Basic Linear Algebra

Yunhai Xiang

August 9, 2019

Contents

1 Vector Spaces and Linear Transformations

2 Linear Independance and Span

3 Null Spaces, Column Spaces, and Rank-Nullity Theorem

4 Systems of Equations and RREF

5 Trace and Determinants

6 Eigenvectors and Eigenvalues

Theorem 6.1 (Gershgorin's Circle Theorem)**.**

7 Invariant Subspaces and Minimal Polynomials

Definition 7.1. Let V be a finite dimensional vector space, W be a subspace of V, and T be a linear operation. We say that W is T**-invariant** if $T(W) \subseteq W$. If W is T-invariant, we use the notation $T_W : W \to W$ to represent the linear operation defined as $T_W(x) = T(x)$ for all $x \in W$.

Proposition 7.2. The eigenspaces of T are T-invariant.

Definition 7.3. Let V be a finite dimensional vector space, T be a linear operation, and $x \in V$ a nonzero vector. Then the T**-cyclic vector space generated by** x is

$$\mathcal{Z}(x, T) = \operatorname{span}\{T^n(x) \mid n \in \mathbf{N}\}$$

Proposition 7.4. Let V be a vector space over $\mathbf{F}$ with dimension n, let T be a linear operation, and let $x \in V$ be nonzero.

1. The subspace $\mathcal{Z}(x, T)$ is the smallest T-invariant subspace of V containing x.
2. The set $\{x, T(x), \dots, T^{n-1}(x)\}$ is a basis for $\mathcal{Z}(x, T)$
3. If $f \in \mathbf{F}[x]$ is monic with $\deg f = n$ and $f(T)(x) = 0$ then the characteristic polynomial of $T_{\mathcal{Z}(x,T)}$ is $(-1)^n f$
4. Let V be a finite dimensional vector space over $\mathbf{F}$, T be a linear operation, and W be a T-invariant subspace of V, then the characteristic polynomial of T_W divides that of T.

Proof. We will only prove 2,3,4.

2. Let $\beta = (v_1, \dots, v_m)$ be an ordered basis for W and we extend it to a basis $\gamma = (v_1, \dots, v_m, v_{m+1}, \dots, v_n)$ for V. Let $A = [T]_\beta$ and $B = [T_W]_\beta$, then

$$B = \begin{pmatrix} A & * \\ O & C \end{pmatrix}_{\text{block}}$$

for some $C \in \mathcal{M}_{n-m}(\mathbf{F})$. Therefore

$$\det(B - \lambda I) = \det(A - \lambda I)\det(C - \lambda I)$$

Hence the characteristic polynomial of T_W divides that of T.

□

Theorem 7.5 (Cayley-Hamilton)**.** Let V be a finite dimensional vector space over $\mathbf{F}$ and T a linear operator with characteristic polynomial $f \in \mathbf{F}[x]$, then $f(T) = 0$.

Proof. Since $f(T)(0) = 0$, let $v \in V$ nonzero we claim that $f(T)(v) = 0$, let $W = \mathcal{Z}(T, v)$ with dimension k, and let $a_0, \ldots, a_k$ not all zero such that $a_k T_k(v) + \cdots + a_0 v = 0$ where we assume wlog that $a_k = 1$, then we have $g(T)(v) = 0$ where $g(x) = x^k + a_{k-1}x^{k-1} + a_0$, so the characteristic polynomial of T_W is $h(x) = (-1)^k g(x)$, since $h(T)(v) = 0$ and $h \mid f$ we have $f(T)(v) = 0$ and thus $f(T) = 0$ □

Definition 7.6. Let V be a finite dimensionnal vector space over $\mathbf{F}$ and T a linear operator, then the **minimal polynomial** of T is the monic polynomial that generates the principle ideal $\{f \in \mathbf{F}[x] \mid f(T) = 0\}$ of $\mathbf{F}[x]$,

Proposition 7.7 (Properties of Minimal Polynomials and Characteristic Polynomials)**.** Let V be a finite dimensionnal vector space over $\mathbf{F}$ and T a linear operator with minimal polynomial $m \in \mathbf{F}[x]$,

1. m divides any $f \in \mathbf{F}[x]$ with $f(T) = 0$,
2. m divides the characteristic polynomial of T,
3. m has the same roots (ignoring multiplicity) as the characteristic polynomial of T,
4. if $V = \mathcal{Z}(T, v)$ for some nonzero $v \in V$ then the characteristic polynomial is $(-1)^{\dim V} m$,

Theorem 7.8. Let V be a finite dimensionnal vector space over $\mathbf{F}$ and T a linear operator with minimal polynomial $m \in \mathbf{F}[x]$, then T is diagonalizable iff $m(x) = \prod_{k=1}^{n}(x - \lambda_k)$ where $\lambda_1, \ldots, \lambda_n$ are all distinct eigenvalues of T.

