Linear Algebra and Applications
12 November 2014

Lect. Erke Arıbaş

References:
-***Elementary Linear Algebra-Applications Version"***,
Howard Anton and Chris Rorres, 9^{th} Edition, Wiley, 2010.

-Inner Product Spaces

-Orthonormal Bases, Gram-Schmidt Process

-QR Decomposition

-Eigenvalue and Eigenvectors

-Diagonalization

Chapters 6 and 7 in the course's textbook.

DEFINITION 1 An *inner product* on a real vector space V is a function that associates a real number $\langle \mathbf{u}, \mathbf{v} \rangle$ with each pair of vectors in V in such a way that the following axioms are satisfied for all vectors $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ in V and all scalars k.

1. $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$ **[Symmetry axiom]**
2. $\langle \mathbf{u} + \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{w} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle$ **[Additivity axiom]**
3. $\langle k\mathbf{u}, \mathbf{v} \rangle = k\langle \mathbf{u}, \mathbf{v} \rangle$ **[Homogeneity axiom]**
4. $\langle \mathbf{v}, \mathbf{v} \rangle \geq 0$ and $\langle \mathbf{v}, \mathbf{v} \rangle = 0$ if and only if $\mathbf{v} = \mathbf{0}$ **[Positivity axiom]**

A real vector space with an inner product is called a *real inner product space*.

$$\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u} \cdot \mathbf{v} = u_1v_1 + u_2v_2 + \cdots + u_nv_n$$

Definition

- If V is an inner product space, then the norm (or length) of a vector $\mathbf{u}$ in V is denoted by $\|\mathbf{u}\|$ and is defined by

$$\|\mathbf{u}\| = \langle \mathbf{u}, \mathbf{u} \rangle^{1/2}$$

- The distance between two points (vectors) $\mathbf{u}$ and $\mathbf{v}$ is denoted by $d(\mathbf{u},\mathbf{v})$ and is defined by

$$d(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\|$$

Norm and Distance in R^n

- If $\mathbf{u} = (u_1, u_2, \ldots, u_n)$ and $\mathbf{v} = (v_1, v_2, \ldots, v_n)$ are vectors in R^n with the Euclidean inner product, then

$$\|\mathbf{u}\| = <\mathbf{u},\ \mathbf{u}>^{1/2} = (\mathbf{u},\ \mathbf{u})^{1/2} = \sqrt{u_1^2 + u_2^2 + \ldots + u_n^2}$$

$$d(\mathbf{u},\ \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\| = <\mathbf{u} - \mathbf{v},\ \mathbf{u} - \mathbf{v}>^{1/2} = \left[(\mathbf{u} - \mathbf{v}) \cdot (\mathbf{u} - \mathbf{v})\right]^{1/2}$$
$$= \sqrt{(u_1 - v_1)^2 + (u_2 - v_2)^2 + \ldots + (u_n - v_n)^2}$$

Algebraic Properties of Inner Products

THEOREM 6.1.2 *If* $\mathbf{u}$, $\mathbf{v}$, *and* $\mathbf{w}$ *are vectors in a real inner product space* V, *and if* k *is a scalar, then:*

(*a*) $\langle \mathbf{0}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{0} \rangle = 0$

(*b*) $\langle \mathbf{u}, \mathbf{v} + \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{u}, \mathbf{w} \rangle$

(*c*) $\langle \mathbf{u}, \mathbf{v} - \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle - \langle \mathbf{u}, \mathbf{w} \rangle$

(*d*) $\langle \mathbf{u} - \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{w} \rangle - \langle \mathbf{v}, \mathbf{w} \rangle$

(*e*) $k\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{u}, k\mathbf{v} \rangle$

