 5 \\ -1 & 1 & -4 \\ -1 & 4 & -3 \\ 1 & -4 & 7 \\ 1 & 2 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 5 \\ 0 & 3 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$\uparrow$ has 3 pivot columns

Thus, column vector of A are linearly independent, they form a basis for col A.

2. Find an orthonormal basis for Col A.

Apply the Gram-Schmidt process.

Denote

$$\vec{x_1} = \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \\ 1 \end{bmatrix} \qquad \vec{x_2} = \begin{bmatrix} 2 \\ 1 \\ 4 \\ -4 \\ 2 \end{bmatrix} \qquad \vec{x_3} = \begin{bmatrix} 5 \\ -4 \\ -3 \\ 7 \\ 1 \end{bmatrix}.$$

$$\vec{v_1} = \vec{x_1} = \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \\ 1 \end{bmatrix}, \qquad \vec{v_2} = \vec{x_2} - \frac{\vec{x_2}\cdot\vec{v_1}}{\vec{v_1}\cdot\vec{v_1}}\vec{v_1} = \vec{x_2} - \frac{-5}{5}\vec{v_1} = \begin{bmatrix} 3 \\ 0 \\ 3 \\ -3 \\ 3 \end{bmatrix}$$

$$\vec{v_3} = \vec{x_3} - \frac{\vec{x_3}\cdot\vec{v_1}}{\vec{v_1}\cdot\vec{v_1}}\vec{v_1} - \frac{\vec{x_3}\cdot\vec{v_2}}{\vec{v_2}\cdot\vec{v_2}}\vec{v_2}$$

$$= \vec{x_3} - \frac{20}{5}\vec{v_1} - \frac{-12}{36}\vec{v_2} = \begin{bmatrix} 2 \\ 0 \\ 2 \\ 2 \\ -2 \end{bmatrix}$$

Normalize $\{\vec{v_1}, \vec{v_2}, \vec{v_3}\}$, and take them as columns of Q:

$$Q = \begin{bmatrix} 1/\sqrt{5} & 1/2 & 1/2 \\ -1/\sqrt{5} & 0 & 0 \\ -1/\sqrt{5} & 1/2 & 1/2 \\ 1/\sqrt{5} & -1/2 & 1/2 \\ 1/\sqrt{5} & 1/2 & -1/2 \end{bmatrix}$$

(Q here is orthonormal, so $Q^TQ = I_3$)

3. If $A = QR$, now find R

Since $A = QR$,

$$Q^TA = Q^TQR = R$$

$$\therefore R = \begin{bmatrix} 1/\sqrt{5} & -1/\sqrt{5} & -1/\sqrt{5} & 1/\sqrt{5} & 1/\sqrt{5} \\ 1/2 & 0 & 1/2 & -1/2 & 1/2 \\ 1/2 & 0 & 1/2 & 1/2 & -1/2 \end{bmatrix} \begin{bmatrix} 1 & 2 & 5 \\ -1 & 1 & -4 \\ -1 & 4 & -3 \\ 1 & -4 & 7 \\ 1 & 2 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} \sqrt{5} & -\sqrt{5} & 4\sqrt{5} \\ 0 & 6 & -2 \\ 0 & 0 & 4 \end{bmatrix}$$

(4. Check if $A = QR$.)

Example 3

State whether each of the following statement is true or false. (If it is true, give a brief justification; if it is false, give a counterexample.)

(1) Let U be an orthogonal matrix. If $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is an orthogonal set in $\mathbb{R}^n$, then $\{U\mathbf{v}_1, U\mathbf{v}_2, U\mathbf{v}_3\}$ is also an orthogonal set.

(2) Let U and W be subspaces of $\mathbb{R}^n$, and $U \subseteq W$. Then $U^\perp \subseteq W^\perp$.

(3) If U is a square matrix with orthonormal columns, then the rows of U are also orthonormal.

(4) Suppose $\mathbf{x}_1, \mathbf{x}_2$, and $\mathbf{x}_3$ are vectors in $\mathbb{R}^n$. If $W = \text{Span}\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$ with $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$ linearly independent, and if $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is an orthogonal set in W, then $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is a basis for W.

(1) True.

U is orthogonal. we have $U\vec{v_i} \cdot U\vec{v_j} = \vec{v_i} \cdot \vec{v_j}$.

(2) False.

eg. $U = \text{span}\{\vec{e_1}\}$, $W = \text{span}\{\vec{e_1}, \vec{e_2}\}$ be subspaces of $\mathbb{R}^3$, so $U \subseteq W$.

