CHAPTER 6: INNER PRODUCT SPACES

Section 6.1: Inner Products

Definition. 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 $\mathbf{u}, \mathbf{v}, \mathbf{w}$ 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**.

Definition. If V is a real inner product space, then the **norm** (or **length**) of a vector $\mathbf{v}$ in V is defined by $||\mathbf{v}|| = \sqrt{\langle \mathbf{v}, \mathbf{v}\rangle}$. The **distance** between two vectors is defined by $d(\mathbf{u}, \mathbf{v}) = ||\mathbf{u} - \mathbf{v}||\ \sqrt{\langle \mathbf{u} - \mathbf{v}, \mathbf{u} - \mathbf{v}\rangle}$. A vector of norm 1 is called a **unit vector**.

Theorem. If $\mathbf{u}$ and $\mathbf{v}$ are vectors in a real inner product space V, and if k is a scalar, then:
(a) $||\mathbf{v}|| \geq 0$ with equality if and only if $\mathbf{v} = \mathbf{0}$
(b) $||k\mathbf{v}|| = |k|\ ||\mathbf{v}||$
(c) $d(\mathbf{u},\mathbf{v}) = d(\mathbf{v},\mathbf{u})$
(d) $d(\mathbf{u},\mathbf{v}) \geq 0$ with equality if and only if $\mathbf{u} = \mathbf{v}$

- $\langle \mathbf{u}, \mathbf{v}\rangle = w_1u_1v_1 + w_2u_2v_2 + \ldots + w_nu_nv_n$
- $\langle \mathbf{u}, \mathbf{v}\rangle = A\mathbf{u} \bullet A\mathbf{v}$
- $\langle \mathbf{u}, \mathbf{v}\rangle = \mathbf{v}^TA^TA\mathbf{u}$
- $\langle U, V\rangle = \mathrm{tr}(U^TV)$

Theorem. 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) $k\langle \mathbf{u}, \mathbf{v}\rangle = \langle \mathbf{u}, k\mathbf{v}\rangle$

Section 6.2: Angle and Orthogonality in Inner Product Spaces

Theorem. **Cauchy-Schwarz Inequality**: If $\mathbf{u}$ and $\mathbf{v}$ are vectors in a real inner product space V, then $|\langle \mathbf{u}, \mathbf{v}\rangle| \leq ||\mathbf{u}||\ ||\mathbf{v}||$.

- $\langle \mathbf{u}, \mathbf{v}\rangle^2 \leq \langle \mathbf{u}, \mathbf{u}\rangle\langle \mathbf{v}, \mathbf{v}\rangle$
- $\langle \mathbf{u}, \mathbf{v}\rangle^2 \leq ||\mathbf{u}||^2\ ||\mathbf{v}||^2$
- $\theta = \cos^{-1}[\langle \mathbf{u}, \mathbf{v}\rangle/(||\mathbf{u}||\ ||\mathbf{v}||)]$

Theorem. If $\mathbf{u}, \mathbf{v}$ and $\mathbf{w}$ are vectors in a real inner product space V, and if k is any scalar, then:
(a) $||\mathbf{u}+\mathbf{v}|| \leq ||\mathbf{u}|| + ||\mathbf{v}||$ (**Triangle Inequality for Vectors**)
(b) $d(\mathbf{u},\mathbf{v}) \leq d(\mathbf{u},\mathbf{w}) + d(\mathbf{w},\mathbf{v})$ (**Triangle Inequality for Distances**)

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

Theorem. **Generalized Theorem of Pythagoras.**
If $\mathbf{u}$ and $\mathbf{v}$ are orthogonal vectors in an inner product space, then $||\mathbf{u} + \mathbf{v}||^2 = ||\mathbf{u}||^2 + ||\mathbf{v}||^2$

Definition. If W is a subspace of an inner product space V, then the set of all vectors in V that are orthogonal to every vector in W is called the **orthogonal complement** of W and is denoted by the symbol $W^\perp$.

Theorem. If W is a subspace of an inner product space V, then:
(a) $W^\perp$ is a subspace of V
(b) $W \cap W^\perp = \{\mathbf{0}\}$

Theorem. If W is a subspace of a finite-dimensional inner product space V, then the orthogonal complement of $W^\perp$ is W; that is, $(W^\perp)^\perp = W$.

Section 6.3: Gram-Schmidt Process; *QR*-Decomposition

Definition. A set of two or more vectors in a real inner product space is said to be **orthogonal** if all pairs of distinct vectors in the set are orthogonal. An orthogonal set in which each vector has norm 1 is said to be **orthonormal**.