Inner Products Generated by Matrices

- Let $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$ be vectors in R^n (expressed as $n \times 1$ matrices), and let A be an invertible $n \times n$ matrix.
- If $\mathbf{u} \cdot \mathbf{v}$ is the Euclidean inner product on R^n, then the formula

$$\langle \mathbf{u}, \mathbf{v} \rangle = A\mathbf{u} \cdot A\mathbf{v}$$

  defines an inner product; it is called the inner product on R^n generated by A.
- Recalling that the Euclidean inner product $\mathbf{u} \cdot \mathbf{v}$ can be written as the matrix product $\mathbf{v}^T\mathbf{u}$, the above formula can be written in the alternative form $\langle \mathbf{u}, \mathbf{v} \rangle = (A\mathbf{v})^T A\mathbf{u}$, or equivalently,

$$\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{v}^T A^T A \mathbf{u}$$

Theorems

- Theorem 6.2.1 (Cauchy-Schwarz Inequality)
  - If $\mathbf{u}$ and $\mathbf{v}$ are vectors in a real inner product space, then

$$|\langle \mathbf{u}, \mathbf{v} \rangle| \leq \|\mathbf{u}\| \|\mathbf{v}\|$$

- Theorem 6.2.2 (Properties of Length)
  - If $\mathbf{u}$ and $\mathbf{v}$ are vectors in an inner product space V, and if k is any scalar, then :
    - $\| \mathbf{u} \| \geq 0$
    - $\| \mathbf{u} \| = 0$ if and only if $\mathbf{u} = \mathbf{0}$
    - $\| k\mathbf{u} \| = | k | \| \mathbf{u} \|$
    - $\| \mathbf{u} + \mathbf{v} \| \leq \| \mathbf{u} \| + \| \mathbf{v} \|$ (Triangle inequality)
- Theorem 6.2.3 (Properties of Distance)
  - If $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ are vectors in an inner product space V, and if k is any scalar, then:
    - $d(\mathbf{u}, \mathbf{v}) \geq 0$
    - $d(\mathbf{u}, \mathbf{v}) = 0$ if and only if $\mathbf{u} = \mathbf{v}$
    - $d(\mathbf{u}, \mathbf{v}) = d(\mathbf{v}, \mathbf{u})$
    - $d(\mathbf{u}, \mathbf{v}) \leq d(\mathbf{u}, \mathbf{w}) + d(\mathbf{w}, \mathbf{v})$ (Triangle inequality)

Remarks

- The Cauchy-Schwarz inequality for R^n (Theorem 4.1.3) follows as a special case of Theorem 6.2.1 by taking $\langle \mathbf{u}, \mathbf{v} \rangle$ to be the Euclidean inner product $\mathbf{u} \cdot \mathbf{v}$.

- The angle between vectors in general inner product spaces can be defined as

$$\cos\theta = \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\|\mathbf{u}\|\|\mathbf{v}\|} \quad \text{and} \quad 0 \le \theta \le \pi$$

- Example
  - Let R^4 have the Euclidean inner product. Find the cosine of the angle θ between the vectors $\mathbf{u} = (4, 3, 1, -2)$ and $\mathbf{v} = (-2, 1, 2, 3)$.

θ: the angle between u and v

$$\theta = \cos^{-1}\left(\frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\|\mathbf{u}\|\,\|\mathbf{v}\|}\right)$$

DEFINITION 1 Two vectors $\mathbf{u}$ and $\mathbf{v}$ in an inner product space are called ***orthogonal*** if $\langle \mathbf{u}, \mathbf{v} \rangle = 0$.

Orthogonality

- Definition
  - Two vectors **u** and **v** in an inner product space are called orthogonal if $\langle \mathbf{u}, \mathbf{v} \rangle = 0$.

- Example
  - If M_{22} has the inner project defined previously, then the matrices

$$U = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \text{ and } V = \begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix}$$

    are orthogonal, since $\langle U, V \rangle = 1(0) + 0(2) + 1(0) + 1(0) = 0$.

Orthonormal Basis

- Definition
  - A set of vectors in an inner product space is called an orthogonal set if all pairs of distinct vectors in the set are orthogonal.
  - An orthogonal set in which each vector has norm 1 is called orthonormal.

