UCLA Basic: Algebra Notes

Completed on September 13, 2021

Based on Notes/Lectures from Ben

Anonymous

September 13, 2021

Contents

Basic Concepts, Nothing Special

Basis
Let V be a vector space. A set of vectors $\mathcal{B}$ is a basis for V if any $v \in V$ can be written as a linear combination of vectors from $\mathcal{B}$, in a unique way.

Theorem. Every vector space has a basis.

Proof. The proof of this is not simple or fun and should not appear on the basic.

Note : Unless otherwise stated, we will assume that any vector space has a finite basis.

Theorem. (Steinitz Replacement) : Let $\{y_1, \dots, y_m\} \subset V$ be linearly independent. Let $V = \text{span}\{x_1, \dots, x_n\}$. Then $m \leq n$ and V has a basis of the form $y_1, \dots, y_n, x_{i_1}, \dots, x_{i_\ell}$ for some $\ell \leq n - m$.

Span
The span of a subset $S \subset V$, denote $\text{span}(S)$, is the smallest subspace of V containing S.

Commutator
Let $T, S : V \to V$ be linear operators. The commutator $[T, S]$ is the operator

$$[T, S] = TS - ST$$

We say that the operators T, S *commute* if $[T, S] = 0$.

Affine Subspace
A set $A \subset V$ is said the be an affine subspace if there exists $w \in A$ such that

$$A - w = \{v - w : v \in A\}$$

is a subspace of V. The dimension of A is defined to be the dimension of $A - w$.

Rank

Rank and Nullity
Let $T : V \to W$ be a linear transformation between vector spaces. Then

$$\text{rank}(T) = \dim(\text{im}(T)) \qquad \text{nullity}(T) = \dim(\ker(T))$$

Theorem. (Rank Nullity Theorem) : Let $T : V \to W$ be a linear transformation between vector spaces. Then

$$\text{rank}(T) + \text{nullity}(T) = \dim(V)$$

Proof. There are two main proofs of the rank nullity theorem, both of which I will present.

(1) By definition, $\ker T$ is a subspace of V. Therefore, it has a basis $w_1, \ldots, w_m$ for some $m \leq \dim(V)$. By the Steinitz exchange theorem, we may extend $w_1, \ldots, w_m$ to a full basis by adding vectors $v_{m+1}, \ldots, v_n$ where $n = \dim(V)$.

We claim the $T(v_i)$ is a basis of $\text{Im}(T)$. To show this, we must show that $\{T(v_i)\}$ is linearly independent and generating.

Generating : Consider an arbitrary $y \in W$. By definition, there exists some $x \in V$ such that $T(x) = y$. Because $w_1, \ldots, w_m, \ldots, v_n$ is a basis of V, there exists some $\alpha_1, \ldots, \alpha_n$ such that $x = \alpha_1 w_1 + \cdots + \alpha_n v_n$. By the linearity of T and the definition of w_i, this implies that

$$\begin{aligned} T(x) &= y \\ \alpha_1 \underbrace{T(w_1)}_{=0} + \cdots + \alpha_m \underbrace{T(w_m)}_{=0} + \alpha_{m+1} T(v_{m+1}) + \cdots + \alpha_n T(v_n) &= y \\ \alpha_{m+1} T(v_{m+1}) + \cdots + \alpha_n T(v_n) &= y \end{aligned}$$

Therefore $\{T(v_i)\}$ is generating.

Linearly Independent : Let there exist $\alpha_{m+1}, \ldots, \alpha_n$ such that

$$\alpha_{m+1} T(v_{m+1}) + \cdots + \alpha_n T(v_n) = 0$$

Then by the linearity of T,

$$T(\alpha_{m+1} v_{m+1}) + \cdots + \alpha_n v_n) = 0$$

and so $\alpha_{m+1} v_{m+1}) + \cdots + \alpha_n v_n \in \ker T$. Because $w_1, \ldots, w_m$ are a basis for $\ker T$, this implies that there exists $-\alpha_1, \ldots, -\alpha_m$ such that

$$\begin{aligned} \alpha_{m+1} v_{m+1}) + \cdots + \alpha_n v_n &= -\alpha_1 w_1) - \cdots - \alpha_m w_m \\ \alpha_1 w_1) + \cdots + \alpha_m w_m + \alpha_{m+1} v_{m+1}) + \cdots + \alpha_n v_n &= 0 \end{aligned}$$

Because $\{w_i, v_i\}$ is a basis, it is linearly independent. Therefore $\alpha_i = 0$ for all i, namely for $i = m+1, \ldots, n$. This implies that $\{T(v_i)\}$ are linearly independent.