8 Jordan Canonical Form

Definition 8.1. Let $n \in \mathbf{Z}^+$ and $\mathbf{F}$ a field, a **Jordan block** is a matrix $A \in \mathcal{M}_n(\mathbf{F})$ such that

$$A = \begin{pmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda & 1 \\ & & & & \lambda \end{pmatrix}$$

where $\lambda \in \mathbf{F}$. We say that $J \in \mathcal{M}_N(\mathbf{F})$ is a **Jordan matrix** if there exists $n_1, \ldots, n_k \in \mathbf{Z}^+$ such that $N = n_1 + \cdots + n_k$ and Jordan blocks $A_1 \in \mathcal{M}_{n_1}(\mathbf{F}), \ldots, A_k \in \mathcal{M}_{n_k}(\mathbf{F})$ such that

$$J = \begin{pmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_k \end{pmatrix}_{\text{block}}$$

Definition 8.2. Let V be a finite dimensional vector space with some eigenvalue λ and T be a linear operator. A **generalilzed eigenvector** of the eigenvalue λ is a vector $v \in V$ such that $(T-\lambda I)^m(v) = 0$ for some $m \in \mathbf{Z}^+$. And we define $\mathcal{K}_\lambda = \{v \in V \mid \exists m \in \mathbf{Z}^+, (T-\lambda I)^m(v) = 0\}$ to be the **generalized λ-eigenspace** of V.

Proposition 8.3 (Properties of Generlized Eigenspace)**.** Let V be a finite dimensional vector space with some eigenvalue λ

1. $\mathcal{K}_\lambda = \bigcup_{n=1}^{\infty} \ker(T - \lambda I)^n$,
2. $\mathcal{K}_\lambda$ is T-invariant,
3. if λ, μ are distinct eigenvalues of T then $(T - \lambda I) : \mathcal{K}_\mu \to \mathcal{K}_\mu$ is a bijection,
4. if λ, μ are distinct eigenvalues then $\mathcal{K}_\lambda \cap \mathcal{K}_\mu = \{0\}$.
5. let m be the algebraic multiplicity of λ then $\dim \mathcal{K}_\lambda \leq m$
6. let m be the algebraic multiplicity of λ then $\mathcal{K}_\lambda = \ker(T - \lambda I)^m$
7. assume the characteristic polynomial of T splits with distinct eigenvalues $\lambda_1, \ldots, \lambda_k$ and respective algebraic multiplicity $m_1, \ldots, m_k$, and let β_i be an ordered basis for $\mathcal{K}_{\lambda_i}$ for $1 \leq i \leq k$, then $\beta_1, \ldots, \beta_k$ are pairwise disjoint and $\beta = \bigcup_{i=1}^{k} \beta_i$ is a basis for V. Moreover, $\dim \mathcal{K}_{\lambda_i} = m_i$.

Theorem 8.4. Let V be a finite dimensional vector space and T a linear operator. If $\lambda_1, \ldots, \lambda_k$ are all distinct eigenvalues of T then for $x \in V$ there exists $v_i \in \mathcal{K}_{\lambda_i}$ for $1 \leq i \leq k$ such that $x = v_1 + \cdots + v_k$.

Algorithm 8.5 (Finding the Jordan Canonical Form)**.**

9 Inner Product Space and Orthogonal Projection

Definition 9.1. An **inner product space** is a vector space V over $\mathbf{C}$ with a map $\langle \cdot, \cdot \rangle : V \times V \to \mathbf{C}$ called the **inner product** such that for $x, y, z \in V$ and $a \in \mathbf{C}$,

1. $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$,
2. $\langle ax, y \rangle = a\langle x, y \rangle$,
3. $\overline{\langle x, y \rangle} = \langle y, x \rangle$, and
4. $\langle x, x \rangle \in \mathbf{R}_{\geq 0}$ and $\langle x, x \rangle = 0$ iff $x = 0$.

Moreover, if $V = \mathbf{C}^n$ for some $n \in \mathbf{Z}^+$ then the **standard inner product** is defined as

$$\langle x, y \rangle = a_1\overline{b_1} + a_2\overline{b_2} + \cdots + a_n\overline{b_n}$$

where $x = (a_1, \ldots, a_n)$ and $y = (b_1, \ldots, b_n)$.

