Chapter 7

7.1

Remember that $E_1 \cdots E_k A = R$.

Four fundamental subspaces of an $m \times n$ matrix A:

1. Column space: $\{A\vec{x} \in \mathbb{R}^m \mid \vec{x} \in \mathbb{R}^n\}$
2. Row space: $\{A^T\vec{x} \in \mathbb{R}^n \mid \vec{x} \in \mathbb{R}^m\}$
3. Null space: $\{\vec{x} \in \mathbb{R}^n \mid A\vec{x} = \vec{0}\}$
4. Left null space: $\{\vec{x} \in \mathbb{R}^m \mid A^T\vec{x} = \vec{0}\}$

Note that $\mathrm{Col}(A^T) = \mathrm{Row}(A)$ and $\mathrm{Col}(A)$ is the span of the columns of A.

Theorem 7.1.1: If $A \in M_{m\times n}(\mathbb{R})$, then the column space and left nullspace are subspaces of $\mathbb{R}^m$ and the row space and null space are subspaces of $\mathbb{R}^n$.

Theorem 7.1.2: Let the RREF of A be R. The columns of A which correspond to pivot columns in R form a basis for $\mathrm{Col}(A)$ and $\dim \mathrm{Col}(A) = \mathrm{rank}\, A$. Prove this by first showing that the pivot columns of R form a basis for $\mathrm{Col}(R)$.

Theorem 7.1.3: If R is an $m \times n$ matrix and E is an invertible $n \times n$ matrix, then $\mathrm{Col}(RE) = \mathrm{Col}(R)$, or $\{RE\vec{x} \mid \vec{x} \in \mathbb{R}^n\} = \{R\vec{y} \mid \vec{y} \in \mathbb{R}^n\}$.

Theorem 7.1.4: If A is an $m \times n$ matrix, then the non-zero rows of the RREF of A form a basis for $\mathrm{Row}(A)$. And $\dim \mathrm{Row}(A) = \mathrm{rank}\, A$. As well, $\mathrm{Row}(A) = \mathrm{Row}(R)$, where R is the RREF of A, since by Theorem 7.1.3, we have $EA = R$, or $A^T = R^T(E^{-1})^T$, where $(E^{-1})^T$ is invertible.

Theorem 7.1.5: For any matrix $A, \mathrm{rank}\, A = \mathrm{rank}\, A^T$.

Theorem 7.1.6: If A is an $m \times n$ matrix, then $\mathrm{rank}\, A + \dim \mathrm{Null}(A) = n$. The outline of the proof is as follows: Create a k dimensional basis for $\mathrm{Null}(A)$ then extend to n vectors which is a basis for $\mathbb{R}^n$. Then, show that $\{A\vec{v}_{k+1}, \ldots, A\vec{v}_n\}$ is a basis for $\mathrm{Col}(A)$.

Note: Be careful reducing a matrix, I almost always make a small mistake somewhere.

Chapter 8: Linear Mappings

8.1 General Linear Mappings

Let $\mathbb{V}, \mathbb{W}$ be real vector spaces. A mapping $L : \mathbb{V} \to \mathbb{W}$ is linear if $L(s\vec{x} + t\vec{y}) = sL(\vec{x}) + tL(\vec{y})$ for all $\vec{x}, \vec{y} \in \mathbb{V}$ and $s, t \in \mathbb{R}$. Two linear mappings are equal iff $L(\vec{v}) = M(\vec{v})$ for all $\vec{v} \in \mathbb{V}$. If $L : \mathbb{V} \to \mathbb{V}$, then L is a linear operator.

Let $L : \mathbb{V} \to \mathbb{W}$ and $M : \mathbb{V} \to \mathbb{W}$ be linear mappings. Then $L+M : \mathbb{V} \to \mathbb{W}$ is defined $(L+M)(\vec{v}) = L(\vec{v})+M(\vec{v})$ and for any $t \in \mathbb{R}$, we define $tL : \mathbb{V} \to \mathbb{W}$ as $(tL)(\vec{v}) = tL(\vec{v})$.

Theorem 8.1.1: Let $\mathbb{V}, \mathbb{W}$ be vector spaces. The set $\mathbb{L}$ of all linear mappings $L : \mathbb{V} \to \mathbb{W}$ with standard addition and scalar multiplication of mappings is a vector space.

Let $L : \mathbb{V} \to \mathbb{W}$ and $M : \mathbb{W} \to \mathbb{U}$ be linear mappings. Then $M \circ L : \mathbb{V} \to \mathbb{U}$ is defined $(M \circ L)(\vec{v}) = M(L(\vec{v}))$ for all $\vec{v} \in \mathbb{V}$.

Theorem 8.1.2: If $L : \mathbb{V} \to \mathbb{W}$ and $M : \mathbb{W} \to \mathbb{U}$ are linear mappings, then $M \circ L : \mathbb{V} \to \mathbb{U}$ is a linear mapping.

Invertible Mapping: Let $L : \mathbb{V} \to \mathbb{W}$ and $M : \mathbb{W} \to \mathbb{V}$ be linear mappings. If $(M \circ L)(\vec{v}) = \vec{v}\ \forall \vec{v} \in \mathbb{V}$ and $(L \circ M)(\vec{w}) = \vec{w}\ \forall \vec{w} \in \mathbb{W}$, then L and M are invertible and $M = L^{-1}, L = M^{-1}$. All invertible mappings are isomorphisms. Thus, $\dim \mathbb{V} = \dim \mathbb{W}$.

8.2 Rank-Nullity Theorem

Let $L : \mathbb{V} \to \mathbb{W}$ be a linear mapping. Then:

$$\mathrm{Ker}(L) = \{\vec{v} \in \mathbb{V} | L(\vec{v}) = \vec{0}_{\mathbb{W}}\}, \mathrm{Range}(L) = \{L(\vec{v}) \in \mathbb{W} | \vec{v} \in \mathbb{V}\}$$

Theorem 8.2.2: If $L : \mathbb{V} \to \mathbb{W}$ is a linear mapping, then $\mathrm{Ker}(L)$ is a subspace of $\mathbb{V}$ and $\mathrm{Range}(L)$ is a subspace of $\mathbb{W}$.

For a linear mapping L, $\mathrm{rank}(L) = \dim \mathrm{Range}(L)$ and $\mathrm{nullity}(L) = \dim \mathrm{Ker}(L)$. Note that to show that for two linear mappings L, M, to show that $\mathrm{rank}(L) \leq \mathrm{rank}(M)$, we can show that $\mathrm{Range}(L) \subseteq \mathrm{Range}(M)$.

Theorem 8.2.3: Rank-Nullity Theorem: Let $\mathbb{V}$ be an n-dimensional vector space and let $\mathbb{W}$ be a vector space. If $L : \mathbb{V} \to \mathbb{W}$ is linear, then $\mathrm{rank}(L) + \mathrm{nullity}(L) = n$. The proof of this is basically the same as Dimension Theorem.

8.3 Matrix of a Linear Mapping

Every linear mapping $L : \mathbb{V} \to \mathbb{W}$ can be represented as a matrix mapping. Using coordinates, we can write a matrix mapping representation for a linear mapping.

Matrix of a linear mapping: Suppose $\mathcal{B} = \{\vec{v}_1, \dots, \vec{v}_n\}$ is any basis for a vector space $\mathbb{V}$ and $\mathcal{C}$ is any basis for a finite dimensional vector space $\mathbb{W}$. For a linear mapping $L : \mathbb{V} \to \mathbb{W}$, the matrix of L with respect to $\mathcal{B}$ and $\mathcal{C}$ is: ${}_{\mathcal{C}}[L]_{\mathcal{B}} = \begin{bmatrix} [L(\vec{v}_1)]_{\mathcal{C}} & \cdots & [L(\vec{v}_n)]_{\mathcal{C}} \end{bmatrix}$ and satisfies $[L(\vec{x})]_{\mathcal{C}} = {}_{\mathcal{C}}[L]_{\mathcal{B}}[\vec{x}]_{\mathcal{B}}$

Matrix of a linear operator: Suppose $\mathcal{B} = \{\vec{v}_1, \dots, \vec{v}_n\}$ is any basis for a vector space $\mathbb{V}$ and let $L : \mathbb{V} \to \mathbb{V}$ be a linear operator. Then the $\mathcal{B}-$matrix of L is: $[L]_{\mathcal{B}} = \begin{bmatrix} [L(\vec{v}_1)]_{\mathcal{B}} & \cdots & [L(\vec{v}_n)]_{\mathcal{B}} \end{bmatrix}$ and satisfies $[L(\vec{x})]_{\mathcal{B}} = [L]_{\mathcal{B}}[\vec{x}]_{\mathcal{B}}$. Note that this matrix is $n \times n$. Important properties:

1. $[L(\vec{v})]_C = {}_C[L]_B[\vec{v}]_B$ for all $\vec{v} \in \mathbb{V}$
2. ${}_C[L + M]_B = {}_C[L]_B + {}_C[M]_B$ for two linear mappings $L, M : \mathbb{V} \to \mathbb{W}$
3. ${}_C[sL]_B = s({}_C[L]_B)$ for all $s \in \mathbb{R}$
4. Suppose $L : \mathbb{V} \to \mathbb{W}, M : \mathbb{W} \to \mathbb{U}$ and B, C, D are bases for V, W, U respectively. Then, ${}_D[M \circ L]_B = {}_D[M]_{C\,C}[L]_B$

8.4 Isomorphisms

Let $L : \mathbb{V} \to \mathbb{W}$ be a linear mapping. L is one to one if $\forall \vec{u}, \vec{v} \in \mathbb{V}$ such that $L(\vec{u}) = L(\vec{v})$, then $\vec{u} = \vec{v}$. L is onto if $\forall \vec{w} \in \mathbb{W}, \exists \vec{v} \in \mathbb{V}$ such that $L(\vec{v}) = \vec{w}$. Or, Range$(L) = \mathbb{W}$.