- Example
  - Let $\mathbf{u}_1 = (0, 1, 0)$, $\mathbf{u}_2 = (1, 0, 1)$, $\mathbf{u}_3 = (1, 0, -1)$ and assume that R^3 has the Euclidean inner product.
  - It follows that the set of vectors $S = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ is <u>orthogonal</u> since
  $$\langle \mathbf{u}_1, \mathbf{u}_2 \rangle = \langle \mathbf{u}_1, \mathbf{u}_3 \rangle = \langle \mathbf{u}_2, \mathbf{u}_3 \rangle = 0.$$
  - The Euclidean norms of the vectors are $\|\mathbf{u}_1\| = 1,\ \|\mathbf{u}_2\| = \sqrt{2},\ \|\mathbf{u}_3\| = \sqrt{2}$
  - Normalizing $\mathbf{u}_1$, $\mathbf{u}_2$, and $\mathbf{u}_3$ yields
  $$\mathbf{v}_1 = \frac{\mathbf{u}_1}{\|\mathbf{u}_1\|} = (0,1,0), \quad \mathbf{v}_2 = \frac{\mathbf{u}_2}{\|\mathbf{u}_2\|} = (\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}), \quad \mathbf{v}_3 = \frac{\mathbf{u}_3}{\|\mathbf{u}_3\|} = (\frac{1}{\sqrt{2}}, 0, -\frac{1}{\sqrt{2}})$$
  - The set $S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is <u>orthonormal</u> since
  $$\langle \mathbf{v}_1, \mathbf{v}_2 \rangle = \langle \mathbf{v}_1, \mathbf{v}_3 \rangle = \langle \mathbf{v}_2, \mathbf{v}_3 \rangle = 0 \quad and \quad \|\mathbf{v}_1\| = \|\mathbf{v}_2\| = \|\mathbf{v}_3\| = 1$$

Orthonormal Basis

- Theorem 6.3.1*
  - If $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is an orthonormal basis for an inner product space V, and $\mathbf{u}$ is any vector in V, then

$$\mathbf{u} = \langle \mathbf{u}, \mathbf{v}_1 \rangle \mathbf{v}_1 + \langle \mathbf{u}, \mathbf{v}_2 \rangle \mathbf{v}_2 + \cdots + \langle \mathbf{u}, \mathbf{v}_n \rangle \mathbf{v}_n$$

- Remark
  - The scalars $\langle \mathbf{u}, \mathbf{v}_1 \rangle, \langle \mathbf{u}, \mathbf{v}_2 \rangle, \ldots, \langle \mathbf{u}, \mathbf{v}_n \rangle$ are the coordinates of the vector $\mathbf{u}$ relative to the orthonormal basis $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ and

$$(\mathbf{u})_S = (\langle \mathbf{u}, \mathbf{v}_1 \rangle, \langle \mathbf{u}, \mathbf{v}_2 \rangle, \ldots, \langle \mathbf{u}, \mathbf{v}_n \rangle)$$

    is the coordinate vector of $\mathbf{u}$ relative to this basis

Example

- Let $\mathbf{v}_1 = (0, 1, 0)$, $\mathbf{v}_2 = (-4/5, 0, 3/5)$, $\mathbf{v}_3 = (3/5, 0, 4/5)$.
  It is easy to check that $S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is an orthonormal basis for R^3 with the Euclidean inner product.
  Express the vector $\mathbf{u} = (1, 1, 1)$ as a linear combination of the vectors in S, and find the coordinate vector $(\mathbf{u})_S$.

- Solution:
  - $\langle \mathbf{u}, \mathbf{v}_1 \rangle = 1$, $\langle \mathbf{u}, \mathbf{v}_2 \rangle = -1/5$, $\langle \mathbf{u}, \mathbf{v}_3 \rangle = 7/5$
  - Therefore, by Theorem 6.3.1 we have $\mathbf{u} = \mathbf{v}_1 - 1/5\ \mathbf{v}_2 + 7/5\ \mathbf{v}_3$
  - That is, $(1, 1, 1) = (0, 1, 0) - 1/5\ (-4/5, 0, 3/5) + 7/5\ (3/5, 0, 4/5)$
  - The coordinate vector of $\mathbf{u}$ relative to S is

$$(\mathbf{u})_S = (\langle \mathbf{u}, \mathbf{v}_1 \rangle, \langle \mathbf{u}, \mathbf{v}_2 \rangle, \langle \mathbf{u}, \mathbf{v}_3 \rangle) = (1, -1/5, 7/5)$$