Example 9.2 (Frobenius Inner Product)**.**

Definition 9.3. A **normed vector space** is a vector space V over $\mathbf{C}$ with a map $\|\cdot\| : V \to \mathbf{R}$ called the **norm** or the **metric** such that for $x, y \in V$ and $a \in \mathbf{C}$,

1. $\|x\| \geq 0$ and $\|x\| = 0$ iff $x = 0$,
2. $\|ax\| = |a|\|x\|$, and,
3. $\|x + y\| \leq \|x\| + \|y\|$.

Moreover, if V is an inner product space, then we define the **standard norm** as $\|x\| = \sqrt{|\langle x, x \rangle|}$. In particular, if $V = \mathbf{C}^n$ for some $n \in \mathbf{Z}^+$ then the standard norm is

$$\|x\| = \sqrt{|a_1|^2 + |a_2|^2 + \cdots + |a_n|^2}$$

where $x = (a_1, \ldots, a_n)$.

Definition 9.4. In an inner product space V, we say $x, y \in V$ are **orthogonal** if $\langle x, y \rangle = 0$. We say that a subset $S \subseteq V$ is orthogonal if for $x, y \in V$ with $x \neq y$, then x, y are orthogonal. Moreover, if V is also a normed vector space, we say that $S \subseteq V$ is **orthonormal** if S is orthogonal and for $x \in V$, we have $\|x\| = 1$.

Proposition 9.5 (Facts about orthogonality and inner products)**.** Let V be an inner product space with inner product $\langle \cdot, \cdot \rangle$ and standard norm $\| \cdot \|$, then

1. if $x, y, z \in V$, then $\langle x, y + z \rangle = \langle x, y \rangle + \langle x, z \rangle$,
2. if $x \in V$ and $a \in \mathbf{C}$, then $\langle x, ay \rangle = \overline{a}\langle x, y \rangle$

3. if $x \in V$, then $\langle x, 0\rangle = \langle 0, x\rangle = 0$,
4. if $x, y \in V$ with $y \neq 0$, then $\langle x, y\rangle = 0$ iff $x = 0$,
5. if $x, y \in V$ are orthogonal then $\|x + y\| = \|x\| + \|y\|$,
6. if $x, y \in V$ then $\|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2)$
7. if $S \subseteq V$ be an orthogonal set of nonzero vectors, then S is linearly independent.

Proof. We will only prove 7. Let $s_1, \ldots, s_n \in S$ and $a_1 \ldots, a_n \in \mathbf{C}$ and let $a_1 s_1 + \cdots + a_n s_n = 0$ for $1 \leq i \leq n$, since

$$0 = \langle 0, s_i\rangle = \langle a_1 s_1 + \cdots + a_n s_n, s_i\rangle = a_1\langle s_1, s_i\rangle + \cdots + a_n\langle s_n, s_i\rangle = a_i\langle s_i, s_i\rangle$$

and since $\langle s_i, s_i\rangle > 0$ we have $a_i = 0$ for all $1 \leq i \leq n$. □

Theorem 9.6 (Cauchy-Schwarz inequality)**.** Let V be a inner product space then for $x, y \in V$

$$|\langle x, y\rangle|^2 \leq \langle x, x\rangle\langle y, y\rangle$$

wich equality iff x, y are linearly dependant.

Proof. If $y = 0$ then trivial. Assume $y \neq 0$, and let $a = \frac{\langle x,y\rangle}{\langle y,y\rangle}$

$$\begin{aligned}
\langle x - ay, x - ay\rangle &= \langle x, x\rangle - \overline{a}\langle x, y\rangle - a\langle y, x\rangle + a\overline{a}\langle y, y\rangle \\
&= \langle x, x\rangle - \overline{\frac{\langle x, y\rangle}{\langle y, y\rangle}}\langle x, y\rangle - \frac{\langle x, y\rangle}{\langle y, y\rangle}\langle y, x\rangle + \frac{\langle x, y\rangle}{\langle y, y\rangle}\overline{\frac{\langle x, y\rangle}{\langle y, y\rangle}}\langle y, y\rangle \\
&= \langle x, x\rangle - \frac{\langle x, y\rangle\langle y, x\rangle}{\langle y, y\rangle}
\end{aligned}$$

and since $\langle x - ay, x - ay\rangle \geq 0$, with equality iff $x = ay$, we have

$$\begin{aligned}
\frac{\langle x, y\rangle\langle y, x\rangle}{\langle y, y\rangle} &\leq \langle x, x\rangle \\
\langle x, y\rangle\langle y, x\rangle &\leq \langle x, x\rangle\langle y, y\rangle \\
\langle x, y\rangle\overline{\langle x, y\rangle} &\leq \langle x, x\rangle\langle y, y\rangle \\
|\langle x, y\rangle|^2 &\leq \langle x, x\rangle\langle y, y\rangle
\end{aligned}$$