Theorem 8.4.1: A linear mapping is one to one iff $\mathrm{Ker}(L) = \{\vec{0}\}$. The easiest way to show one-to-one is to show that $\mathrm{Ker}(L) = \{\vec{0}\}$.

The purpose of an isomorhpism between two vector spaces is to have a linear mapping such that each vector in one space is uniquely identified with a unique vector in the other, so that the two spaces have the same structure.

A vector space $\mathbb{V}$ is isomorphic to $\mathbb{W}$ if there is a linear mapping between the two that is one to one and onto. This mapping is an isomorphism from $\mathbb{V}$ to $\mathbb{W}$. To show that two vector spaces are isomorphic, construct an isomorhpism between the two.

Theorem 8.4.2: Let $\mathbb{V}, \mathbb{W}$ be finite dimensional vector spaces. $\mathbb{V}$ is isomorphic to $\mathbb{W}$ iff $\dim \mathbb{V} = \dim \mathbb{W}$. To prove this, consider a mapping which maps the basis vectors of $\mathbb{V}$ to the basis vectors of $\mathbb{W}$.

Notes: If $\mathbb{V}$ is isomorphic to $\mathbb{W}$, then $\mathbb{W}$ is isomorphic to $\mathbb{V}$. We can make an isomorphism from $\mathbb{V}$ to $\mathbb{W}$ by mapping basis vectors of $\mathbb{V}$ to basis vectors of $\mathbb{W}$.

Theorem 8.4.3: If $\mathbb{V}, \mathbb{W}$ are the same dimension and $L : \mathbb{V} \to \mathbb{W}$ is linear, then L is one to one iff L is onto. From Theorem 8.4.2, we know that $\mathbb{V}$ and $\mathbb{W}$ must be isomorphic, however, this theorem does not say that every mapping $L : \mathbb{V} \to \mathbb{W}$ is an isomorphism.

Theorem 8.4.4: Let $\mathbb{V}, \mathbb{W}$ be isomorphic vector spaces and let $\{\vec{v}_1, \dots, \vec{v}_n\}$ be a basis for $\mathbb{V}$. Then a linear mapping $L : \mathbb{V} \to \mathbb{W}$ is an isomorphism iff $\{L(\vec{v}_1), \dots, L(\vec{v}_n)\}$ is a basis for $\mathbb{W}$.

More generally, suppose $L : \mathbb{V} \to \mathbb{W}$ and $B = \{\vec{v}_1, \dots, \vec{v}_n\}$ is a basis for $\mathbb{V}$. Then $\mathrm{Range}(L) = \mathrm{Span}\{L(\vec{v}_1), \dots, L(\vec{v}_n)\}$.

Facts about isomorhpisms:

1. If $L : \mathbb{V} \to \mathbb{W}, M : \mathbb{W} \to \mathbb{U}$ are isomorphisms, then $(M \circ L) : \mathbb{V} \to \mathbb{U}$ is an isomorphism.
2. All n-dimensional vector spaces $\mathbb{V}$ are isomorphic to $\mathbb{R}^n$. Let $\mathcal{B}$ be a basis for $\mathbb{V}$. Then $T : \mathbb{V} \to \mathbb{R}^n$ where $T(\vec{v}) = [\vec{v}]_B$ is an isomorphism.

3. If $\mathcal{B}$ is a basis for $\mathbb{V}$ and $\mathcal{C}$ is a basis for $\mathbb{W}$, then $L : \mathbb{V} \to \mathbb{W}$ is an isomorphism iff ${}_C[L]_B$ is invertible. Where $({}_C[L]_B)^{-1} = {}_B[L^{-1}]_C$

Chapter 9: Inner Products

9.1: Inner Product Spaces

Let $\mathbb{V}$ be a finite or infinite dimensional vector space. An inner product is a function $\langle\ ,\ \rangle : \mathbb{V} \times \mathbb{V} \to \mathbb{R}$ such that:

1. $\langle \vec{v}, \vec{v} \rangle \geq 0$ and $\langle \vec{v}, \vec{v} \rangle = 0$ iff $\vec{v} = \vec{0}$ - positive definite
2. $\langle \vec{v}, \vec{w} \rangle = \langle \vec{w}, \vec{v} \rangle$ - symmetric
3. $\langle s\vec{v} + t\vec{u}, \vec{w} \rangle = s\langle \vec{v}, \vec{w} \rangle + t\langle \vec{u}, \vec{w} \rangle$ - left linear
4. $\langle \vec{w}, s\vec{v} + t\vec{u} \rangle = s\langle \vec{w}, \vec{v} \rangle + t\langle \vec{w}, \vec{u} \rangle$
5. $\langle s\vec{v} + t\vec{u}, p\vec{x} + q\vec{y} \rangle = sp\langle \vec{v}, \vec{x} \rangle + sq\langle \vec{v}, \vec{y} \rangle + tp\langle \vec{u}, \vec{x} \rangle + tq\langle \vec{u}, \vec{y} \rangle$

To verify that a function is an inner product, verify: property 1, 2, and 3.

A vector space with an inner product is an inner product space. The standard inner product on $\mathbb{R}^n$ is the dot product and the standard inner product on $M_{m\times n}(\mathbb{R})$ is $\mathrm{tr}(B^T A)$.

Let $B = [\vec{b}_1 \cdots \vec{b}_n], A = [\vec{a}_1 \cdots \vec{a}_n]$. Then $\mathrm{tr}(B^T A) = \vec{b}_1 \cdot \vec{a}_1 + \cdots + \vec{b}_n \cdot \vec{a}_n$.

Theorem 9.1.1: If $\mathbb{V}$ is a vector space with inner product $\langle\ ,\ \rangle$, then $\forall\, \vec{v} \in \mathbb{V}$, then $\langle \vec{v}, \vec{0} \rangle = 0$. To prove this, notice that $\langle \vec{v}, \vec{0} \rangle = \langle \vec{v}, -\vec{0} \rangle = -\langle \vec{v}, \vec{0} \rangle$.

Note that if $p(x) \in P_n(\mathbb{R})$ has $n+1$ roots, then $p(x) = \vec{0}$.

9.2: Orthogonality and Length

Let $\mathbb{V}$ be an inner product space with inner product $\langle\ ,\ \rangle$. Then the length (norm) of any $\vec{v} \in \mathbb{V}$ is $||\vec{v}|| = \sqrt{\langle \vec{v}, \vec{v} \rangle}$. A vector with length 1 is a unit vector.

Theorem 9.2.1: Let $\mathbb{V}$ be an inner product space. For any $\vec{v}, \vec{y} \in \mathbb{V}$ and $t \in \mathbb{R}$, then:

1. $||\vec{v}|| \geq 0$ and $||\vec{v}|| = 0$ iff $\vec{v} = \vec{0}$
2. $||t\vec{v}|| = |t|\,||\vec{v}||$
3. $\langle \vec{v}, \vec{y} \rangle \leq ||\vec{v}||\,||\vec{y}||$, Cauchy-Schwarz Inequality
4. $||\vec{v} + \vec{y}|| \leq ||\vec{v}|| + ||\vec{y}||$, triangle inequality

Normalizing a vector: Dividing a vector by its length creates a vector in the same direction but with length 1.

Orthogonality: If two vectors in the same inner product space have inner product 0, then they are orthogonal. Whether two vectors are orthogonal depends on the definition of the inner product space.

Orthogonal set: If $\mathcal{S} = \{\vec{v}_1, \ldots, \vec{v}_k\}$ is a set of vectors in an inner product space such that $\langle \vec{v}_i, \vec{v}_j \rangle = 0$ for all $i \neq j$, then $\mathcal{S}$ is an orthogonal set.