Converting an arbitrary basis into orthogonal basis: Gram-Schmidt Process

The Gram–Schmidt Process

To convert a basis $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_r\}$ into an orthogonal basis $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$, perform the following computations:

Step 1. $\mathbf{v}_1 = \mathbf{u}_1$

Step 2. $\mathbf{v}_2 = \mathbf{u}_2 - \dfrac{\langle \mathbf{u}_2, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|^2}\mathbf{v}_1$

Step 3. $\mathbf{v}_3 = \mathbf{u}_3 - \dfrac{\langle \mathbf{u}_3, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 - \dfrac{\langle \mathbf{u}_3, \mathbf{v}_2 \rangle}{\|\mathbf{v}_2\|^2}\mathbf{v}_2$

Step 4. $\mathbf{v}_4 = \mathbf{u}_4 - \dfrac{\langle \mathbf{u}_4, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 - \dfrac{\langle \mathbf{u}_4, \mathbf{v}_2 \rangle}{\|\mathbf{v}_2\|^2}\mathbf{v}_2 - \dfrac{\langle \mathbf{u}_4, \mathbf{v}_3 \rangle}{\|\mathbf{v}_3\|^2}\mathbf{v}_3$

$\vdots$

(continue for r steps)

Optional Step. To convert the orthogonal basis into an orthonormal basis $\{\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_r\}$, normalize the orthogonal basis vectors.

Example (Gram-Schmidt Process)

- Consider the vector space R^3 with the Euclidean inner product. Apply the Gram-Schmidt process to transform the basis vectors

$$\mathbf{u}_1 = (1, 1, 1),\ \mathbf{u}_2 = (0, 1, 1),\ \mathbf{u}_3 = (0, 0, 1)$$

into an orthogonal basis $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$; then normalize the orthogonal basis vectors to obtain an orthonormal basis $\{\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3\}$.

- Solution:
  - Step 1: Let $\mathbf{v}_1 = \mathbf{u}_1$. That is, $\mathbf{v}_1 = \mathbf{u}_1 = (1, 1, 1)$
  - Step 2: Let $\mathbf{v}_2 = \mathbf{u}_2 - \text{proj}_{W_1}\mathbf{u}_2$. That is,

$$\mathbf{v}_2 = \mathbf{u}_2 - \text{proj}_{w_1}\mathbf{u}_2 = \mathbf{u}_2 - \frac{<\mathbf{u}_2, \mathbf{v}_1>}{\|\mathbf{v}_1\|^2}\mathbf{v}_1$$

$$= (0, 1, 1) - \frac{2}{3}(1, 1, 1) = (-\frac{2}{3}, \frac{1}{3}, \frac{1}{3})$$

Example (Gram-Schmidt Process)

We have two vectors in W_2 now!

- Step 3: Let $\mathbf{v}_3 = \mathbf{u}_3 - \text{proj}_{W_2}\mathbf{u}_3$. That is,

$$\mathbf{v}_3 = \mathbf{u}_3 - \text{proj}_{W_2}\mathbf{u}_3 = \mathbf{u}_3 - \frac{<\mathbf{u}_3, \mathbf{v}_1>}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 - \frac{<\mathbf{u}_3, \mathbf{v}_2>}{\|\mathbf{v}_2\|^2}\mathbf{v}_2$$

$$= (0, 1, 1) - \frac{1}{3}(1, 1, 1) - \frac{1/3}{2/3}(-\frac{2}{3}, \frac{1}{3}, \frac{1}{3}) = (0, -\frac{1}{2}, \frac{1}{2})$$