with equality iff $x = ay$. □

Theorem 9.7 (Gram-Schmit)**.** Let V be a finite dimensional inner product space and $\{v_1, \ldots, v_n\} \subseteq V$ be linearly independent. Define $e_1 = v_1$ and

$$e_k = v_k - \sum_{i=1}^{k-1} \frac{\langle v_k, e_i\rangle}{\langle e_i, e_i\rangle} e_i$$

for $2 \leq k \leq n$, then $\{e_1, \ldots, e_n\}$ is orthogonal and $\text{span}\{e_1, \ldots, e_n\} = \text{span}\{v_1, \ldots, v_n\}$.

Proof. □

Proposition 9.8. Let V be a finite dimensional inner product space and $\{v_1, \ldots, v_n\} \subseteq V$ be linearly independent. Let

$$A = \begin{pmatrix} \text{———} & v_1 & \text{———} \\ \text{———} & v_2 & \text{———} \\ & \vdots & \\ \text{———} & v_n & \text{———} \end{pmatrix}$$

Then the Row-Echelon form of the augmented matrix $[AA^* \mid A]$ will produce Gram-Schmit output in place of rows of A.

Proposition 9.9. Let V be a finite dimensional vector space and let $\{v_1, \ldots, v_n\}$ be linearly independent. Run Gram-Schmit on $\{v_1, \ldots, v_n\}$ to get $\{e_1, \ldots, e_n\}$. Then the change of basis matrix from $(v_1, \ldots, v_n)$ to $(e_1, \ldots, e_n)$ is upper triangular.

Definition 9.10. Let V be an inner product space and S a nonempty subset then the **orthogonal complement** of S is defined as $S^\perp = \{v \in V \mid \forall x \in S, \langle x, v\rangle = 0\}$.

Proposition 9.11 (Properties of orthogonal complements)**.** Let V be an inner product space and S a nonempty subset, then

1. $V = S \oplus S^\perp$,

Definition 9.12. Let V be an finite dimensional inner product space and W a subspace with orthonormal basis $\{e_1, \ldots, e_n\}$, then the **orthogonal projection** of a vector $v \in V$ onto W is defined as

$$\text{proj}_W(v) = \sum_{i=1}^{n} \langle v, e_i\rangle e_i$$

Note that the projection is independent of the choice of basis.

10 The Adjoint and Least Square Approximation

Definition 10.1. A **linear functional** on a vector space V over $\mathbf{F}$ is a linear operator $T : V \to \mathbf{F}$. The **dual space** of V, denoted V^*, is the set of all linear functionals on V.

Theorem 10.2 (Riesz's Representation Theorem)**.** Let T be a linear functional on the finite dimensional vector space V, then for all $x \in V$, there exists $y \in V$ such that $T(x) = \langle x, y\rangle$.

Definition 10.3. Let V be a finite dimensional inner product space and T a linear operator, the **adjoint operator** of T is a linear operator T^* such that $\langle T(x), y\rangle = \langle x, T^*(y)\rangle$ for all $x, y \in V$.

We can prove that T^* always exists and is unique.

Definition 10.4. Let $A \in \mathcal{M}_n(\mathbf{C})$, then the **adjoint matrix** or the **conjugate transpose** of A, is defined as $A^* = (\overline{a_{j,i}})_{i,j=1}^n$ where $A = (a_{i,j})_{i,j=1}^n$.

Note that if β is an orthonormal basis then $[T^*]_\beta = [T]_\beta^*$

Proposition 10.5 (Properties of Adjoint)**.** Let V be a finite dimensional vector space over $\mathbf{F}$ and T, U two linear operators, then for $a \in \mathbf{F}$

1. $(T + U)^* = T^* + U^*$,
2. $(aT)^* = \overline{a}T^*$,
3. $(UT)^* = T^*U^*$,
4. $(T^*)^* = T$
5. $I^* = I$

Let $(x_1, y_1), \ldots, (x_m, y_m)$ be a set of points, we wish to find the polynomial $f(x) = a_n x^n + \cdots + a_0$ such that $\sum_{i=1}^m (f(x_i) - y_i)^2$ is minimal

Theorem 10.6 (Least Square Approximation)**.** Let $(x_1, y_1), \ldots, (x_n, y_n)$ be a set of points, let $y = (y_1, y_2, \ldots, y_m)$ and let the matrix

$$A = \begin{pmatrix} x_1^n & x_1^{n-1} & \cdots & x_1 & 1 \\ x_2^n & x_2^{n-1} & \cdots & x_2 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_m^n & x_m^{n-1} & \cdots & x_m & 1 \end{pmatrix}$$

then the vector $x = x_0$ where $x_0 = (A^*A)^{-1}A^*y$ minimizes $\|Ax - y\|$.