Theorem 9.2.2: If $\{\vec{v}_1, \ldots, \vec{v}_k\}$ is an orthogonal set with inner product space $\mathbb{V}$, then $||\vec{v}_1 + \cdots + \vec{v}_k||^2 = ||\vec{v}_1||^2 + \cdots + ||\vec{v}_k||^2$.

Theorem 9.2.3: If a set of vectors in an inner product space is orthogonal and does not contain the zero vector, then it is linearly independent. Thus, if we have such an orthogonal set with n non-zero vectors in an n-dimensional vector space, then it is a basis. Note that the same set may not be orthogonal under another inner product, but as long as it is with one, then it is linearly independent.

Orthogonal basis: an orthogonal set in an inner product space which is a basis for the inner product space.

Theorem 9.2.4: If $\mathcal{S} = \{\vec{v}_1, \ldots, \vec{v}_n\}$ is an orthogonal basis for an inner product space $\mathbb{V}$ and $\vec{v} \in \mathbb{V}$, then $\vec{v} = \frac{\langle \vec{v}, \vec{v}_1 \rangle}{||\vec{v}_1||^2}\vec{v}_1 + \cdots + \frac{\langle \vec{v}, \vec{v}_n \rangle}{||\vec{v}_n||^2}\vec{v}_n$. The coefficient of $\vec{v}_i$ is $\frac{\langle \vec{v}, \vec{v}_i \rangle}{||\vec{v}_i||^2}$. We can use this to find the $\mathcal{B}$-coordinate vector of $\vec{v}$ easily.

Orthonormal set: An orthogonal set in an inner product space where all the vectors have length 1.

Orthonormal basis: A basis for an inner product space which is an orthonormal set. Given an orthonormal set, we can make it an orthonormal basis by dividing each vector by the length.

If $\{\vec{v}_1, \ldots, \vec{v}_n\}$ is an orthonormal basis for $\mathbb{V}$ with respect to two inner products $\langle , \rangle$ and $\langle , \rangle_2$, then they are the same inner product.

Theorem 9.2.5: If $\mathbb{V}$ is an inner product space and $\vec{v} \in \mathbb{V}$ and $\mathcal{B} = \{\vec{v}_1, \ldots, \vec{v}_n\}$ is an orthonormal basis for $\mathbb{V}$, then $\vec{v} = \langle \vec{v}, \vec{v}_1 \rangle \vec{v}_1 + \cdots + \langle \vec{v}, \vec{v}_n \rangle \vec{v}_n$.

Orthogonal matrices: Theorem 9.2.6: For an $n \times n$ matrix P, the following are equivalent:

1. The columns of P form an orthonormal basis for $\mathbb{R}^n$
2. $P^T = P^{-1}$
3. The rows of P form an orthonormal basis for $\mathbb{R}^n$

Theorem 9.2.7: If P, Q are $n \times n$ orthogonal matrices and $\vec{x}, \vec{y} \in \mathbb{R}^n$, then

1. $(P\vec{x}) \cdot (P\vec{y}) = \vec{x} \cdot \vec{y}$
2. $||P\vec{x}|| = ||\vec{x}||$
3. $\det P = \pm 1$
4. All real eigenvalues of P are ± 1, complex eigenvalues $z = a + bi$ have $|z| = \sqrt{a^2 + b^2} = 1$.
5. PQ is an orthogonal matrix

9.3: Gram-Schmidt Orthogonalization Procedure

GSOP: given a spanning set (may be basis) for some finite dimensional inner product space, produces an orthogonal basis for that space.

GSOP: Let $\{\vec{w}_1, \ldots, \vec{w}_n\}$ be a basis for an inner product space $\mathbb{W}$. If we define these vectors: $\vec{v}_1 = \vec{w}_1, \vec{v}_i = \vec{w}_i - \frac{\langle \vec{w}_i, \vec{v}_1 \rangle}{\langle \vec{v}_1, \vec{v}_1 \rangle}\vec{v}_1 - \cdots - \frac{\langle \vec{w}_i, \vec{v}_{i-1} \rangle}{\langle \vec{v}_{i-1}, \vec{v}_{i-1} \rangle}\vec{v}_{i-1}$ for $2 \leq i \leq n$, then $\{\vec{v}_1, \ldots, \vec{v}_k\}$ is an orthogonal basis for $\text{Span}\{\vec{w}_1, \ldots, \vec{w}_k\}$ for $1 \leq k \leq n$. Note that if the order of the given basis is changed, the end result orthogonal basis will also likely be different.

Given an orthogonal, linearly independent set of non-zero vectors, if we are told to extend this to an orthogonal basis of some vector space, first extend the set using regular vectors then apply GSOP. Also, note that the set given by GSOP must be a basis - there are no zero vectors inside.

9.4: General Projections

Let $\mathbb{W}$ be a subspace of an inner product space $\mathbb{V}$. The orthogonal complement $\mathbb{W}^\perp$ of $\mathbb{W}$ in $\mathbb{V}$ is:

$$\mathbb{W}^\perp = \{\vec{v} \in \mathbb{V} | \langle \vec{w}, \vec{v} \rangle = 0, \ \forall \ \vec{w} \in \mathbb{W}\}$$

Theorem 9.4.1: Let $\{\vec{v}_1, \ldots, \vec{v}_k\}$ be a spanning set for a subspace $\mathbb{W}$ for an IPS $\mathbb{V}$ and let $\vec{x} \in \mathbb{V}$. Then $\vec{x} \in \mathbb{W}^\perp$ iff $\langle \vec{x}, \vec{v}_i \rangle = 0$ for $1 \leq i \leq k$. Thus, to check that a vector is in $\mathbb{W}^\perp$, we can just check against a spanning set, instead of every vector in $\mathbb{W}$.

Theorem 9.4.2: If $\mathbb{W}$ is a subspace of an IPS $\mathbb{V}$, then

1. $\mathbb{W}^\perp$ is a subspace of $\mathbb{V}$
2. If $\dim \mathbb{V} = n$, then $\dim \mathbb{W} + \dim \mathbb{W}^\perp = n$
3. If $\dim \mathbb{V} = n$, then $(\mathbb{W}^\perp)^\perp) = W$
4. $\mathbb{W} \cap \mathbb{W}^\perp = \{\vec{0}\}$
5. If $\dim \mathbb{V} = n, \{\vec{v}_1, \ldots, \vec{v}_k\}$ is an orthogonal basis for $\mathbb{W}$ and $\{\vec{v}_{k+1}, \ldots, \vec{v}_n\}$ is an orthogonal basis for $\mathbb{W}^\perp$, then $\{\vec{v}_1, \ldots, \vec{v}_k, \vec{v}_{k+1}, \ldots, \vec{v}_n\}$ is an orthogonal basis for $\mathbb{V}$.

Note that properties 2, 3, 5 require that $\mathbb{V}$ is a finite dimensional vector space.

We want to define the projection of a vector $\vec{v}$ onto a subspace $\mathbb{W}$ of an inner product space $\mathbb{V}$ such that $\vec{v} = \text{proj}_{\mathbb{W}}(\vec{v}) + \text{perp}_{\mathbb{W}}(\vec{v})$ and $\text{proj}_{\mathbb{W}}(\vec{v}) \in \mathbb{W}, \text{perp}_{\mathbb{W}}(\vec{v}) \in \mathbb{W}^\perp$. Let $\{\vec{v}_1, \ldots, \vec{v}_k\}$ be an orthogonal basis for $\mathbb{W}$ and $\{\vec{v}_{k+1}, \ldots, \vec{v}_n\}$ be an orthogonal basis for $\mathbb{W}^\perp$. Then $\{\vec{v}_1, \ldots, \vec{v}_n\}$ is an orthogonal basis for $\mathbb{V}$. Thus, for all $\vec{v} \in \mathbb{V}, \vec{v} = \frac{\langle \vec{v}, \vec{v}_1 \rangle}{||\vec{v}_1||^2}\vec{v}_1 + \cdots + \frac{\langle \vec{v}, \vec{v}_n \rangle}{||\vec{v}_n||^2}\vec{v}_n$ where $\text{proj}_{\mathbb{W}}(\vec{v}) = \frac{\langle \vec{v}, \vec{v}_1 \rangle}{||\vec{v}_1||^2}\vec{v}_1 + \cdots + \frac{\langle \vec{v}, \vec{v}_k \rangle}{||\vec{v}_k||^2}\vec{v}_k$ and $\text{perp}_{\mathbb{W}}(\vec{v}) = \vec{v} - \text{proj}_{\mathbb{W}}(\vec{v})$.