(2) Consider the matrix form of T with respect to any basis. Putting T into echelon form, we know that

$$\begin{aligned} \#\text{ of columns of } T \text{ with pivots } &= \dim \operatorname{im}(T) = \operatorname{rank} T \\ \#\text{ of columns of } T \text{ without pivots } &= \dim \ker(T) = \operatorname{nullity} T \\ \#\text{ of columns of } T &= \dim(V) \end{aligned}$$

Because each column must either have a pivot or not have a pivot, this implies that

$$\#\text{ of columns with pivots } + \#\text{ of columns without pivots } = \#\text{ of columns}$$
$$\operatorname{rank} T + \operatorname{nullity} T = \dim(V)$$

Isomorphic
Two vector spaces V, W are isomorphic if there exists linear maps $L : V \to W$ and $K : W \to V$ so that

$$KL = I_V \qquad LK = I_W$$

Subspace Theorem

Quotient Space
Let $W \subset V$ be a subspace. The quotient space, V/W is the vector space of equivalence classes under the equivalence relation $v\ w$ if $v - w \in W$, endowed with the induced scalar multiplication and addition operations

$$\lambda[v] = [\lambda v] \qquad [v] + [w] = [v + w]$$

Where $[v] = v + W$ is the corresponding equivalence class.

This leads to the

Theorem. (Subspace Theorem) : Let $W \subset V$ be a subspace. Then

$$\dim V = \dim W + \dim V/W$$

Proof. Define a linear transformation $T : V \to V/W$ such that $T(v) = [v]$. Then $\ker T = W$ and $\operatorname{im} T = V/W$. Therefore by the rank nullity theorem,

$$\dim(V) = \dim\ker T + \dim\operatorname{im} T = \dim W + \dim V/W$$

Which is what was to be shown.

Rank Theorem

Row Rank and Column Rank
Let A be an $n \times m$ matrix. Then the row rank of A is the maximal number of linearly independent rows and the column rank is the maximal number of linearly independent rows.
Equivalently, row rank is the dimension of the span of the rows and column rank is the dimension of the span of the columns. By definition of matrix multiplication, this implies that the rank of A is the column rank.

Theorem. (Rank Theorem) : Let A be an $n \times m$ matrix. Then the row rank of A is equal to the column rank.

Proof. This follows immediately by counting pivots in the echelon form of A.
Alternatively,
Let $x_1, \dots, x_c$ be a basis for the column space of A. Therefore, for each v in the column space, there exists a unique $\alpha_1, \dots, \alpha_c$ such that $v = \alpha_1 x_1 + \dots + \alpha_c x_c$. Define a linear transformation B : column space of A $\to \mathbb{F}^c$ such that $Bv = (\alpha_1, \dots, \alpha_c)^T$ where $v = \alpha_1 x_1 + \dots + \alpha_c x_c$. With this definition,

$$A = \begin{bmatrix} x_1 & x_2 & \dots & x_c \end{bmatrix} B$$

Taking the transpose,

$$A^T = B^T \begin{bmatrix} x_1 & x_2 & \dots & x_c \end{bmatrix}^T$$

Therefore $\dim \operatorname{im} A^T \leq \dim(\text{column space of } A)$. Because $\operatorname{im} A^T$ is the span of the columns of A^T, it is the span of the rows of A, and so is the row space of A. Therefore, the row rank of A is at most the column rank of A. Repeating this argument but starting with A^T, we find the reverse inequality. Therefore the row rank and column rank are equivalent.

Dual Space

Dual Space
If V is a vector space then

$$V' = \hom(V, \mathbb{F}) := \{f : V \to \mathbb{F} \mid f \text{ is linear } \}$$

Note : if V is infinite dimensional then V' is the algebraic dual space, not the continuous dual space.

Theorem. (Basis of Dual Space) : If V has basis $v_1, \dots, v_n$ then V' has basis $\phi_1, \dots, \phi_n$ where $\phi_i(v_j) = \delta_{ij}$. In particular, $\dim(V) = \dim(V')$.

Proof. The proof of this is straightforward manipulation of bases and is not useful.

Annihilator
If $U \subset V$ is a subset then

$$U^0 = U^\perp = \{f \in V' : f(u) = 0 \text{ for all } u \in U\}$$

Theorem. Let V be a finite dimensional vector space with subspace W. Then

$$\dim(V) = \dim(W) + \dim(W^0)$$

Proof. Define a linear transformation $T : W \to V$ such that $T(v) = v$. We can then define its dual map $T' : V' \to W'$ as

$$T'(f) = f \circ T$$

Then $\ker T' = W^0$ because $\operatorname{im} T = W$ and $\operatorname{im} T' = W'$ because T is the identity on W. Therefore by rank nullity,

$$\dim V' = \operatorname{nullity} T' + \operatorname{rank} T' = \dim W^0 + \dim W'$$

Because we are working with finite dimensional vector spaces,

$$\dim V = \dim W^0 + \dim W$$

Which is what was to be shown.