- Thus, $\mathbf{v}_1 = (1, 1, 1)$, $\mathbf{v}_2 = (-2/3, 1/3, 1/3)$, $\mathbf{v}_3 = (0, -1/2, 1/2)$ form an orthogonal basis for R^3. The norms of these vectors are

$$\|\mathbf{v}_1\| = \sqrt{3}, \ \|\mathbf{v}_2\| = \frac{\sqrt{6}}{3}, \ \|\mathbf{v}_3\| = \frac{1}{\sqrt{2}}$$

so an orthonormal basis for R^3 is

$$\mathbf{q}_1 = \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|} = (\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}), \ \mathbf{q}_2 = \frac{\mathbf{v}_2}{\|\mathbf{v}_2\|} = (-\frac{2}{\sqrt{6}}, \frac{1}{6}, \frac{1}{\sqrt{6}}),$$

$$\mathbf{q}_3 = \frac{\mathbf{v}_3}{\|\mathbf{v}_3\|} = (0, -\frac{1}{\sqrt{2}}, \frac{1}{2})$$

Theorems

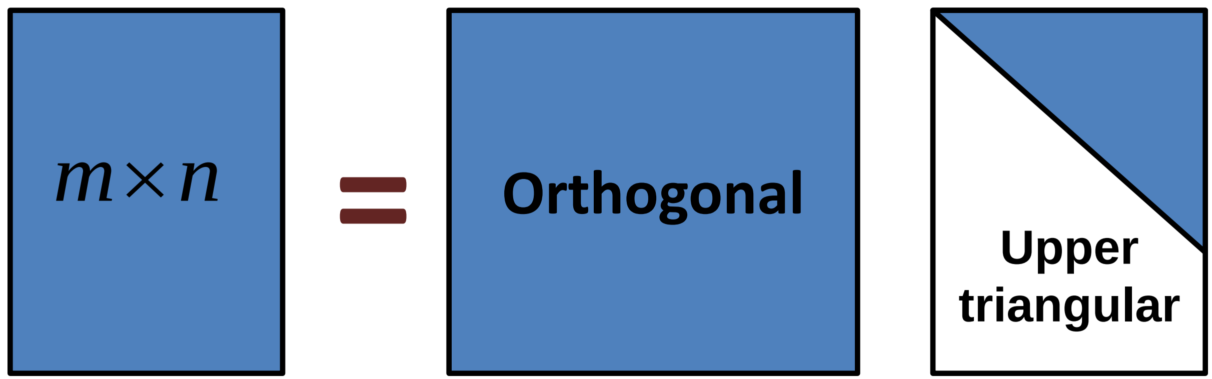
- Theorem 6.3.7 (QR-Decomposition)
  - If A is an $m \times n$ matrix with linearly independent column vectors, then A can be factored as

$$A = QR$$

    where Q is an $m \times n$ matrix with orthonormal column vectors, and R is an $n \times n$ invertible upper triangular matrix.

- Remark
  - In recent years the QR-decomposition has assumed growing importance as the mathematical foundation for a wide variety of practical algorithms, including a widely used algorithm for computing eigenvalues of large matrices.

$$A = QR$$



QR-Decomposition of a 3×3 Matrix

- Find the QR-decomposition of $A = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}$
- Solution:
  - The column vectors A are

$$\mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{u}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{u}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

  - Applying the Gram-Schmidt process with subsequent normalization to these column vectors yields the orthonormal vectors

$$\mathbf{q}_1 = \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}, \mathbf{q}_2 = \begin{bmatrix} -2/\sqrt{6} \\ 1/\sqrt{6} \\ 1/\sqrt{6} \end{bmatrix}, \mathbf{q}_3 = \begin{bmatrix} 0 \\ -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} \quad \Rightarrow Q$$

QR-Decomposition of a 3×3 Matrix

- The matrix R is

$$R=\begin{bmatrix}\langle \mathbf{u}_1,\mathbf{q}_1\rangle & \langle \mathbf{u}_2,\mathbf{q}_1\rangle & \langle \mathbf{u}_3,\mathbf{q}_1\rangle \\ 0 & \langle \mathbf{u}_2,\mathbf{q}_2\rangle & \langle \mathbf{u}_3,\mathbf{q}_2\rangle \\ 0 & 0 & \langle \mathbf{u}_3,\mathbf{q}_3\rangle\end{bmatrix}=\begin{bmatrix}3/\sqrt{3} & 2/\sqrt{3} & 1/\sqrt{3} \\ 0 & 2/\sqrt{6} & 1/\sqrt{6} \\ 0 & 0 & 1/\sqrt{2}\end{bmatrix}$$