Theorem. If $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is an orthogonal set of nonzero vectors in an inner product space, then S is linearly independent.

Theorem. (a) If $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is an orthogonal basis for an inner product space V, and if $\mathbf{u}$ is any vector in V, then
$\mathbf{u} = [\langle \mathbf{u}, \mathbf{v}_1\rangle/||\mathbf{v}_1||^2]\mathbf{v}_1 + [\langle \mathbf{u}, \mathbf{v}_2\rangle/||\mathbf{v}_2||^2]\mathbf{v}_2 + \ldots + [\langle \mathbf{u}, \mathbf{v}_n\rangle/||\mathbf{v}_n||^2]\mathbf{v}_n$

(*b*) If $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is an orthonormal basis for an inner product space V, and if $\mathbf{u}$ is any vector in V, then $\mathbf{u} = \langle \mathbf{u}, \mathbf{v}_1\rangle\mathbf{v}_1 + \langle \mathbf{u}, \mathbf{v}_2\rangle\mathbf{v}_2 + \ldots + \langle \mathbf{u}, \mathbf{v}_n\rangle\mathbf{v}_n$

Theorem. **Projection Theorem**: If W is a finite-dimensional subspace of an inner product space V, then every vector $\mathbf{u}$ in V can be expressed in exactly one way as $\mathbf{u} = \mathbf{w}_1 + \mathbf{w}_2$ where $\mathbf{w}_1$ is in W and $\mathbf{w}_2$ is in $W^{\perp}$.

- $\mathbf{u} = \text{proj}_W\,\mathbf{u} + \text{proj}_{W\perp}\,\mathbf{u} = \text{proj}_W\,\mathbf{u} + (\mathbf{u} - \text{proj}_W\,\mathbf{u})$

Theorem. Let W be a finite-dimensional subspace of an inner product space V.
(*a*) If $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$ is an orthogonal basis for W, and $\mathbf{u}$ is any vector in V, then
$\text{proj}_W\,\mathbf{u} = [\langle \mathbf{u}, \mathbf{v}_1\rangle/\|\mathbf{v}_1\|^2]\mathbf{v}_1 + [\langle \mathbf{u}, \mathbf{v}_2\rangle/\|\mathbf{v}_2\|^2]\mathbf{v}_2 + \ldots + [\langle \mathbf{u}, \mathbf{v}_r\rangle/\|\mathbf{v}_r\|^2]\mathbf{v}_r$.
(*b*) If $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$ is an orthonormal basis for W, and $\mathbf{u}$ is any vector in V, then
$\text{proj}_W\,\mathbf{u} = \langle \mathbf{u}, \mathbf{v}_1\rangle\mathbf{v}_1 + \langle \mathbf{u}, \mathbf{v}_2\rangle\mathbf{v}_2 + \ldots + \langle \mathbf{u}, \mathbf{v}_r\rangle\mathbf{v}_r$.

Theorem. Every nonzero finite-dimensional inner product space has an orthonormal basis.

- **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: $\mathbf{v}_2 = \mathbf{u}_2 - [\langle \mathbf{u}_2, \mathbf{v}_1\rangle/\|\mathbf{v}_1\|^2]\mathbf{v}_1$
Step 3: $\mathbf{v}_3 = \mathbf{u}_3 - [\langle \mathbf{u}_3, \mathbf{v}_1\rangle/\|\mathbf{v}_1\|^2]\mathbf{v}_1 - [\langle \mathbf{u}_3, \mathbf{v}_2\rangle/\|\mathbf{v}_2\|^2]\mathbf{v}_2$
Step 4: $\mathbf{v}_4 = \mathbf{u}_4 - [\langle \mathbf{u}_4, \mathbf{v}_1\rangle/\|\mathbf{v}_1\|^2]\mathbf{v}_1 - [\langle \mathbf{u}_4, \mathbf{v}_2\rangle/\|\mathbf{v}_2\|^2]\mathbf{v}_2 - [\langle \mathbf{u}_4, \mathbf{v}_3\rangle/\|\mathbf{v}_3\|^2]\mathbf{v}_3$
$\vdots$
(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.

Theorem. If W is a finite-dimensional inner product space, then:
(*a*) Every orthogonal set of nonzero vectors in W can be enlarged to an orthogonal basis for W.
(*b*) Every orthonormal set in W can be enlarged to an orthonormal basis for W.

- **Problem**: If A is an $m \times n$ matrix with linearly independent column vectors, and if Q is the matrix that results by applying the Gram-Schmidt process to the column vectors of A, what relationship, if any, exists between A and Q?

Theorem. ***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.