Theorem 9.4.3, 4, 5: If $\mathbb{W}$ is a k-dimensional subspace of an inner product space $\mathbb{V}$, then:

1. For any $\vec{v} \in \mathbb{V}$, $\text{perp}_{\mathbb{W}}(\vec{v}) \in \mathbb{W}^\perp$
2. $\text{proj}_{\mathbb{W}}$ is a linear operator on $\mathbb{V}$ with kernel $\mathbb{W}^\perp$
3. For any $\vec{v} \in \mathbb{V}, \text{proj}_{\mathbb{W}^\perp}(\vec{v}) = \text{perp}_{\mathbb{W}}(\vec{v})$

All vectors in an inner product space can be written $\vec{v} = \text{perp}_{\mathbb{W}}(\vec{v}) + \text{proj}_{\mathbb{W}}(\vec{v})$. Or, the sum of one vector in $\mathbb{W}$ and another in $\mathbb{W}^\perp$.

Additional notes about projection:

1. $\vec{v} \in \mathbb{W}$ iff $\text{proj}_{\mathbb{W}}(\vec{v}) = \vec{v}$, or $\text{perp}_{\mathbb{W}}(\vec{v}) = 0$
2. $\text{proj}_{\mathbb{W}}(\vec{v}) \in \mathbb{W}$
3. Projection onto a subspace is a linear mapping
4. $\langle \text{proj}_{\mathbb{W}}(\vec{v}), \text{perp}_{\mathbb{W}}(\vec{v}) \rangle = 0, \langle \text{perp}_{\mathbb{W}}(\vec{v}), \vec{x} \rangle = 0 \ \forall \vec{x} \in \mathbb{W}$.
5. In GSOP, $\vec{v}_i = \vec{w}_i - \text{proj}_{S_{i-1}}(\vec{w}_i)$, where $S_{i-1} = \text{Span}\{\vec{w}_1, \ldots, \vec{w}_{i-1}\} = \text{Span}\{\vec{v}_1, \ldots, \vec{v}_{i-1}\}$
6. $\text{perp}_{\mathbb{W}} : \mathbb{V} \to \mathbb{W}^\perp$ is linear and onto and $\text{Ker}(\text{perp}_{\mathbb{W}}) = \mathbb{W}$.
7. We can see immediately that $\text{proj}_{\mathbb{W}}(\vec{w})$ is onto since if $\vec{w} \in \mathbb{W}$, then $\vec{w} = \text{proj}_{\mathbb{W}}(\vec{w})$.
8. $\text{proj}_{\mathbb{W}}(\vec{v}), \text{perp}_{\mathbb{W}}(\vec{v})$ are unique.

Suppose $\mathbb{V}$ is an IPS and $\mathbb{W}_1, \mathbb{W}_2$ are subspaces of $\mathbb{V}$. Then if $\mathbb{W}_1 \subseteq \mathbb{W}_2$, then $\mathbb{W}_2^\perp \subseteq \mathbb{W}_1^\perp$.

9.5: Fundamental Theorem

Let $\mathbb{V}$ be a vector space, $\mathbb{U}, \mathbb{W}$ be subspaces of $\mathbb{V}$ such that $\mathbb{U} \cap \mathbb{W} = \{\vec{0}\}$. The direct sum of $\mathbb{U}, \mathbb{W}$ is $\mathbb{U} \oplus \mathbb{W} = \{\vec{u} + \vec{w} \in \mathbb{V} \mid \vec{u} \in \mathbb{U}, \vec{w} \in \mathbb{W}\}$. This is a subspace of $\mathbb{V}$.

Theorem 9.5.1, 2: Let $\mathbb{V}$ be a vector space, $\mathbb{U}, \mathbb{W}$ be subspaces of $\mathbb{V}$ such that $\mathbb{U} \cap \mathbb{W} = \{\vec{0}\}$. If $\{\vec{v}_1, \ldots, \vec{v}_k\}$ is a basis for $\mathbb{U}$ and $\{\vec{w}_1, \ldots, \vec{w}_l\}$ is a basis for $\mathbb{W}$, then $\{\vec{v}_1, \ldots, \vec{v}_k, \vec{w}_1, \ldots, \vec{w}_l\}$ is a basis for $\mathbb{U} \oplus \mathbb{W}$. Thus,

$\dim(\mathbb{U} \oplus \mathbb{W}) = \dim \mathbb{U} + \dim \mathbb{W}$. As well, for all $\vec{v} \in \mathbb{U} \oplus \mathbb{W}$, there are unique $\vec{u} \in \mathbb{U}$ and $\vec{w} \in \mathbb{W}$ such that $\vec{v} = \vec{u} + \vec{w}$.

Note that $\mathbb{U} \subseteq \mathbb{U} \oplus \mathbb{W}$ and same for $\mathbb{W}$.

Theorem 9.5.3: If $\mathbb{V}$ is a finite dimensional vector space and $\mathbb{W}$ is a subspace of $\mathbb{V}$, then $\mathbb{W} \oplus \mathbb{W}^{\perp} = \mathbb{V}$.

Fundamental Theorem of Linear Algebra: If $A \in M_{m \times n}$, then $\mathrm{Col}(A)^{\perp} = \mathrm{Null}(A^T)$ and $\mathrm{Row}(A)^{\perp} = \mathrm{Null}(A)$. As well, $\mathbb{R}^n = \mathrm{Row}(A) \oplus \mathrm{Null}(A)$ and $\mathbb{R}^m = \mathrm{Col}(A) \oplus \mathrm{Null}(A^T)$. Thus, every vector in $\mathbb{R}^n$ can be written as the sum of any vector in the rowspace of an $m \times n$ matrix and any vector in the nullspace. Similarly for $\mathbb{R}^m$. Note that orthogonality in this case is with the dot product.

In summary, $(\mathrm{Row}(A))^{\perp} = \mathrm{Null}(A), (\mathrm{Col}(A))^{\perp} = \mathrm{Null}(A^T)$.

9.6: The Method of Least Squares

Overdetermined system: A system of equations with more equations than unknowns.

Purpose: Let A be an $m \times n$ matrix with $m > n$ and let $A\vec{x} = \vec{b}$ be inconsistent. We want to find a $\vec{x}$ such that $||\vec{b} - A\vec{x}||$ is minimized. From the Appromation Theorem, we need $A\vec{x} = \mathrm{proj}_{\mathrm{Col}\,A}(\vec{b})$.

Approximation Theorem: Let $\mathbb{W}$ be a finite dimensional subspace of an inner product space $\mathbb{V}$. If $\vec{v} \in \mathbb{V}$, then the vector closest to $\vec{v}$ in $\mathbb{W}$ is $\mathrm{proj}_{\mathbb{W}}(\vec{v})$. Thus, for all $\vec{w} \in \mathbb{W}, w \neq \mathrm{proj}_{\mathbb{W}}(\vec{v})$

$$||\vec{v} - \mathrm{proj}_{\mathbb{W}}(\vec{v})|| < ||\vec{v} - \vec{w}||$$

Normal system: $A^T A\vec{x} = A^T\vec{b}$ is a normal system, the individual equations are the normal equations. This system is consistent by construction however there may be an infinite number of solutions. By solving this equation for $\vec{x}$, we can get the vector which minimizes $||b - A\vec{x}||$.

Method of Least Squares: From above, we can now minimize $||\vec{b} - A\vec{x}|| = \left|\left| \begin{bmatrix} v_1 \\ \vdots \\ v_m \end{bmatrix} \right|\right| = \sqrt{v_1^2 + \cdots + v_m^2}$

Theorem 9.6.2: Let m data points $(x_1, y_1), \ldots, (x_m, y_m)$ be given and write

$$\vec{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix}, X = \begin{bmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^n \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_m & x_m^2 & \cdots & x_m^n \end{bmatrix}$$

If $\vec{a} = \begin{bmatrix} a_0 \\ \vdots \\ a_n \end{bmatrix}$ is any solution to the normal system $X^T X\vec{a} = X^T\vec{y}$ then the polynomial $p(x) = a_0 + a_1 x + \cdots + a_n x^n$ is the best fitting polynomial of degree n for the given data. Moreover, if at least $n+1$ of $x_1, \ldots, x_m$ are distinct, then $X^T X$ is invertible and thus $\vec{a}$ is unique with $\vec{a} = (X^T A)^{-1} X^T \vec{y}$. This is because if $n+1$ x values are distinct, then the columns of X are linearly independent. Thus, $X^T X$ is invertible.

Note that if we were able to find a solution directly for $X\vec{a} = \vec{y}$, then $\vec{a}$ represents the coefficients of a perfectly fitting polynomial.

Chapter 10: Applications of Orthogonal Matrices

10.1: Orthogonal Similarity

A, B are orthogonally similar if there exists an orthogonal matrix P such that $P^T AP = B$. All the properties of similar matrices still hold for orthogonally similar matrices.