- Thus, the QR-decomposition of A is

$$\underbrace{\begin{bmatrix}1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1\end{bmatrix}}_{A}=\underbrace{\begin{bmatrix}1/\sqrt{3} & -2/\sqrt{6} & 0 \\ 1/\sqrt{3} & 1/\sqrt{6} & -1/\sqrt{2} \\ 1/\sqrt{3} & 1/\sqrt{6} & 1/\sqrt{3}\end{bmatrix}}_{Q}\underbrace{\begin{bmatrix}3/\sqrt{3} & 2/\sqrt{3} & 1/\sqrt{3} \\ 0 & 2/\sqrt{6} & 1/\sqrt{6} \\ 0 & 0 & 1/\sqrt{2}\end{bmatrix}}_{R}$$

QR Decomposition

Example: Find the QR decomposition of

$$A = \begin{bmatrix} 1 & -1 & -1 \\ 1 & 0 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

Calculation of QR Decomposition

Applying Gram-Schmidt process of computing QR decomposition

1st Step:

$$r_{11} = \|a_1\| = \sqrt{3}$$

$$q_1 = \frac{1}{\|a_1\|} a_1 = \begin{pmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \\ 0 \end{pmatrix}$$

2nd Step:

$$r_{12} = q_1^T a_2 = -2/\sqrt{3}$$

3rd Step:

$$\hat{q}_2 = a_2 - q_1 q_1^T a_2 = a_2 - q_1 r_{12} = \begin{pmatrix} -1 \\ 0 \\ -1 \\ 0 \end{pmatrix} - (-2/\sqrt{3}) \begin{pmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \\ 0 \end{pmatrix} = \begin{pmatrix} -1/3 \\ 2/3 \\ -1/3 \\ 0 \end{pmatrix}$$

$$r_{22} = \|\hat{q}_2\| = \sqrt{2/3}$$

$$q_2 = \frac{1}{\|\hat{q}_2\|} \hat{q}_2 = \begin{pmatrix} -1/\sqrt{6} \\ \sqrt{2/3} \\ -1/\sqrt{6} \\ 0 \end{pmatrix}$$

Calculation of QR Decomposition

4th Step:

$$r_{13} = q_1^T a_3 = -1\big/\sqrt{3}$$

5th Step:

$$r_{23} = q_2^T a_3 = 1\big/\sqrt{6}$$

6th Step:

$$\hat{q}_3 = a_3 - q_1 q_1^T a_3 - q_2 q_2^T a_3 = a_3 - r_{13} q_1 - r_{23} q_2 = \begin{pmatrix} -1/2 \\ 0 \\ 1/2 \\ -1 \end{pmatrix}$$

$$r_{33} = \left\| \hat{q}_3 \right\| = \sqrt{6}/2$$

$$q_3 = \frac{1}{\left\| \hat{q}_3 \right\|} \hat{q}_3 = \begin{pmatrix} -1/\sqrt{6} \\ 0 \\ 1/\sqrt{6} \\ -2/\sqrt{6} \end{pmatrix}$$

Calculation of QR Decomposition

Therefore, $A=QR$

$$\begin{bmatrix} 1 & -1 & -1 \\ 1 & 0 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{3} & -1/\sqrt{6} & -1/\sqrt{6} \\ 1/\sqrt{3} & 2/\sqrt{6} & 0 \\ 1/\sqrt{3} & -1/\sqrt{6} & 1/\sqrt{6} \\ 0 & 0 & -2/\sqrt{6} \end{bmatrix} \begin{bmatrix} \sqrt{3} & -2/\sqrt{3} & -1/\sqrt{3} \\ 0 & 2/\sqrt{6} & 1/\sqrt{6} \\ 0 & 0 & \sqrt{6}/2 \end{bmatrix}$$

Uses: QR decomposition is widely used in computer codes to find the eigenvalues of a matrix, to solve linear systems, and to find least squares approximations.

