Eigenvalue eigenvector

Instructor: Dr. Avijit Pal

Linear algebra- II (IC152)

Observation :4

- Let $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ be an orthonormal subset of $\mathbb{R}^n$.
- Let $\mathcal{B} = (\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n)$ be the standard ordered basis of $\mathbb{R}^n$. Then there exist real numbers $\alpha_{ij},\ 1 \le i \le k,\ 1 \le j \le n$ such that

$$[\mathbf{v}_i]_{\mathcal{B}} = (\alpha_{1i}, \alpha_{2i}, \dots, \alpha_{ni})^t.$$

- Let

$$A = [\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k].$$

- Then in the ordered basis $\mathcal{B}$, we have

$$A = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1k} \\ \vdots & \vdots & \dots & \vdots \\ \alpha_{n1} & \alpha_{n2} & \cdots & \alpha_{nk} \end{bmatrix}$$

is an $n \times k$ matrix.

- Also, observe that the conditions $\|\mathbf{v}_i\| = 1$ and $\langle \mathbf{v}_i, \mathbf{v}_j \rangle = 0$ for $1 \le i \ne j \le n$, implies that implies that

$$1 = \|\mathbf{v}_i\|^2 = \langle \mathbf{v}_i, \mathbf{v}_i \rangle = \sum_{j=1}^{n} \alpha_{ji}^2 \text{ and } 0 = \langle \mathbf{v}_i, \mathbf{v}_j \rangle = \sum_{s=1}^{n} \alpha_{si}\alpha_{sj}. \tag{1}$$

- Note that,

$$A^tA = \begin{bmatrix} \mathbf{v}_1^t \\ \mathbf{v}_2^t \\ \vdots \\ \mathbf{v}_k^t \end{bmatrix} \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_k \end{bmatrix} = \begin{bmatrix} \|\mathbf{v}_1\|^2 & \langle \mathbf{v}_1, \mathbf{v}_2 \rangle & \cdots & \langle \mathbf{v}_1, \mathbf{v}_k \rangle \\ \vdots & \vdots & \ldots & \vdots \\ \langle \mathbf{v}_k, \mathbf{v}_1 \rangle & \langle \mathbf{v}_k, \mathbf{v}_2 \rangle & \cdots & \|\mathbf{v}_k\|^2 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} = I_k.$$

- Or using (1), in the language of matrices, we get

$$A^tA = \begin{bmatrix} \alpha_{11} & \alpha_{21} & \cdots & \alpha_{n1} \\ \alpha_{12} & \alpha_{22} & \cdots & \alpha_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{1k} & \alpha_{2k} & \cdots & \alpha_{nk} \end{bmatrix} \begin{bmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{n1} & \alpha_{n2} & \cdots & \alpha_{nk} \end{bmatrix} = I_k.$$

- Notice that the inverse of A is its transpose.
- Such matrices are called orthogonal matrices and they have a special role to play.

Definition

A $n \times n$ real matrix A is said to be an orthogonal matrix if

$$A\,A^t = A^t A = I_n.$$

Theorem (QR Decomposition)

Let A be a square matrix of order n. Then there exist matrices Q and R such that Q is orthogonal and R is upper triangular with $A = QR$.
In case, A is non-singular, the diagonal entries of R can be chosen to be positive. Also, in this case, the decomposition is unique.

- We prove the theorem when A is non-singular.
- Let the columns of A be $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$.
- The Gram-Schmidt orthogonalisation process applied to the vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ gives the vectors $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$ satisfying
  1. $\|\mathbf{u}_i\| = 1$ for $1 \leq i \leq n$,
  2. $\langle \mathbf{u}_i, \mathbf{u}_j \rangle = 0$ for $1 \leq i \neq j \leq n$, and
  3. $L(\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n) = L(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)$.
- Now, consider the ordered basis $\mathcal{B} = (\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n)$.
- So, we can find scalars $\alpha_{ji}, 1 \leq j \leq i$ such that

$$\mathbf{x}_i = \alpha_{1i}\mathbf{u}_1 + \alpha_{2i}\mathbf{u}_2 + \cdots + \alpha_{ii}\mathbf{u}_i = \left[(\alpha_{1i}, \ldots, \alpha_{ii}, 0 \ldots, 0)^t\right]_{\mathcal{B}}. \qquad (2)$$

Outline of the proof cont.

- Let $Q = [\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n]$. Then Q is an orthogonal matrix.
- We now define an $n \times n$ upper triangular matrix R by

$$R = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ 0 & \alpha_{22} & \cdots & \alpha_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \alpha_{nn} \end{bmatrix}.$$

- By using (2), we get

$$\begin{aligned} QR &= [\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n] \begin{bmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ 0 & \alpha_{22} & \cdots & \alpha_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \alpha_{nn} \end{bmatrix} \\ &= \left[\alpha_{11}\mathbf{u}_1,\ \alpha_{12}\mathbf{u}_1 + \alpha_{22}\mathbf{u}_2, \dots, \sum_{i=1}^{n} \alpha_{in}\mathbf{u}_i\right] \\ &= [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n] = A. \end{aligned}$$

Outline of the proof cont.