11 Bilinear Form

12 Normal Operators and Schur's Theorem

Definition 12.1. For $n \in \mathbf{Z}^+$ and a field $\mathbf{F}$, if $A \in \mathcal{M}_n(\mathbf{F})$ is such that $A^*A = AA^*$, then A is said to be **normal**. If $A^*A = AA^* = I$ then A is **orthogonal**.

Lemma 12.2. Let V be a finite dimensional inner product space and T a linear operator, then if T has an eigenvector, T^* has an eigenvector.

Theorem 12.3. Let V be a finite dimensional inner product space and T a linear operator, if the characteristic polynomial of T splits, then there exists an orthonormal basis β such that $[T]_\beta$ is upper triangular.

Corollary 12.4. For $n \in \mathbf{Z}^+$, a field $\mathbf{F}$, if $A \in \mathcal{M}_n(\mathbf{F})$ is such that its characteristic polynomial splits, then there exists orthogonal matrix U and upper triangular matrix B such that $A = UBU^*$.

Proposition 12.5. For a finite dimensionall inner product space V and normal linear operator T,

1. for $x \in V$, $\|T(x)\| = \|T^*(x)\|$,
2. every λ-eigenvector of T is a $\overline{\lambda}$-eigenvector of T^*,
3. if x is a λ-eigenvector of T and y is a μ-eigenvector of T with $\lambda \neq \mu$ then $\langle x, y \rangle = 0$

Theorem 12.6. For a finite dimensionall inner product space V and a linear operator T, T is normal iff there exists an orthonormal basis consisting of eigenvectors of T.

13 Hermitian Operators, Unitary Operators, and Rigid Motion

Definition 13.1. Let V be an inner product space, then a linear operator T is called **Hermitian** if $T^* = T$

Proposition 13.2. Let V be an finite dimensional inner product space with a Hermitian linear operator T, then

1. Every eigenvalue of T is real,
2. The characteristic polynomial of T splits.

Theorem 13.3. Let V be a finite dimensional inner product over $\mathbf{R}$ with linear operator T. Then T is Hermitian iff there exists an orthonormal basis consisting of eigenvectors of T.

Corollary 13.4. For $n \in \mathbf{Z}^+$, $A \in \mathcal{M}_n(\mathbf{R})$ is Hermitian iff there exists symmetric matrix U and diagonal matrix D such that $A = UDU^T$.

Proposition 13.5. Let V be a finite dimensional inner product with linear operator T. The following are equivalent,

1. T is orthogonal
2. for all $x, y \in V$, $\langle x, y \rangle = \langle T(x), T(y) \rangle$
3. if β is an orthonormal basis for V then $T(\beta)$ is also an orthonormal basis for V.
4. for all $x \in V$, $\|T(x)\| = \|x\|$

Corollary 13.6. Every eigenvalue of an orthogonal linear operator has absolute value 1.

Definition 13.7. Let V be a finite dimensional real inner product space, then an operator f is a **rigid motion** if it satisfies $\|f(x) - f(y)\| = \|x - y\|$ for all $x, y \in V$.

Definition 13.8. A **translation** on a vector space V is an operator $f : V \to V$ defined as $f(x) = x + v$ for some $v \in V$.

Theorem 13.9. Let V be a finite dimensional real inner product space. If f is a rigid motion then there exists unique orthogonal operator T and a unique translation g such that $f = g \cdot T$.

14 Spectral Theorem

Definition 14.1. Let V be a vector space with subspaces W_1, W_2 such that $V = W_1 \oplus W_2$. If the linear operator T is such that for $v \in V$, $T(v) = x$ where $v = x + y$ where $x \in W_1$ and $y \in W_2$ (which is unique) then T is a **projection** on W_1 along W_2.

Let T be a projection on W_1 along W_2 then $\operatorname{im} T = W_1$ and $\ker T = W_2$

Proposition 14.2. The linear operator is a projection iff it is idempotent.

Theorem 14.3 (Spectral Theorem)**.**

15 Polar Decomposition and Singular Value Decomposition

16 Tensor Product

17 Tensor Algebra and Exterior Algebra

18 Introduction to Functional Analysis

Definition 18.1. A **Banach space** is a normed vector space that is complete. If, further, that the norm is $\|\cdot\| = \sqrt{\langle \cdot, \cdot \rangle}$ for some inner product $\langle \cdot, \cdot \rangle$ then the vector space is called a **Hilbert space**.