1. $\operatorname{rank} A = \operatorname{rank} B$
2. $\operatorname{tr} A = \operatorname{tr} B$
3. $\det A = \det B$ and A, B have the same eigenvalues

Triangularization Theorem: If A is an $n \times n$ matrix with only real eigenvalues, then A is orthogonally similar to an upper triangular matrix T. The eigenvalues of a T are the diagonal entries. Thus, A also has these eigenvalues. With these eigenvalues, we can then find the eigenvectors of A.

10.2: Orthogonal Diagonalization

A square matrix A is orthogonally diagonalizable if there exists an orthogonal matrix P and diagonal matrix D such that $P^T AP = D$. This means A is orthogonally similar to a diagonal matrix.

Theorem 10.2.1: If A is orthogonally diagonalizable, then A is symmetric: $A^T = A$.

Theorem 10.2.2: If A is symmetric with real entries, then all of its eigenvalues are real.

Principal Axis Theorem: Every symmetric matrix is orthogonally diagonalizable. Combined with 10.2.1, this means a matrix is orthogonally diagonalizable iff it is symmetric.

Theorem 10.2.4: A matrix A is symmetric iff $\vec{x} \cdot (A\vec{y}) = (A\vec{x}) \cdot \vec{y}$ for all $\vec{x}, \vec{y} \in \mathbb{R}^n$.

Theorem 10.2.5: If $\vec{v}_1, \vec{v}_2$ are eigenvectors of a symmetric matrix A corresponding to different eigenvalues λ_1, λ_2, then $\vec{v}_1$ is orthogonal to $\vec{v}_2$. Thus, if A has n distinct eigenvalues, the basis of eigenvectors which diagonalizes A will naturally be orthogonal. After normalizing these vectors, we can orthogonally diagonalize A right away. If A has eigenvalues of geometric multiplicity more than 1, than we apply Gram-Schmidt to find an orthonormal basis for this eigenspace. Then, combining all the vectors for different eigenspaces, we have an orthonormal basis.

10.3: Quadratic Forms

Let $A \in M_{n\times n}(\mathbb{R})$. A quadratic form is a function $Q : \mathbb{R}^n \to \mathbb{R}$ of the form:

$$Q(\vec{x}) = \vec{x}^T A\vec{x} = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j = \vec{x}^T A^T \vec{x}$$

Consider a symmetric matrix S defined as having the same values along the diagonal as A but $(S)_{ij} = (S)_{ji} = \frac{(A)_{ij}+(A)_{ji}}{2}$. Then $\vec{x}^T A\vec{x} = \vec{x}^T S\vec{x}$. Quadratic forms are closely tied to a corresponding symmetric matrix. A diagonal quadratic form has a diagonal corresponding symmetric matrix.

Quadratic form classification: Note that $Q(\vec{0}) = 0$. A quadratic form $Q(\vec{x})$ is:

1. Positive definite if $Q(\vec{x}) > 0$ for all $\vec{x} \neq \vec{0}$
2. Negative definite if $Q(\vec{x}) < 0$ for all $\vec{x} \neq \vec{0}$

3. Indefinite if $Q(\vec{x}) > 0$ for some vectors and $Q(\vec{x}) < 0$ for others.
4. Positive semidefinite if $Q(\vec{x}) \geq 0$ for all $\vec{x} \neq \vec{0}$. All positive definite is also positive semidefinite.
5. Negative semidefinite if $Q(\vec{x}) \leq 0$ for all $\vec{x} \neq \vec{0}$

The symmetric matrices corresponding to these quadratic forms are classified in the same way. For example, A is positive definite if the corresponding quadratic form to A is also positive definite.

Theorem 10.3.1: Let A be the symmetric matrix corresponding to a quadratic form $Q(\vec{x}) = \vec{x}^T A\vec{x}$. If P is an orthogonal matrix that diagonalizes A, then $Q(\vec{x}) = \lambda_1 y_1^2 + \cdots + \lambda_n y_n^2$, where $\vec{y} = P^T\vec{x}$ and $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of A corresponding to the columns of P. Notice that since P^T is invertible, if $\vec{x} \neq \vec{0}$, then $\vec{y} \neq \vec{0}$. Moreover, $||\vec{y}|| = ||P^T\vec{x}|| = ||\vec{x}||$. As well, since A is symmetric, then this P is guaranteed to exist. So every quadratic form is equivalent to a diagonal quadratic form using a change of variables $\vec{y} = P^T\vec{x}$.

As well, if B, C are both symmetric matrices such that $\vec{x}^T A\vec{x} = \vec{x}^T B\vec{x}$, then $B = C$.

Theorem 10.3.2: We now classify quadratic forms with eigenvalues. $Q(\vec{x}) = \vec{x}^T A\vec{x}$ is:

1. Positive definite iff all eigenvalues of A are positive.
2. Negative definite iff all eigenvalues of A are negative.
3. Indefinite iff some eigenvalues of A are positive and some are negative.
4. Positive semidefinite iff all eigenvalues are non-negative.
5. Negative semidefinite iff all eigenvalues are non-positive.

10.4: Graphing Quadratic Forms

Theorem 10.4.1: If $Q(\vec{x}) = ax_1^2 + bx_1x_2 + cx_2^2$ with a, b, c not all zero, then there exists an orthogonal matrix P which corresponds to a rotation in $\mathbb{R}^2$ such that the change of variables $\vec{y} = P^T\vec{x}$ brings $Q(\vec{x})$ into diagonal form.

Procedure for graphing quadratic forms:

1. Create a symmetric matrix A from the quadratic form
2. Find an orthogonal matrix P which diagonalizes A
3. Apply the change of variables $\vec{y} = P^T\vec{x}$ to change the quadratic form into a diagonal version.

10.5: Optimizing Quadratic Forms

Theorem 10.5.1: Let $Q(\vec{x})$ be a quadratic form on $\mathbb{R}^n$ with corresponding symmetric matrix A. Let $\lambda_1 \leq \cdots \leq \lambda_n$ be the eigenvalues of A, arranged in order. Let $\vec{v}_1, \ldots, \vec{v}_n$ be unit eigenvectors corresponding to the eigenvalues. Given a constraint $||\vec{x}|| = a$, then the min value of $Q(\vec{x})$ is $\lambda_1 a^2$ when $\vec{x} = a\vec{v}_1$ and the max is $\lambda_n a^2$ when $\vec{x} = a\vec{v}_n$.

10.6: Singular Value Decomposition

The purpose is to find the min/max of $||A\vec{x}||$ for $A \in M_{m\times n}(\mathbb{R})$ subject to $||\vec{x}|| = a$. Notice that maximizing $||A\vec{x}||$ is the same as maximizing $||A\vec{x}||^2 = \vec{x}^T A^T A\vec{x}$. This is a quadratic form. If $\lambda_1 \geq \cdots \geq \lambda_n$ are the eigenvalues of $A^T A$, then the max of $||A\vec{x}||$ is $a\sqrt{\lambda_1}$ and occurs when $\vec{x} = a\vec{v}_1$, where $\vec{v}_1$ is the corresponding eigenvector.

Theorem 10.6.1: If $A \in M_{m\times n}(\mathbb{R})$ and $\lambda_1, \dots, \lambda_n$ are eigenvalues of $A^T A$ with corresponding unit eigenvectors $\vec{v}_1, \dots, \vec{v}_n$, then $\lambda_1, \dots, \lambda_n$ are all non-negative. In particular, $||A\vec{v}_i||= \sqrt{\lambda_i}$.

The singular values $\sigma_1, \dots, \sigma_n$ of an $m \times n$ matrix A are the square roots of the eigenvalues of $A^T A$ arranged so that $\sigma_1 \geq \sigma_2, \geq \cdots \geq \sigma_n \geq 0$.

Theorem 10.6.2: If $A \in M_{m\times n}(\mathbb{R})$, then $\text{Null}(A^T A) = \text{Null}(A)$.

Theorem 10.6.3: If A is an $m \times n$ matrix, then $\text{rank}(A^T A) = \text{rank}(A)$.

Theorem 10.6.4: If $A \in M_{m\times n}(\mathbb{R})$ and $\text{rank}(A) = r$, then A has r non-zero singular values. This is because the number of non-zero eigenvalues of $A^T A$ is equal to $\text{rank}(A^T A)$ which is equal to $\text{rank}(A)$.

Singular vectors: Let $A \in M_{m\times n}(\mathbb{R})$. If $\vec{v} \in \mathbb{R}^n, \vec{u} \in \mathbb{R}^m$ are unit vectors and $\sigma \neq 0$ is a singular value of A such that $A\vec{v} = \sigma\vec{u}$ and $A^T\vec{u} = \sigma\vec{v}$, then $\vec{u}$ is a left singular vector of A corresponding to σ and $\vec{v}$ is a right singular vector of A corresponding to σ. However, if $\sigma = 0$, then if $\vec{u}$ is a unit vector and $A^T\vec{u} = 0$, then $\vec{u}$ is a left singular vector of A in $\text{Null}(A^T)$ corresponding to σ. Similarly, if $\vec{v}$ is a unit vector such that $A\vec{v} = \vec{0}$, then $\vec{v}$ is a right singular vector corresponding to σ.