However, $W^\perp = \text{span}\{\vec{e_3}\}$, $U^\perp = \text{span}\{\vec{e_2}, \vec{e_3}\}$, i.e. $W^\perp \subseteq U^\perp$.

Actually, if $U \subseteq W$, then $W^\perp \subseteq U^\perp$.

proof: $\forall\ \vec{v} \in W^\perp$, we have $\vec{v} \cdot \vec{p} = 0 \quad \forall\ \vec{p} \in W$

Since $U \subseteq W$,

$\therefore\ \vec{v} \cdot \vec{p} = 0$ for any $\vec{p} \in U$

$\therefore\ \vec{v} \in U^\perp$

Thus, $W^\perp \subseteq U^\perp$.

(3) True.

$\because$ Rows of U is the columns of U^T,

$\therefore$ it suffices showing U^T is an orthogonal matrix:

$$(U^T)^T U^T = UU^T = UU^{-1} = I.$$

(4) False.

$\{\vec{v_1}, \vec{v_2}, \vec{v_3}\}$ may contain the zero vector. In this case, can't be a basis.

Example 4

Given data $(x_1, y_1), \ldots, (x_n, y_n)$, for a least-squares problem $y = \hat{\beta}_0 + \hat{\beta}_1 x$. And $(\hat{\beta}_0, \hat{\beta}_1)$ is the least-squares solution to

$$\underset{X}{\begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix}} \underset{\beta}{\begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}.$$

From linear algebra perspective, prove the formula for regression coefficients $\hat{\beta}_0, \hat{\beta}_1$ from statistics:

$$\hat{\beta}_1 = \frac{\mathrm{SS}_{xy}}{\mathrm{SS}_{xx}}, \quad \hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$

where,

$$\mathrm{SS}_{xy} = \Sigma xy - \frac{(\Sigma x)(\Sigma y)}{n}, \quad \mathrm{SS}_{xx} = \Sigma xx - \frac{(\Sigma x)^2}{n}, \quad \bar{y} = \frac{(\Sigma y)}{n}, \quad \bar{x} = \frac{(\Sigma x)}{n}$$

Example 4 - Continued

(1) By considering the normal equations, find a matrix M such that $M\begin{bmatrix}\hat{\beta}_0\\ \hat{\beta}_1\end{bmatrix}=\begin{bmatrix}\Sigma y\\ \Sigma xy\end{bmatrix}$. (The entries of M will depend on n, x and y.)

(2) Assume that $x_1, \ldots, x_n$ are not all the same, explain why this indicates $\hat{\beta}_0$ and $\hat{\beta}_1$ are unique.

(3) The uniqueness of $\hat{\beta}_0$ and $\hat{\beta}_1$ implies that the matrix M is invertible. By inverting M, show that $\hat{\beta}_1$ has the formula given above, then show that $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1\bar{x}$.

(1) Normal equation : $X^T X \beta = X^T \vec{y}$, $\hat{\beta}_0, \hat{\beta}_1$ are solutions to this.

$$\begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_n \end{bmatrix} \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_n \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

$$\therefore \underbrace{\begin{bmatrix} n & \Sigma x \\ \Sigma x & \Sigma x^2 \end{bmatrix}}_{M} \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \end{bmatrix} = \begin{bmatrix} \Sigma y \\ \Sigma xy \end{bmatrix},$$

(2) By theorem, the solution is unique iff columns of A are linearly independent. From questions assumption, x_i's are not all the same,

$$\therefore \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \neq c \cdot \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \text{ for some } c \in \mathbb{R}.$$

(3) If invert M, we have :

$$\begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \end{bmatrix} = \begin{bmatrix} n & \Sigma x \\ \Sigma x & \Sigma x^2 \end{bmatrix}^{-1} \begin{bmatrix} \Sigma y \\ \Sigma xy \end{bmatrix}$$

$$= \frac{1}{n\Sigma x^2 - (\Sigma x)^2} \begin{bmatrix} \Sigma x^2 & -\Sigma x \\ -\Sigma x & n \end{bmatrix}^{-1} \begin{bmatrix} \Sigma y \\ \Sigma xy \end{bmatrix}$$

$$= \frac{1}{n\Sigma x^2 - (\Sigma x)^2} \begin{bmatrix} \Sigma x^2 \Sigma y - \Sigma x \Sigma xy \\ -\Sigma x \Sigma y + n \Sigma xy \end{bmatrix}$$

