Chapter 5 Notes - LA

John Yang

April 5, 2022

Contents

5 Orthogonality

5.1 Orthogonality in $\mathbb{R}^n$

- A set of vectors $\{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k\}$ in $\mathbb{R}^n$ is called an orthogonal set if all pairs of distinct vectors in the set are orthogonal; that is, if

$$\mathbf{v}_i \cdot \mathbf{v}_j = 0 \quad \text{whenever} \quad i \neq j \quad \text{for } i, j = 1, 2, \cdots, k$$

- The standard basis of $\mathbb{R}^n$ is an orthogonal set.

- If $\{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k\}$ is an orthogonal set of nonzero vectors in $\mathbb{R}^n$, then these vectors are linearly independent.

- An orthogonal basis for a subspace W of $\mathbb{R}^n$ is a basis of W that is an orthogonal set.

- Let $\{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k\}$ be an orthogonal basis for a subspace W of $\mathbb{R}^n$ and let $\mathbf{w}$ be any vector in W. Then the unique scalars $c_1, \cdots, c_k$ such that

$$\mathbf{w} = c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k$$

are given by

$$c_i = \frac{\mathbf{w} \cdot \mathbf{v}_i}{\mathbf{v}_i \cdot \mathbf{v}_i} \qquad \text{for } i = 1, \cdots, k$$

- A set of vectors $\mathbb{R}^n$ is an orthonormal set if it is an orthogonal set of unit vectors. An orthonormal basis for a subspace W of $\mathbb{R}^n$ is a basis of W that is an orthonormal set.

- Let $\{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k\}$ be an orthogonal basis for a subspace W of $\mathbb{R}^n$ and let $\mathbf{w}$ be any vector in W. Then

$$\mathbf{w} = (\mathbf{w} \cdot \mathbf{q}_1)\mathbf{q}_1 + (\mathbf{w} \cdot \mathbf{q}_2)\mathbf{q}_2 + \cdots + (\mathbf{w} \cdot \mathbf{q}_k)\mathbf{q}_k$$

and this representation is unique.

- The columns of an $m \times n$ matrix Q form an orthonormal set IFF $Q^TQ = I_n$

- An $n \times n$ matrix Q whose columns form an orthonormal set is called an orthogonal matrix.

- A square matrix Q is orthogonal IFF $Q^{-1} = Q^T$
- Let Q be an $n \times n$ matrix. The following are equivalent:
  - Q is orthogonal.
  - $||Q\mathbf{x}|| = ||\mathbf{x}||$ for every $\mathbf{x}$ in $\mathbb{R}^n$
  - $Q\mathbf{x} \cdot Q\mathbf{y} = \mathbf{x} \cdot \mathbf{y}$ for every $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{R}^n$
- If Q is an orthogonal matrix, then its rows form an orthonormal set.
- Let Q be an orthogonal matrix.
  - Q^{-1} is orthogonal.
  - $\det Q = \pm 1$
  - If λ is an eigenvalue of Q, then $|\lambda| = 1$
  - If Q_1 and Q_2 are orthogonal $n \times n$ matrices, then so is $Q_1 Q_2$.