Theorem 10.6.5: Let $A \in M_{m\times n}(\mathbb{R})$. If $\vec{v} \in \mathbb{R}^n$ is a unit eigenvector of $A^T A$ corresponding to a non-zero σ, then $\vec{u} = \frac{1}{\sigma}A\vec{v}$ is a left singular vector of A corresponding to σ.

Theorem 10.6.6: Let $A \in M_{m\times n}(\mathbb{R})$. Then $\vec{v} \in \mathbb{R}^n$ is a right singular vector of A iff $\vec{v}$ is a unit eigenvector of $A^T A$. And $\vec{u} \in \mathbb{R}^m$ is a left singular vector of A iff $\vec{u}$ is a unit eigenvector of AA^T.

Theorem 10.6.7: Let $A \in M_{m\times n}(\mathbb{R})$ with $\text{rank}(A) = r$ and singular values $\sigma_1 \geq \sigma_2 \ldots \geq \sigma_n$. If $\{\vec{v}_1, \dots, \vec{v}_n\}$ is an orthonormal basis for $\mathbb{R}^n$ consisting of the eigenvectors of $A^T A$ arranged corresponding to the singular values, then $\left\{\frac{1}{\sigma_1}A\vec{v}_1, \dots, \frac{1}{\sigma_r}A\vec{v}_r\right\}$ is an orthonormal basis for $\text{Col}(A)$. These are left singular values. Note that $A\vec{v}_{r+1}, \dots, A\vec{v}_n = 0$ since $\sigma_{r+1}, \dots, \sigma_n = 0$.

Theorem 10.6.8: If $A \in M_{m\times n}(\mathbb{R})$ with rank r and singular values $\sigma_1, \dots, \sigma_n$, then there exists an orthonormal basis $\{\vec{v}_1, \dots, \vec{v}_n\}$ for $\mathbb{R}^n$ of right singular vectors of A and an orthonormal basis $\{\vec{u}_1, \dots, \vec{u}_m\}$ for $\mathbb{R}^m$ of left singular vectors of A such that $A\vec{v}_i = \sigma_i\vec{u}_i, 1 \leq i \leq \min(m, n)$. Take $\{\vec{v}_1, \dots, \vec{v}_n\}$ as the eigenvectors of $A^T A$.

Let A be m by n. To find an orthonormal basis for $\mathbb{R}^n$ of right singular vectors of A, find an orthonormal basis of eigenvectors of $A^T A$.

To find an orthonormal basis of $\mathbb{R}^m$ of left singular vectors of A, apply Theorem 10.6.7. If we do not have enough vectors (since the basis must have m vectors but Theorem 10.6.7 only gives us r, or in cases where $m > n$), find an orthonormal basis for $\text{Null}(A^T)$. Since $\text{Col}(A) \oplus \text{Null}(A^T) = \mathbb{R}^m$, then by combining these two bases, we have an orthonormal basis of left singular vectors of A.

Singular value decomposition: Let $A \in M_{m\times n}(\mathbb{R})$ with rank r and non-zero singular values $\sigma_1, \dots, \sigma_r$. SVD is a factorization of the form $A = U\Sigma V^T$ where U is an orthogonal matrix containing left singular vectors of A, V is an orthogonal matrix containing right singular vectors of A, and Σ is the $m \times n$ matrix with $(\Sigma)_{ii} = \sigma_i$ for $1 \leq i \leq r$ and all other entries 0. Note that U is $m \times m$ and V is $n \times n$.

SVD Algorithm: Let $\text{rank}(A) = r$. Find the eigenvalues of $A^T A$, arranged from greatest to least and a corresponding set of orthonormal eigenvectors $\{\vec{v}_1, \dots, \vec{v}_n\}$. Define $\sigma_i = \sqrt{\lambda_i}$ for $1 \leq i \leq r$. Let $V = [\vec{v}_1 \cdots \vec{v}_n]$ and let Σ be $(\Sigma)_{ii} = \sigma_i$ for $1 \leq i \leq r$ and all other entries 0. Lastly, find left singular vectors of A by computing $\vec{u}_i = \frac{1}{\sigma_i}A\vec{v}_i$ for $1 \leq i \leq r$ and then extend this set to an orthonormal basis for $\mathbb{R}^m$ by taking vectors in $\text{Null}(A^T)$. Set up $U = [\vec{u}_1 \cdots \vec{u}_m]$. Then, $A = U\Sigma V^T$.

Fundamental Subspaces from SVD:

Let $A \in M_{m\times n}(\mathbb{R})$ with rank r and $V = [\vec{v}_1 \cdots \vec{v}_n], U = [\vec{u}_1 \cdots \vec{u}_m]$. Basis for the subspaces are:

1. $\text{Col}(A) : \{\vec{u}_1, \dots, \vec{u}_r\}$, from Theorem 10.6.7

2. $\text{Null}(A^T) : \{\vec{u}_{r+1}, \ldots, \vec{u}_m\}$, from basis of $\text{Col}(A)$ and Fundamental Theorem
3. $\text{Null}(A) : \{\vec{v}_{r+1}, \ldots, \vec{v}_n\}$, from $\sigma_{r+1}, \ldots, \sigma_n = 0$
4. $\text{Row}(A) : \{\vec{v}_1, \ldots, \vec{v}_r\}$, from basis of $\text{Null}(A)$ and Fundamental Theorem

Chapter 11: Complex Vector Spaces

11.1: Complex Number Review

$i^2 = -1$. All complex numbers are in the form $z = a + bi, a, b \in \mathbb{R}$ where $\text{Re } z = a, \text{Im } z = b$. $\mathbb{C}$ is a two dimensional real vector space. The modulus or absolute value of a complex number $z = a+bi$ is $|z|= \sqrt{a^2 + b^2}$.

Theorem 11.1.1: If $w, z \in \mathbb{C}$, then $|z|\geq 0, |z|= 0$ iff $z = 0, |w + z|\leq |w|+|z|$.

Theorem 11.1.2: If $z_1, z_2, z_3 \in \mathbb{C}$, then:

1. $z_1 + z_2 \in \mathbb{C}$
2. $z_1 + z_2 = z_2 + z_1$
3. $z_1 z_2 = z_2 z_1$
4. $z_1 + (z_2 + z_3) = (z_1 + z_2) + z_3$
5. $z_1(z_2 z_3) = (z_1 z_2)z_3$
6. $z_1(z_2 + z_3) = z_1 z_2 + z_1 z_3$
7. $z_1 + 0 = z_1$
8. $1z_1 = z_1$
9. For each $z \in \mathbb{C}$, there exists $(-z) \in \mathbb{C}$ such that $z + (-z) = 0$
10. For each $z \in \mathbb{C}$ with $z \neq 0$, there exists $(1/z) \in \mathbb{C}$ such that $z(1/z) = 1$

Complex conjugate $\overline{z}$ of $z = a + bi \in \mathbb{C}$ is: $\overline{z} = a - bi$ and $z\overline{z} = a^2 + b^2 = |a + bi|^2= |z|^2$

For complex numbers, its not necessarily true that $|z|^2= z^2$

Theorem 11.1.3: If $w, z \in \mathbb{C}$ with $z = a + bi$, then

1. $\overline{\overline{z}} = z$
2. z is real iff $\overline{z} = z$
3. z is imaginary iff $\overline{z} = -z, z \neq 0$
4. $\overline{z \pm w} = \overline{z} \pm \overline{w}$
5. $\overline{zw} = \overline{z}\,\overline{w}$

To divide by a complex number, multiply by its conjugate over its conjugate: $\frac{z}{w} = \frac{z}{w} \cdot \frac{\overline{w}}{\overline{w}} = \frac{z\overline{w}}{|w|^2}$.

11.2: Complex Vector Spaces

The set $\mathbb{C}$ can be thought of as a two dimensional real vector space but sometimes we want to use complex scalar coefficients. We need complex vector spaces.

A complex vector space $\mathbb{V}$ with operations of addition and scalar multiplication satisfies the typical 10 properties of vector spaces except it is now for any vector in $\mathbb{V}$ and any scalar in $\mathbb{C}$ rather than just in $\mathbb{R}$.

Theorem 11.2.1: If $\vec{z} \in \mathbb{C}^n$, there exists $\vec{x}, \vec{y} \in \mathbb{R}^n$ such that $\vec{z} = \vec{x}+i\vec{y}$. Similarly for $A \in M_{m\times n}(\mathbb{C}), A = B+iC$, for $B, C \in M_{m\times n}(\mathbb{R})$.