TACOMA Bridge.. Collapsed in 1940.. Eigenvalues???

Eigenvalue and Eigenvector

- Definition
  - If A is an $n \times n$ matrix, then a nonzero vector $\mathbf{x}$ in R^n is called an eigenvector of A if $A\mathbf{x}$ is a scalar multiple of $\mathbf{x}$; that is, $A\mathbf{x} = \lambda\mathbf{x}$ for some scalar λ.
  - The scalar λ is called an eigenvalue of A, and $\mathbf{x}$ is said to be an eigenvector of A corresponding to λ.
- Remark
  - To find the eigenvalues of an $n \times n$ matrix A we rewrite $A\mathbf{x} = \lambda\mathbf{x}$ as $A\mathbf{x} = \lambda I\mathbf{x}$ or equivalently, $(\lambda I - A)\mathbf{x} = \mathbf{0}$.
  - For λ to be an eigenvalue, there must be a nonzero solution of this equation. However, by Theorem 6.4.5, the above equation has a nonzero solution if and only if $\det(\lambda I - A) = 0$.
  - This is called the characteristic equation of A; the scalar satisfying this equation are the eigenvalues of A. When expanded, the determinant $\det(\lambda I - A)$ is a polynomial p in λ called the characteristic polynomial of A.

Example

- Find the eigenvalues of

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 4 & -17 & 8 \end{bmatrix}$$

- Solution:
  - The characteristic polynomial of A is

$$\det(\lambda I - A) = \det \begin{bmatrix} \lambda & -1 & 0 \\ 0 & \lambda & -1 \\ -4 & 17 & \lambda - 8 \end{bmatrix} = \lambda^3 - 8\lambda^2 + 17\lambda - 4$$

  - The eigenvalues of A must therefore satisfy the cubic equation

$$\lambda^3 - 8\lambda^2 + 17\lambda - 4 = 0$$

Finding Bases for Eigenspaces

- The eigenvectors of A corresponding to an eigenvalue λ are the nonzero $\mathbf{x}$ that satisfy $A\mathbf{x} = \lambda\mathbf{x}$.

- Equivalently, the eigenvectors corresponding to λ are the nonzero vectors in the solution space of $(\lambda I - A)\mathbf{x} = \mathbf{0}$.

- We call this solution space the eigenspace of A corresponding to λ.

Example

- Find bases for the eigenspaces of $A = \begin{bmatrix} 0 & 0 & -2 \\ 1 & 2 & 1 \\ 1 & 0 & 3 \end{bmatrix}$

- Solution:
  - The characteristic equation of matrix A is $\lambda^3 - 5\lambda^2 + 8\lambda - 4 = 0$, or in factored form, $(\lambda - 1)(\lambda - 2)^2 = 0$; thus, the eigenvalues of A are $\lambda = 1$ and $\lambda = 2$, so there are two eigenspaces of A.
  - $(\lambda I - A)\mathbf{x} = \mathbf{0} \Rightarrow \begin{bmatrix} \lambda & 0 & 2 \\ -1 & \lambda - 2 & -1 \\ -1 & 0 & \lambda - 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$ (3)
  - If $\lambda = 2$, then (3) becomes $\begin{bmatrix} 2 & 0 & 2 \\ -1 & 0 & -1 \\ -1 & 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

Example

$$\begin{bmatrix} 2 & 0 & 2 \\ -1 & 0 & -1 \\ -1 & 0 & -1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

- Solving the system yield

$$x_1 = -s,\ x_2 = t,\ x_3 = s$$

- Thus, the eigenvectors of A corresponding to $\lambda = 2$ are the nonzero vectors of the form

$$\mathbf{x} = \begin{bmatrix} -s \\ t \\ s \end{bmatrix} = \begin{bmatrix} -s \\ 0 \\ s \end{bmatrix} + \begin{bmatrix} 0 \\ t \\ 0 \end{bmatrix} = s\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} + t\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

- The vectors $[-1\ 0\ 1]^T$ and $[0\ 1\ 0]^T$ are linearly independent and form a basis for the eigenspace corresponding to $\lambda = 2$.
- Similarly, the eigenvectors of A corresponding to $\lambda = 1$ are the nonzero vectors of the form $\mathbf{x} = s\ [-2\ 1\ 1]^T$
- Thus, $[-2\ 1\ 1]^T$ is a basis for the eigenspace corresponding to $\lambda = 1$.