- Thus, we see that $A = QR$, where Q is an orthogonal matrix and R is an upper triangular matrix.
- The proof doesn't guarantee that for $1 \leq i \leq n$, α_{ii} is positive.
- But this can be achieved by replacing the vector $\mathbf{u}_i$ by $-\mathbf{u}_i$ whenever α_{ii} is negative.
- **Uniqueness** :
  1. Suppose $Q_1R_1 = Q_2R_2$ then $Q_2^{-1}Q_1 = R_2R_1^{-1}$. Observe the following properties of upper triangular matrices.
  2. The inverse of an upper triangular matrix is also an upper triangular matrix, and product of upper triangular matrices is also upper triangular.
  3. Thus the matrix $R_2R_1^{-1}$ is an upper triangular matrix. Also, the matrix $Q_2^{-1}Q_1$ is an orthogonal matrix.
  4. Hence, $R_2R_1^{-1} = I_n$. So, $R_2 = R_1$ and therefore $Q_2 = Q_1$.

Generalised QR Decomposition

- Suppose we have matrix $A = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k]$ of dimension $n \times k$ with rank $(A) = r$.
- Then by the application of the Gram-Schmidt orthogonalisation process yields a set $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_r\}$ of orthonormal vectors of $\mathbb{R}^n$.
- In this case, for each $i,\ 1 \leq i \leq r$, we have

$$L(\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_i) = L(\mathbf{x}_1, \ldots, \mathbf{x}_j), \quad \text{for some } j, \quad i \leq j \leq k.$$

- Hence, proceeding on the lines of the above theorem, we have the following result.

Theorem (Generalised QR Decomposition)

Let A be an $n \times k$ matrix of rank r. Then $A = QR$, where

1. *Q is an $n \times r$ matrix with $Q^t Q = I_r$. That is, the columns of Q form an orthonormal set,*
2. *If $Q = [\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_r]$, then $L(\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_r) = L(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k)$,*
3. *R is an $r \times k$ matrix with rank $(R) = r$.*

Example 1

- Let $A = \begin{bmatrix} 1 & 0 & 1 & 2 \\ 0 & 1 & -1 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \end{bmatrix}$.
- We want to find an orthogonal matrix Q and an upper triangular matrix R such that $A = QR$.
- We know that

$$\mathbf{v}_1 = \frac{1}{\sqrt{2}}(1,0,1,0),\ \mathbf{v}_2 = \frac{1}{\sqrt{2}}(0,1,0,1),\ \mathbf{v}_3 = \frac{1}{\sqrt{2}}(0,-1,0,1). \quad (3)$$

- We now compute $\mathbf{w}_4$. If we denote $\mathbf{u}_4 = (2,1,1,1)^t$ then by the Gram-Schmidt process,

$$\begin{aligned} \mathbf{w}_4 &= \mathbf{u}_4 - \langle \mathbf{u}_4, \mathbf{v}_1 \rangle \mathbf{v}_1 - \langle \mathbf{u}_4, \mathbf{v}_2 \rangle \mathbf{v}_2 - \langle \mathbf{u}_4, \mathbf{v}_3 \rangle \mathbf{v}_3 \\ &= \frac{1}{2}(1,0,-1,0)^t. \end{aligned} \quad (4)$$

Example 1 cont.

- Thus, using (3) and (4), we get

$$Q = \left[\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\right] = \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & 0 & \frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & 0 & 0 & -\frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \end{bmatrix}$$

and

$$R = \begin{bmatrix} \sqrt{2} & 0 & \sqrt{2} & \frac{3}{\sqrt{2}} \\ 0 & \sqrt{2} & 0 & \sqrt{2} \\ 0 & 0 & \sqrt{2} & 0 \\ 0 & 0 & 0 & -\frac{1}{\sqrt{2}} \end{bmatrix}$$

and

- Check that $A = QR$.

Example 2

- Let $A = \begin{bmatrix} 1 & 1 & 1 & 0 \\ -1 & 0 & -2 & 1 \\ 1 & 1 & 1 & 0 \\ 1 & 0 & 2 & 1 \end{bmatrix}$.
- We wish to find a 4×3 matrix Q satisfying $Q^t Q = I_3$ and an upper triangular matrix R such that $A = QR$.
- Let us apply the Gram Schmidt orthogonalisation to the columns of A .
- Or equivalently to the rows of A^t . So, we need to apply the process to the subset $\{(1,-1,1,1),(1,0,1,0),(1,-2,1,2),(0,1,0,1)\}$ of $\mathbb{R}^4$.
- Let $\mathbf{u}_1 = (1,-1,1,1)$. Define $\mathbf{v}_1 = \dfrac{\mathbf{u}_1}{2}$. Let $\mathbf{u}_2 = (1,0,1,0)$.
- Then $\mathbf{w}_2 = (1,0,1,0) - \langle \mathbf{u}_2, \mathbf{v}_1 \rangle \mathbf{v}_1 = (1,0,1,0) - \mathbf{v}_1 = \dfrac{1}{2}(1,1,1,-1)$.