Subspace: $\mathbb{S}$ is a subset of a complex vector space $\mathbb{V}$ and $\mathbb{S}$ is a complex vector space under the same operations as $\mathbb{V}$ iff $\mathbb{S}$ is a subspace of $\mathbb{V}$. However, note that if the $\mathbb{V}$ is a complex vector space, then to check that $\mathbb{S}$

is closed under scalar multiplication, we must multiply with a complex number. For example, $\mathbb{R}$ is not a subspace of $\mathbb{C}$ as a complex vector space since $2i \notin \mathbb{R}$ so it is not closed.

A basis for $\mathbb{C}$ as a complex vector space is $\{1\}$ since $a + bi = (a + bi)(1)$. Also, note that $\{1, i\}$ is not linearly independent in $\mathbb{C}$ as a complex vector space.

All concepts of linear independence, spanning, bases, dimension are the same in complex vector spaces. As well, standard basis in $\mathbb{C}^n$ is same as in $\mathbb{R}^n$.

Complex conjugates: For $\vec{z} = \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix} \in \mathbb{C}^n$, then $\overline{\vec{z}} = \begin{bmatrix} \overline{z_1} \\ \vdots \\ \overline{z_n} \end{bmatrix}$. Simiarly, the conjugate of a matrix is where each element turns into its conjugate, or $(\overline{A})_{jk} = \overline{(A_{jk})}$. If $A \in M_{n\times n}(\mathbb{R}), A = \overline{A}$.

Theorem 11.2.2: If $A \in M_{m\times n}(\mathbb{C})$ and $\vec{z} \in \mathbb{C}^n$, then $\overline{A\vec{z}} = \overline{A}\,\overline{\vec{z}}$. Similarly, if $\alpha \in \mathbb{C}, \overline{\alpha\vec{z}} = \overline{\alpha}\overline{\vec{z}}$.

11.3: Complex Diagonalization

Complex eigenvalues and eigenvectors behave the same, except now $\lambda \in \mathbb{C}, \vec{z} \in \mathbb{C}^n, \vec{z} \neq \vec{0}$ such that $A\vec{z} = \lambda\vec{z}$. As well, $A \in M_{n\times n}(\mathbb{C})$. All the theory from Chapter 6 regarding similar matrices, eigenvalues, eigenvectors, diagonalization still apply to complex diagonalization.

However, notice that if a matrix has complex parts and is symmetric, its eigenvalues are not necessarily all real, even though for real matrices, this is a theorem (10.2.2).

Theorem 11.3.1: If $A \in M_{n\times n}(\mathbb{R})$ and λ is a non-real eigenvalue of A with corresponding eigenvector $\vec{z}$, then $\overline{\lambda}$ is also an eigenvalue of A with eigenvector $\overline{\vec{z}}$. Note that since A is a real matrix but λ is non-real, then we must have that $\vec{z} \in \mathbb{C}^n, \notin \mathbb{R}^n$. In other words, its not possible that $A\vec{z} = \lambda\vec{z}$ if $\vec{z}, A$ have only real components but λ is non-real.

Theorem 11.3.2: If $A \in M_{n\times n}(\mathbb{R})$ with n odd, then A has at least one real eigenvalue.

Important note: The trace of a matrix is equal to the sum of the eigenvalues.

If $\{\vec{v}_1, \ldots, \vec{v}_k\}$ is linearly independent in $\mathbb{R}^n$, then it is also linearly independent in $\mathbb{C}^n$. The standard basis for $\mathbb{R}^n$ is also the standard basis for $\mathbb{C}^n$. If a set is a basis of $\mathbb{R}^n$, then it is also a basis of $\mathbb{C}^n$ as a complex vector space. This is because every $\vec{z} \in \mathbb{C}^n$ can be written as $\vec{z} = \vec{a} + i\vec{b}, \vec{a}, \vec{b} \in \mathbb{R}^n$.

All $n \times n$ matrices have an n degree characteristic polynomial and n eigenvalues, some of which may be complex.

In general, for $\vec{z} \in \mathbb{C}^n, \{\vec{z}, \overline{\vec{z}}\}$ doesn't have to be linearly independent. Unless $\{\vec{z}, \overline{\vec{z}}\}$ are eigenvalues of a real matrix corresponding to non-real eigenvalues $\lambda, \overline{\lambda}$.

11.4: Complex Inner Products

Standard inner product on $\mathbb{C}^n$: $\langle \vec{z}, \vec{w} \rangle = \vec{z} \cdot \overline{\vec{w}} = z_1\overline{w_1} + \cdots + z_n\overline{w_n}$ for any $\vec{z}, \vec{w} \in \mathbb{C}^n$.

Notice that $\langle \vec{z}, \vec{z} \rangle = \vec{z} \cdot \overline{\vec{z}} = z_1\overline{z_1} + \cdots + z_n\overline{z_n} = |z_1|^2 + \cdots + |z_n|^2$.

Hermitian inner product on a complex vector space $\mathbb{V}$ is a function $\langle , \rangle : \mathbb{V} \times \mathbb{V} \to \mathbb{C}$ such that for all $\vec{v}, \vec{w}, \vec{z} \in \mathbb{V}$ and $\alpha \in \mathbb{C}$, then

1. $\langle \vec{z}, \vec{z} \rangle \in \mathbb{R}, \langle \vec{z}, \vec{z} \rangle \geq 0$, and $\langle \vec{z}, \vec{z} \rangle = 0$ iff $\vec{z} = \vec{0}$
2. $\langle \vec{z}, \vec{w} \rangle = \overline{\langle \vec{w}, \vec{z} \rangle}$

3. $\langle \alpha\vec{v}+\beta\vec{w}, \varphi\vec{y}+\gamma\vec{z}\rangle = \alpha\overline{\varphi}\langle\vec{v},\vec{y}\rangle + \alpha\overline{\gamma}\langle\vec{v},\vec{z}\rangle + \beta\overline{\varphi}\langle\vec{w},\vec{y}\rangle + \beta\overline{\gamma}\langle\vec{w},\vec{z}\rangle$
4. $\langle\vec{v}+\vec{z},\vec{w}\rangle = \langle\vec{v},\vec{w}\rangle + \langle\vec{z},\vec{w}\rangle$
5. $\langle\alpha\vec{z},\vec{w}\rangle = \alpha\langle\vec{z},\vec{w}\rangle$
6. $\langle\vec{z},\alpha\vec{w}\rangle = \overline{\alpha}\langle\vec{z},\vec{w}\rangle$
7. $\overline{\langle\alpha\vec{v},\vec{w}\rangle} = \overline{\alpha}\overline{\langle\vec{v},\vec{w}\rangle}$

Since all real numbers are also complex numbers, then the above properties hold for real inner products as well. All Hermitian inner products are inner products. To verify that something is a Hermitian inner product, very properties 1, 2, 4, 5 above.

Theorem 11.4.3: The standard Hermitian inner product on $M_{m\times n}(\mathbb{C})$ is $\langle A,B\rangle = \operatorname{tr}(\overline{B^T}A)$. To calculate this, take the standard inner product on $\mathbb{C}^n$ of each column of A and B, exactly the same as in real matrices.

Let $\mathbb{V}$ be a Hermitian inner product space with inner product $\langle,\rangle$. Then,

1. For any $\vec{v}\in\mathbb{V}$, the length of $\vec{z}$ is $||\vec{z}||=\sqrt{\langle\vec{z},\vec{z}\rangle}$. If the length is 1, then $\vec{z}$ is a unit vector.
2. For any $\vec{z},\vec{w}\in\mathbb{V}$, they are orthogonal if $\langle\vec{z},\vec{w}\rangle = 0$.
3. A set of vectors, each of which are in $\mathbb{V}$, is orthogonal if each pair is orthogonal. The set is orthonormal if all vectors are orthogonal to each other and all are unit vectors.
4. Let $\{\vec{v}_1,\dots,\vec{v}_n\}$ be an orthogonal basis for $\mathbb{V}$. Then if $\vec{z}\in\mathbb{V}$, $\vec{z} = \frac{\langle\vec{z},\vec{v}_1\rangle}{||\vec{v}_1||^2}\vec{v}_1 + \cdots + \frac{\langle\vec{z},\vec{v}_n\rangle}{||\vec{v}_n||^2}\vec{v}_n$
5. Gram-Schmidt procedure to find an orthogonal basis still applies in Hermitian inner product spaces.

Theorem 11.4.4: If $\mathbb{V}$ is a Hermitian inner product space, then for any $\vec{z},\vec{w}\in\mathbb{V}$ and $\alpha\in\mathbb{C}$, then

1. $||\alpha\vec{z}||=|\alpha|||\vec{z}||$
2. $||\vec{z}+\vec{w}||\leq||\vec{z}||+||\vec{w}||$
3. $\frac{1}{||\vec{z}||}\vec{z}$ is a unit vector in the direction of $\vec{z}$

Theorem 11.4.5: If $S=\{\vec{z},\dots,\vec{z}_k\}$ is an orthogonal set in a Hermitian inner product space $\mathbb{V}$, then $||\vec{z}_1+\cdots+\vec{z}_k||^2=||\vec{z}_1||^2+\cdots+||\vec{z}_k||^2$

Theorem 11.4.6: If S is an orthogonal set of non-zero vectors in a Hermitian inner product space, then it is linearly independent.