Diagonalization

- Definition
  - A square matrix A is called diagonalizable if there is an invertible matrix P such that $P^{-1}AP$ is a diagonal matrix (i.e., $P^{-1}AP = D$); the matrix P is said to diagonalize A.

- Theorem 7.2.1
  - If A is an $n \times n$ matrix, then the following are equivalent.
    - A is diagonalizable.
    - A has n linearly independent eigenvectors.

Procedure for Diagonalizing a Matrix

- The preceding theorem guarantees that an $n \times n$ matrix A with n linearly independent eigenvectors is diagonalizable, and the proof provides the following method for diagonalizing A.
  - **Step 1.** Find n linear independent eigenvectors of A, say, $\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n$.
  - **Step 2.** From the matrix P having $\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n$ as its column vectors.
  - **Step 3.** The matrix $P^{-1}AP$ will then be diagonal with $\lambda_1, \lambda_2, \ldots, \lambda_n$ as its successive diagonal entries, where λ_i is the eigenvalue corresponding to $\mathbf{p}_i$, for $i = 1, 2, \ldots, n$.

Example

- Find a matrix P that diagonalizes $A = \begin{bmatrix} 0 & 0 & -2 \\ 1 & 2 & 1 \\ 1 & 0 & 3 \end{bmatrix}$
- Solution:
  - From the previous example, we have the following bases for the eigenspaces:
    - $\lambda = 2$: $\mathbf{p}_1 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$, $\mathbf{p}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$ $\lambda = 1$: $\mathbf{p}_3 = \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}$
  - Thus,

$$P = \begin{bmatrix} -1 & 0 & -2 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}$$

  - Also,

$$P^{-1}AP = \begin{bmatrix} 1 & 0 & 2 \\ 1 & 1 & 1 \\ -1 & 0 & -1 \end{bmatrix} \begin{bmatrix} 0 & 0 & -2 \\ 1 & 2 & 1 \\ 1 & 0 & 3 \end{bmatrix} \begin{bmatrix} -1 & 0 & -2 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} = D$$

Example (A Non-Diagonalizable Matrix)

- Find a matrix P that diagonalizes $A = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 2 & 0 \\ -3 & 5 & 2 \end{bmatrix}$
- Solution:
  - The characteristic polynomial of A is

$$\det(\lambda I - A) = \begin{vmatrix} \lambda - 1 & 0 & 0 \\ -1 & \lambda - 2 & 0 \\ 3 & -5 & \lambda - 2 \end{vmatrix} = (\lambda - 1)(\lambda - 2)^2$$

  - The bases for the eigenspaces are
    - $\lambda = 1$: $\mathbf{p}_1 = \begin{bmatrix} 1/8 \\ -1/8 \\ 1 \end{bmatrix}$ $\lambda = 2$: $\mathbf{p}_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$
  - Since there are only two basis vectors in total, A is not diagonalizable.

Theorems

- Theorem 7.2.2
  - If $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$, are eigenvectors of A corresponding to distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_k$, then $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ is a linearly independent set.

- Theorem 7.2.3
  - If an $n\times n$ matrix A has n distinct eigenvalues, then A is diagonalizable.

Example

- Since the matrix

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 4 & -17 & 8 \end{bmatrix}$$

  has three distinct eigenvalues, $\lambda = 4,\ \lambda = 2+\sqrt{3},\ \lambda = 2-\sqrt{3}$
- Therefore, A is diagonalizable.
- Further,

$$P^{-1}AP = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 2+\sqrt{3} & 0 \\ 0 & 0 & 2-\sqrt{3} \end{bmatrix}$$

  for some invertible matrix P, and the matrix P can be found using the procedure for diagonalizing a matrix.