5.2 Orthogonal Complements and Orthogonal Projections

- Let W be a subspace of $\mathbb{R}^n$. We say that a vector $\mathbf{v}$ in $\mathbb{R}^n$ is orthogonal to W if $\mathbf{v}$ is orthogonal to every vector in W. The set of all vectors that are orthogonal to W is called the orthogonal complement of W, denoted $W^\perp$. That is,
$$W^\perp = \{\mathbf{v} \text{ in } \mathbb{R}^n : \mathbf{v} \cdot \mathbf{w} = 0 \quad \text{for all } \mathbf{w} \text{ in } W\}$$
- Let W be a subspace of $\mathbb{R}^n$.
  - $W^\perp$ is a subspace of $\mathbb{R}^n$
  - $(W^\perp)^\perp = W$
  - $W \cap W^\perp = \{\mathbf{0}\}$
  - If $W = \text{span}(\mathbf{w}_1, \ldots, \mathbf{w}_k)$, then $\mathbf{v}$ is in $W^\perp$ IFF $\mathbf{v} \cdot \mathbf{w}_i = 0$ for all $i = 1, \ldots, k$.
- Let A be an $m \times n$ matrix. Then the orthogonal complement of the row space of A is the null space of A, and the orthogonal complement of the column space of A is the null space of A^T:
$$(\text{row}(A))^\perp = \text{null}(A) \quad \text{and} \quad (\text{col}(A))^\perp = \text{null}(A^T)$$
- Let W be a subspace of $\mathbb{R}^n$ and let $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ be an orthogonal basis for W. For any vector $\mathbf{v}$ in $\mathbb{R}^n$, the orthogonal projection of $\mathbf{v}$ onto W is defined as
$$\text{proj}_w(\mathbf{v}) = \left(\frac{\mathbf{u}_1 \cdot \mathbf{v}}{\mathbf{u}_1 \cdot \mathbf{u}_1}\right)\mathbf{u}_1 + \cdots + \left(\frac{\mathbf{u}_k \cdot \mathbf{v}}{\mathbf{u}_k \cdot \mathbf{u}_k}\right)\mathbf{u}_k$$
  The complement of $\mathbf{v}$ orthogonal to W is the vector
$$\text{perp}_w(\mathbf{v}) = \mathbf{v} - \text{proj}_w(\mathbf{v})$$
- $\text{proj}_w(\mathbf{v}) = \text{proj}_{\mathbf{u}_1}(\mathbf{v}) + \cdots + \text{proj}_{\mathbf{u}_k}(\mathbf{v})$
- The orthogonal decomposition theorem: Let W be a subspace of $\mathbb{R}^n$ and let $\mathbf{v}$ be a vector in $\mathbb{R}^n$. Then there are unique vectors $\mathbf{w}$ in W and $\mathbf{w}^\perp$ in $W^\perp$ such that
$$\mathbf{v} = \mathbf{w} + \mathbf{w}^\perp$$
- If W is a subspace of $\mathbb{R}^n$ then
$$(W^\perp)^\perp = W$$
- If W is a subspace of $\mathbb{R}^n$ then
$$\text{dim}W + \text{dim}W^\perp = n$$
- The Rank Theorem: If A is an $m \times n$ matrix, then
$$\text{rank}(A) + \text{nullity}(A) = n$$

5.3 The Gram-Schmidt Process and the QR Factorization

- The Gram-Schmidt Process: Let $\{\mathbf{x}_1, \ldots, \mathbf{x}_k\}$ be a basis for a subspace W of $\mathbb{R}^n$ and define the following:

$$\mathbf{v}_1 = \mathbf{x}_1; \qquad W_1 = \text{span}(\mathbf{x}_1)$$

$$\mathbf{v}_2 = \mathbf{x}_2 - \left(\frac{\mathbf{v}_1 \cdot \mathbf{x}_2}{\mathbf{v}_1 \cdot \mathbf{v}_1}\right)\mathbf{v}_1, \qquad W_2 = \text{span}(\mathbf{x}_1, \mathbf{x}_2)$$

$$\mathbf{v}_3 = \mathbf{x}_3 - \left(\frac{\mathbf{v}_1 \cdot \mathbf{x}_3}{\mathbf{v}_1 \cdot \mathbf{v}_1}\right)\mathbf{v}_1 - \left(\frac{\mathbf{v}_2 \cdot \mathbf{x}_3}{\mathbf{v}_2 \cdot \mathbf{v}_2}\right)\mathbf{v}_2, \qquad W_3 = \text{span}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$$

$$\vdots$$

$$\mathbf{v}_k = \mathbf{x}_k - \left(\frac{\mathbf{v}_1 \cdot \mathbf{x}_k}{\mathbf{v}_1 \cdot = v_1}\right)\mathbf{v}_1 - \left(\frac{\mathbf{v}_2 \cdot \mathbf{x}_k}{\mathbf{v}_2 \cdot \mathbf{v}_2}\right)\mathbf{v}_2 - \cdots - \left(\frac{\mathbf{v}_{k-1} \cdot \mathbf{x}_k}{\mathbf{v}_{k-1} \cdot \mathbf{v}_{k-1}}\right)\mathbf{v}_{k-1}, \qquad W_k = \text{span}(\mathbf{x}_1, \ldots, \mathbf{x}_k)$$

  Then for each $i = 1, \ldots, k, \{\mathbf{v}_1, \ldots, \mathbf{v}_i\}$ is an orthogonal basis for W_i. In particular, $\{\mathbf{v}_1, \ldots, \mathbf{v}_i\}$ is an orthogonal basis for W.

- QR Factorization: Let A be an $m \times n$ matrix with linearly independent columns. Then A can be factored as $A = QR$, where Q is an $m \times n$ matrix with orthonormal columns and R is an invertible upper triangular matrix.