Gram-Schmidt is still valid for Hermitian inner product spaces.

Theorem 11.4.7: If $U\in M_{m\times n}(\mathbb{C})$, then the following are equivalent,

1. columns of U form an orthonormal basis for $\mathbb{C}^n$
2. $\overline{U^T}=U^{-1}$
3. rows of U form an orthonormal basis for $\mathbb{C}^n$

A unitary matrix satisfies the above properties. It's like the complex version of orthogonal matrix.

Theorem 11.4.8: If U,W are unitary matrices, then UW is a unitary matrix.

Conjugate transpose of a matrix A is A^* where $A^*=\overline{A}^T$

Theorem 11.4.9: If $A,B\in M_{m\times n}(\mathbb{C}),\vec{z}\in\mathbb{C}^n,\vec{w}\in\mathbb{C}^m,\alpha\in\mathbb{C}$, then

1. $\langle A\vec{z},\vec{w}\rangle = \langle\vec{z},A^*\vec{w}\rangle$
2. $(A^*)^*=A$ ‘

3. $(A+B)^* = A^* + B^*$
4. $(\alpha A)^* = \overline{\alpha} A^*$
5. $(AB)^* = B^* A^*$

11.5: Unitary Diagonalization

Unitarily similar: Let A, B be square matrices. If there is a unitary matrix U such that $U^*AU = B$, then A, B are unitarily similar. As well, if A, B are unitarily similar, then they are similar so all the properties of similar matrices still apply.

A matrix is unitarily diagonalizable if it is unitarily similar to a diagonal matrix.

Hermitian matrix: $A \in M_{n\times n}(\mathbb{C})$ such that $A^* = A$. Hermitian matrices have real elements in the diagonal. As well, all real symmetric matrices are Hermitian. Skew-Hermitian matrices satisfy $A^* = -A$

Schur's Theorem: If $A \in M_{n\times n}(\mathbb{C})$, then A is unitarily similar to an upper triangular matrix T. As well, the diagonal entries of T are the eigenvalues of A.

Spectral Theorem for Hermitian Matrices: If A is Hermitian, then A is unitarily diagonalizable. Prove this using Schur's Theorem. However, the converse is not necessarily true.

Theorem 11.5.3: $A \in M_{n\times n}(\mathbb{C})$ is Hermitian iff for all $\vec{z}, \vec{w} \in \mathbb{C}^n$, we have $\langle A\vec{z}, \vec{w}\rangle = \langle \vec{z}, A\vec{w}\rangle$.

Theorem 11.5.4: If A is Hermitian, then

1. All eigenvalues of A are real. Thus, when unitarily diagonalizing a Hermitian matrix, the diagonal matrix we get must be real.
2. If λ_1, λ_2 are distinct eigenvalues with corresponding eigenvectors $\vec{v}_1, \vec{v}_2$, then $\vec{v}_1, \vec{v}_2$ are orthogonal.

Normal: $A \in M_{n\times n}(\mathbb{C})$ is normal if $AA^* = A^*A$.

Spectral Theorem for Normal Matrices: A is normal iff it is unitarily diagonalizable.

Every upper triangular normal matrix is diagonal.

Theorem 11.5.6: If A is normal,

1. $||A\vec{z}|| = ||A^*\vec{z}||$ for all $\vec{z} \in \mathbb{C}^n$
2. $A - cI$ is normal for every $c \in \mathbb{C}$
3. If $A\vec{z} = \lambda\vec{z}$, then $A^*\vec{z} = \overline{\lambda}\vec{z}$
4. If $\vec{z}_1, \vec{z}_2$ are eigenvectors of A corresponding to distinct eigenvalues λ_1, λ_2 of A, then $\vec{z}_1, \vec{z}_2$ are orthogonal.

11.6: Cayley-Hamilton Theorem

Cayley-Hamilton Theorem: If $A \in M_{n\times n}(\mathbb{C})$, then A is a root of its characteristic polynomial $C(\lambda)$. In this instance, we use $C(X) = (-1)^n X^n + c_{n-1}X^{n-1} + \cdots + c_1 X + c_0 I = (X - \lambda_1 I)\cdots(X - \lambda_n I)$.

Theorem 11.6.2: If $A \in M_{m\times n}(\mathbb{C})$ is invertible, then A^{-1} can be written as a linear combination of powers of A.

Extra things

To prove that a mapping $L : \mathbb{V} \to \mathbb{W}$ is not an isomorphism, show one of these:

1. L is not linear, one-to-one, or onto
2. $\mathrm{Ker}(L) \neq \{\vec{0}\}$, or that $\mathrm{nullity}(L) \neq 0$
3. If $\dim \mathbb{V} \neq \dim \mathbb{W}$, then $\mathbb{V}$ is not isomorphic to $\mathbb{W}$ so L is not an isomorphism
4. If a basis for $\mathbb{V}$ is $\{\vec{v}_1, \ldots, \vec{v}_n\}$, then L is an isomorphism iff $\{L(\vec{v}_1), \ldots, L(\vec{v}_n)\}$ is a basis for $\mathbb{W}$.

On page 79: If $\vec{x}, \vec{y} \in \mathbb{R}^n$, then $\vec{x} \cdot \vec{y} = \vec{x}^T \vec{y}$.

The eigenvalues of A are the same as $A^T : \det(A - \lambda I) = \det((A - \lambda I)^T) = \det(A^T - \lambda I)$.

Characteristic polynomial of similar matrices are the same: Let $B = P^{-1}AP$:

$$\begin{aligned}
\det(P^{-1}AP - \lambda I) &= \det(P^{-1}AP - \lambda P^{-1}IP) = \det(P^{-1}(A - \lambda I)P) \\
&= \det(P^{-1}) \det(A - \lambda I) \det(P) = \det(A - \lambda I)
\end{aligned}$$

Proof that the number of non-zero eigenvalues of $A^T A$ is the rank of $A^T A$ (page 311): Since $A^T A$ is symmetric, then there is some orthogonal matrix P such that $P^T A^T AP = D$, where D is a diagonal matrix of the eigenvalues. Clearly, the rank of D is the number of non-zero eigenvalues and by Theorem 6.1.1, $\mathrm{rank}(A^T A) = \mathrm{rank}(D)$.

Proof of Theorem 10.6.6: Let $A \in M_{m\times n}(\mathbb{R})$. Then $\vec{v} \in \mathbb{R}^n$ is a right singular vector of A iff $\vec{v}$ is a unit eigenvector of $A^T A$. And $\vec{u} \in \mathbb{R}^m$ is a left singular vector of A iff $\vec{u}$ is a unit eigenvector of AA^T.

Suppose $\vec{v} \in \mathbb{R}^n$ is a right singular vector of A. Let σ be a singular value of A. Then,

$$\begin{aligned}
A^T \vec{u} &= \sigma \vec{v} \\
A\vec{v} &= \sigma \vec{u} \\
A^T A\vec{v} &= \sigma A^T \vec{u} = \sigma(\sigma \vec{v}) = \lambda \vec{v}
\end{aligned}$$

Suppose $\vec{v}$ is a unit eigenvector of $A^T A$ corresponding to eigenvalue $\lambda = \sigma^2$. Then $\vec{u} = \frac{1}{\sigma} A\vec{v}$ is a left singular vector of A by Theorem 10.6.5. Then, $A\vec{v} = \sigma \vec{u}$ and,

$$\begin{aligned}
A^T A\vec{v} &= \lambda \vec{v} = \sigma^2 \vec{v} \\
A^T(\sigma \vec{u}) &= \sigma^2 \vec{v} \\
A^T \vec{u} &= \sigma \vec{v}
\end{aligned}$$

Therefore, $\vec{v}$ is a right singular vector of A.

If A is symmetric and invertible, then A^{-1} is also symmetric since $(A^{-1})^T = (A^T)^{-1} = A^{-1}$

Suppose $\mathbb{V}$ is an IPS and $\mathbb{W}_1, \mathbb{W}_2$ are subspaces of $\mathbb{V}$. Prove that if $\mathbb{W}_1 \subseteq \mathbb{W}_2$, then $\mathbb{W}_2^\perp \subseteq \mathbb{W}_1^\perp$. Similarly, if $\mathbb{W}_1^\perp \subseteq \mathbb{W}_2^\perp$, then $\mathbb{W}_2 \subseteq \mathbb{W}_1$.

Since A is normal, it is diagonalizable.

1.