Section 6.4: Best Approximation; Least Squares

- **Least Squares Problem**: Given a linear system $A\mathbf{x} = \mathbf{b}$ of m equations in n unknowns, find a vector $\mathbf{x}$ that minimizes $\|\mathbf{b} - A\mathbf{x}\|$ with respect to the Euclidean inner product on $\boldsymbol{R}^m$. We call such an $\mathbf{x}$ a **least squares solution** of the system, we call $\mathbf{b} - A\mathbf{x}$ the **least squares error vector**, and we call $\|\mathbf{b} - A\mathbf{x}\|$ the **least squares error**.

Theorem. **Best Approximation Theorem**: If W is a finite-dimensional subspace of an inner product space V, and if $\mathbf{b}$ is a vector in V, then $\text{proj}_W\,\mathbf{b}$ is the **best approximation** to $\mathbf{b}$ from W in the sense than $\|\mathbf{b} - \text{proj}_W\,\mathbf{b}\| < \|\mathbf{b} - \mathbf{w}\|$ for every vector $\mathbf{w}$ in W that is different from $\text{proj}_W\,\mathbf{b}$.

Theorem. For every linear system $A\mathbf{x} = \mathbf{b}$, the associated normal system $A^TA\mathbf{x} = A^T\mathbf{b}$ is consistent, and all solutions of the normal system are least squares solutions of $A\mathbf{x} = \mathbf{b}$. Moreover, if W is the column space of A, and $\mathbf{x}$ is any least squares solution of $A\mathbf{x} = \mathbf{b}$, then the orthogonal projection of $\mathbf{b}$ on W is $\text{proj}_W\,\mathbf{b} = A\mathbf{x}$.

Theorem. If A is an $m \times n$ matrix, then the following are equivalent:
(*a*) A has linearly independent column vectors.
(*b*) A^TA is invertible.

Theorem. If A is an $m \times n$ matrix with linearly independent column vectors, then for every $m \times 1$ matrix $\mathbf{b}$, the linear system $A\mathbf{x} = \mathbf{b}$ has a unique least squares solution. This solution is given by $\mathbf{x} = (A^TA)^{-1}A^T\mathbf{b}$. Moreover, if W is the column space of A, then the orthogonal projection of $\mathbf{b}$ on W is $\text{proj}_W\,\mathbf{b} = A\mathbf{x} = A(A^TA)^{-1}A^T\mathbf{b}$.

Theorem. If A is an $m \times n$ matrix with linearly independent column vectors, and if $A = QR$ is a QR-decomposition of A, then for each $\mathbf{b}$ in $\boldsymbol{R}^m$ the system $A\mathbf{x} = \mathbf{b}$ has a unique least squares solution given by $\mathbf{x} = R^{-1}Q^T\mathbf{b}$.

Definition. If W is a subspace of $\boldsymbol{R}^m$, then the transformation P: $\boldsymbol{R}^m \to W$ that maps each vector $\mathbf{x}$ in $\boldsymbol{R}^m$ into its orthogonal projection $\text{proj}_W\,\mathbf{x}$ in W is called the **orthogonal projection of $\boldsymbol{R}^m$ on W**.

- $[P] = A(A^TA)^{-1}A^T$

Theorem. **Equivalent Statements**: If A is an $n \times n$ matrix, then the following statements are equivalent:
(*a*) A is invertible
(*b*) $A\mathbf{x} = \mathbf{0}$ has only the trivial solution
(*c*) the reduced row-echelon form of A is I_n
(*d*) A is expressible as a product of elementary matrices
(*e*) $A\mathbf{x} = \mathbf{b}$ is consistent for every $n \times 1$ matrix $\mathbf{b}$
(*f*) $A\mathbf{x} = \mathbf{b}$ has exactly one solution for every $n \times 1$ matrix $\mathbf{b}$
(*g*) $\det(A) \neq 0$
(*h*) The column vectors of A are linearly independent.

(i) The row vectors of A are linearly independent.
(j) The column vectors of A span $\boldsymbol{R}^n$.
(k) The row vectors of A span $\boldsymbol{R}^n$.
(l) The column vectors of A form a basis for $\boldsymbol{R}^n$.
(m) The row vectors of A form a basis for $\boldsymbol{R}^n$.
(n) A has rank n
(o) A has nullity 0
(p) The orthogonal complement of the null space of A is $\boldsymbol{R}^n$.
(q) The orthogonal complement of the row space of A is $\mathbf{0}$.
(r) The range of T_A is $\boldsymbol{R}^n$
(s) T_A is one-to-one.
(t) $\lambda = 0$ is not an eigenvalue of A.
(u) A^TA is invertible