- Finding the QR factorization: find an orthonormal basis for $\text{col}(A)$ using the Gram-Schmidt Process. Then, $Q = [\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_k]$. Then, use the fact that $A = QR$ and $Q^T Q = I$ since Q has orthonormal columns. Therefore $Q^T A = Q^T QR = IR = R$

5.4 Orthogonal Diagonalization of Symmetric Matrices

- A square matrix A is orthogonally diagonalizable if there exists an orthogonal matrix Q and a diagonal matrix D such that $Q^T AQ = D$

- If A is orthogonally diagonalizable, then A is symmetric.

- If A is a real symmetric matrix, then the eigenvalues of A are real.

- If A is a symmetric matrix, then any two eigenvectors corresponding to distinct eigenvalues of A are orthonormal.

- The spectral theorem: Let A be an $n \times n$ real matrix. Then A is symmetric IFF it is orthogonally diagonalizable.

- Spectral decomposition:

$$A = \lambda_1 \mathbf{q}_1 \mathbf{q}_1^T + \lambda_2 \mathbf{q}_2 \mathbf{q}_2^T + \cdots + \lambda_n \mathbf{q}_n \mathbf{q}_n^T$$

5.5 Applications

- A quadratic form in n variables is a function $f : \mathbb{R}^n \to \mathbb{R}$ of the form

$$f(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$$

  where A is a symmetric $n \times n$ matrix and $\mathbf{x}$ is in $\mathbb{R}^n$. We refer to A as the matrix associated with f.

- The principal axes theroem: Every quadratic form can be diagonalized. Specifically, if A is the $n \times n$ symmetric matrix associated with the quadratic form $\mathbf{x}^T A\mathbf{x}$ and if Q is an orthogonal matrix such that $Q^T AQ = D$ is a diagonal matrix, then the change of variable $\mathbf{x} = Q\mathbf{y}$ transforms the quadratic form $\mathbf{x}^T A\mathbf{x}$ into the quadratic form $\mathbf{y}^T D\mathbf{y}$, which has no cross-product terms. If the eigenvalues of A are $\lambda_1, \ldots, \lambda_n$ and $\mathbf{y} = [y_1 \cdots y_n]^T$, then

$$\mathbf{x}^T A\mathbf{x} = \mathbf{y}^T D\mathbf{y} = \lambda_1 y_a^2 + \cdots + \lambda_n y_n^2$$

- A quadratic form $f(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ is classified as one of the following:
  - positive definite if $f(\mathbf{x}) > 0$ for all $\mathbf{x} \neq \mathbf{0}$
  - positive semidefinite if $f(\mathbf{x}) \geq 0$ for all $\mathbf{x}$
  - negative definite if $f(\mathbf{x}) < 0$ for all $\mathbf{x} \neq \mathbf{0}$
  - negative semidefinite if $f(\mathbf{x}) \leq 0$ for all $\mathbf{x}$
  - indefinite if $f(\mathbf{x})$ takes on both positive and negative values
- A symmetric matrix A is called positive definite, positive semidefinite, negative definite, negative semidefinite, or indefinite if the associated quadratic form $f(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ has the corresponding property.
- Let A be an $n \times n$ symmetric matrix. The quadratic form $f(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ is
  - Positive definite IFF all eigenvalues of A are positive.
  - positive semidefinite IFF all eigenvalues are nonnegative.
  - negative definite IFF all eigenvalues are negative
  - negative semidefinite IFF all eigenvalues are nonpositive.
  - indefinite IFF A has both positive and negative eigenvalues.
- Let $f(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ be a quadratic form with associated $n \times n$ symmetric matrix A. Let the eigenvalues of A be $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$. Then the following are true, with the constraint of $||\mathbf{x}|| = 1$:
  - $\lambda_1 \geq f(\mathbf{x}) \geq \lambda_n$
  - The max value of $f(\mathbf{x})$ is λ_1 and occurs when $\mathbf{x}$ is a unit eigenvector corresponding to λ_1
  - The min value of $f(\mathbf{x})$ is λ_n and occurs when $\mathbf{x}$ is a unit eigenvector corresponding to λ_n
- The general form of a quadratic equation in two variables x and y is

$$ax^2 + by^2 + cxy + dx + ey + f = 0$$

- The general form of a quadratic equation in three variables x, y, and z is

$$ax^2 + by^2 + cz^2 + dxy + exz + fyz + gx + hy + iz + j = 0$$