- Hence, $\mathbf{v}_2 = \dfrac{(1,1,1,-1)}{2}$. Let $\mathbf{u}_3 = (1,-2,1,2)$.
- Then

$$\mathbf{w}_3 = \mathbf{u}_3 - \langle \mathbf{u}_3, \mathbf{v}_1 \rangle \mathbf{v}_1 - \langle \mathbf{u}_3, \mathbf{v}_2 \rangle \mathbf{v}_2 = \mathbf{u}_3 - 3\mathbf{v}_1 + \mathbf{v}_2 = \mathbf{0}.$$

- So, we again take $\mathbf{u}_3 = (0,1,0,1)$. Then

$$\mathbf{w}_3 = \mathbf{u}_3 - \langle \mathbf{u}_3, \mathbf{v}_1 \rangle \mathbf{v}_1 - \langle \mathbf{u}_3, \mathbf{v}_2 \rangle \mathbf{v}_2 = \mathbf{u}_3 - 0\mathbf{v}_1 - 0\mathbf{v}_2 = \mathbf{u}_3.$$

So, $\mathbf{v}_3 = \dfrac{(0,1,0,1)}{\sqrt{2}}$.

- Hence,

$$Q = [\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3] = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ -\frac{1}{2} & \frac{1}{2} & \frac{1}{\sqrt{2}} \\ \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

and

$$R = \begin{bmatrix} 2 & 1 & 3 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & \sqrt{2} \end{bmatrix}$$

- Check the following:
  1. rank $(A) = 3$,
  2. $A = QR$ with $Q^t Q = I_3$,
  3. R a 3×4 upper triangular matrix with rank $(R) = 3$.

- Recall that given a k -dimensional vector subspace of a vector space V of dimension n, one can always find an $(n-k)$ -dimensional vector subspace W_0 of V satisfying

$$W + W_0 = V \quad \text{and} \quad W \cap W_0 = \{\mathbf{0}\}.$$

- The subspace W_0 is called the complementary subspace of W in V.
- We now define an important class of linear transformations on an inner product space, called orthogonal projections.

Definition (Projection Operator)

Let V be an n -dimensional vector space and let W be a k -dimensional subspace of V. Let W_0 be a complement of W in V. Then we define a map $P_W : V \longrightarrow V$ by

$$P_W(\mathbf{v}) = \mathbf{w}, \quad \text{whenever} \quad \mathbf{v} = \mathbf{w} + \mathbf{w}_0, \ \mathbf{w} \in W, \ \mathbf{w}_0 \in W_0.$$

The map P_W is called the projection of V onto W along W_0.

Remark

- *The map P is well defined due to the following reasons:*
  1. $W + W_0 = V$ *implies that for every* $\mathbf{v} \in V$, *we can find* $w \in W$ *and* $w_0 \in W_0$ *such that* $\mathbf{v} = \mathbf{w} + \mathbf{w}_0$.
  2. $W \cap W_0 = \{\mathbf{0}\}$ *implies that the expression* $\mathbf{v} = \mathbf{w} + \mathbf{w}_0$ *is unique for every* $\mathbf{v} \in V$.

- The next proposition states that the map defined above is a linear transformation from V to V.
- We omit the proof, as it follows directly from the above remarks.

Proposition

The map $P_W : V \longrightarrow V$ defined above is a linear transformation.

Example 1

- Let $V = \mathbb{R}^3$ and $W = \{(x, y, z) \in \mathbb{R}^3 : x + y - z = 0\}$.
- Let $W_0 = L(\ (1, 2, 2)\ )$. Then $W \cap W_0 = \{\mathbf{0}\}$ and $W + W_0 = \mathbb{R}^3$.
- Also, for any vector $(x, y, z) \in \mathbb{R}^3$, note that $(x, y, z) = \mathbf{w} + \mathbf{w}_0$, where

  $$\mathbf{w} = (z - y, 2z - 2x - y, 3z - 2x - 2y), \quad \text{and} \quad \mathbf{w}_0 = (x + y - z)(1, 2, 2).$$

- So, by definition,

$$P_W((x, y, z)) = (z - y, 2z - 2x - y, 3z - 2x - 2y) = \begin{bmatrix} 0 & -1 & 1 \\ 2 & -1 & 2 \\ -2 & -2 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}.$$

Example 1, cont.

- Let $W_0 = L(\ (1,1,1)\ )$. Then $W \cap W_0 = \{\mathbf{0}\}$ and $W + W_0 = \mathbb{R}^3$.
- Also, for any vector $(x,y,z) \in \mathbb{R}^3$, note that $(x,y,z) = \mathbf{w} + \mathbf{w}_0$, where

$$\mathbf{w} = (z-y, z-x, 2z-x-y), \quad \text{and} \quad \mathbf{w}_0 = (x+y-z)(1,1,1).$$

- So, by definition,

$$P_W(\ (x,y,z)\ ) = (z-y, z-x, 2z-x-y) = \begin{bmatrix} 0 & -1 & 1 \\ -1 & 0 & 1 \\ -1 & -1 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}.$$

Thank You