$$\therefore \hat{\beta}_1 = \frac{-\Sigma x \Sigma y + n\Sigma xy}{n\Sigma x^2 - (\Sigma x)^2} = \frac{\Sigma xy - \frac{\Sigma x \Sigma y}{n}}{\Sigma x^2 - (\Sigma x)^2/n} = \frac{SS_{xy}}{SS_{xx}}$$

$$\therefore \hat{\beta}_0 = \frac{\Sigma x^2 \Sigma y - \Sigma x \Sigma xy}{n\Sigma x^2 - (\Sigma x)^2} = \frac{\Sigma x^2 \frac{\Sigma y}{n} - \Sigma xy \frac{\Sigma x}{n}}{\Sigma x^2 - \frac{(\Sigma x)^2}{n}}$$

$$= \frac{\bar{y}\Sigma x^2 - \bar{x}\Sigma xy}{SS_{xx}}$$

$$= \frac{\bar{y}\left(\Sigma x^2 - \frac{(\Sigma x)^2}{n}\right) - \bar{x}\Sigma xy + \bar{y}\frac{(\Sigma x)^2}{n}}{SS_{xx}}$$

$$\bar{y} = \frac{\Sigma y}{n}$$

$$\frac{\bar{y}(\Sigma x)^2}{n} = \frac{\Sigma y}{n} \cdot \frac{\Sigma x \Sigma x}{n}$$

$$= \frac{\bar{y}\, SS_{xx} - \bar{x}\Sigma xy + \bar{x}\frac{\Sigma x \Sigma y}{n}}{SS_{xx}}$$

$$= \frac{\bar{y}\, SS_{xx} - \bar{x}\, SS_{xy}}{SS_{xx}}$$

$$= \bar{y} - \frac{SS_{xy}}{SS_{xx}}\bar{x}$$

$$= \bar{y} - \hat{\beta}_1 \bar{x}$$

Example 5

Let A be the symmetric matrix

$$A = \begin{bmatrix} 3 & -2 & 4 \\ -2 & 6 & 2 \\ 4 & 2 & 3 \end{bmatrix}$$

$$= \underbrace{\begin{bmatrix} -1 & 2 & 5 \\ 2 & 1 & 0 \\ 0 & -2 & 5 \end{bmatrix}}_{P} \underbrace{\begin{bmatrix} 7 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 7 \end{bmatrix}}_{D} \begin{bmatrix} -1 & 2 & 5 \\ 2 & 1 & 0 \\ 0 & -2 & 5 \end{bmatrix}^{-1}$$

Find an orthogonal matrix Q such that

$$A = Q \begin{bmatrix} -2 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & 7 \end{bmatrix} Q^{-1}.$$

Or explain why this Q doesn't exist.

From digonalization $A = PDP^{-1}$, we know :

when $\lambda = -2$, one corresponding eigenvector is :

$$\vec{v}_2 = \begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix}$$

when $\lambda = 7$, two corresponding linearly independent eigenvectors are :

$$\vec{v}_1 = \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix} \text{ and } \vec{v}_3 = \begin{bmatrix} 5 \\ 0 \\ 5 \end{bmatrix}.$$

Here $S = \{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ are linearly independent, and $\vec{v}_1 \cdot \vec{v}_2 = 0$, $\vec{v}_2 \cdot \vec{v}_3 = 0$

Now we need to find an orthonormal set based on Q.

$$\vec{q}_1 = \vec{v}_1 = \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix}, \quad \vec{q}_3 = \vec{v}_3 - \frac{\vec{v}_3 \cdot \vec{q}_1}{\vec{q}_1 \cdot \vec{q}_1} \vec{q}_1$$

$$= \vec{v}_3 - \frac{-5}{5} \vec{q}_1 = \begin{bmatrix} 4 \\ 2 \\ 5 \end{bmatrix}$$

$$\therefore S' = \left\{ \frac{\vec{q}_1}{\|q_1\|}, \frac{\vec{q}_2}{\|q_2\|}, \frac{\vec{q}_3}{\|q_3\|} \right\} \text{ is an orthonormal set.}$$

Thus, $Q = \begin{bmatrix} \frac{\vec{q}_2}{\|q_2\|} & \frac{\vec{q}_1}{\|q_1\|} & \frac{\vec{q}_3}{\|q_3\|} \end{bmatrix}$ is an orthogonal matrix satisfying the question.

$$= \begin{bmatrix} 2/3 & -1/\sqrt{5} & 4/3\sqrt{5} \\ 1/3 & 2/\sqrt{5} & 2/3\sqrt{5} \\ -2/3 & 0 & 5/3\sqrt{5} \end{bmatrix}$$