Volume Forms and Determinants

Volume Forms

Volume Form
Let V be an n dimensional vector space over a field $\mathbb{F}$. A volume form is a multilinear map

$$\text{vol} : V^n \to \mathbb{F}$$

such that if $v_i = v_j$ for $i \neq j$ then

$$\text{vol}(v_1, \ldots, v_n) = 0$$

Theorem. If $1 \neq -1$ then the volume form is alternating. That is

$$\text{vol}(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_n) = -\text{vol}(v_1, \ldots, v_j, \ldots, v_i, \ldots, v_n)$$

Proof. By definition of a volume form,

$$\begin{aligned} 0 &= \text{vol}(v_1, \ldots, v_i + v_j, \ldots, v_i + v_j, \ldots, v_n) \\ 0 &= \text{vol}(v_1, \ldots, v_i, \ldots, v_i, \ldots, v_n) + \text{vol}(v_1, \ldots, v_j, \ldots, v_i, \ldots, v_n) \\ \text{vol}(v_1, \ldots, v_i, \ldots, v_i, \ldots, v_n) &= -\text{vol}(v_1, \ldots, v_j, \ldots, v_i, \ldots, v_n) \end{aligned}$$

Which is what was to be shown.

Theorem. If $v_1, \ldots, v_n \in V$ satisfy $\text{vol}(v_1, \ldots, v_n) \neq 0$ then $v_1, \ldots, v_n$ is a basis for V.

Proof. If $v_1, \ldots, v_n$ are linearly dependent then $\text{vol}(v_1, \ldots, v_n) = 0$ because v_i can be written as a linear combination of the other vectors. Therefore, $v_1, \ldots, v_n$ are linearly independent. Because there are n vectors, this implies that $v_1, \ldots, v_n$ are a basis.

Theorem. Let $v_1, \ldots, v_n$ be a basis of V. Then there exists a volume form vol such that

$$\text{vol}(v_1, \ldots, v_n) = 1$$

Additionally, if vol' is any other volume form then there exists $\lambda \in \mathbb{F}$ such that

$$\text{vol}' = \lambda \text{vol}$$

Determinant

Determinant
Let $v_1, \ldots, v_n$ be a basis for V. Let vol be a non-trivial volume form. Then we define the determinant of a linear operator $T : V \to V$ to be

$$\det T = \frac{\text{vol}(Tv_1, \ldots, Tv_n)}{\text{vol}(v_1, \ldots, v_n)}$$

(This amounts to taking the volume form of the columns of T)

Theorem. (Laplace Expansion) : *Row and Column Expansion of Determinant*
Let $n \geq 2$ and $A \in M_n(\mathbb{F})$. Denote the $(i,j)^{th}$ minor of A be M_{ij}. Then

$$\det A = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} M_{ij} \text{ for fixed j}$$
$$\det A = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} M_{ij} \text{ for fixed i}$$

Corollary : Repeating this expansion,

$$\det(A)I = (\text{adj}\, A)A = A(\text{adj}\, A)$$

Where $(\text{adj}\, A)_{ij} = (-1)^{i+j} M_{ji}$.

The Characteristic Polynomial

Characteristic Polynomial
Let $T : V \to V$ be a linear operator. The characteristic polynomial of T is $\chi_T(t) = \det(tI - T)$.
Note : This is a monic polynomial of degree $\dim V$.

Eigensection

Theorem. λ is an eigenvalue of T if and only if $\chi_T(\lambda) = 0$.

Proof. We know that $\lambda I - T$ is invertible if and only if $\det(\lambda I - T) = \chi_T(\lambda) = 0$. Additionally, $\lambda I - T$ is invertible if and only if $\ker(\lambda I - T) = \{0\}$, which occurs if and only if there doesn't exist a v such that $Tv = \lambda v$. Therefore, λ is an eigenvalue if and only if $\chi_T(\lambda) = 0$.

Algebraic Multiplicity
Let λ be an eigenvalue of T. If $\chi_T(t) = (t - \lambda)^m p(t)$ where $p(\lambda) \neq 0$ then m is the algebraic multiplicity of λ.

Geometric Multiplicity
Let λ be an eigenvalue of T. The geometric multiplicity of λ is

$$\text{nullity}(\lambda I - T) = \dim E_\lambda \tag{1}$$

Where E_λ is the space of eigenvectors with eigenvalue λ.

Theorem. Let $A \in M_n(\mathbb{C})$ have eigenvalues $\lambda_1, \ldots, \lambda_n$ repeated according to algebraic multiplicity. then

$$\det A = \prod_i \lambda_i$$
$$\operatorname{tr} A = \sum_i \lambda_i$$

The Minimal Polynomial

Minimal Polynomial
Let $T : V \to V$ be a linear operator. The minimal polynomial of T is the smallest order monic non-trivial polynomial

$$\mu_T(t) = t^k + \alpha_{k-1}t^{k-1} + \cdots + \alpha_0$$

Such that $\mu_T(T) = 0$.
As defined, k is the smallest positive integer such that $\{1, T, \ldots, T^k\}$ is linearly dependent.
Note : If $\mu_T(t)$ is the minimal polynomial of T then no non-trivial polynomial of degree $< k$ satisfies $p(T) = 0$.

Theorem. Let $T : V \to V$ be a linear operator with minial polynomial $\mu(t)$. Then

(1) If T satisfies a polynomial $p(t)$ then $\mu | p$.

(2) $\lambda \in \mathbb{F}$ is an eigenvalue of T if and only if $\mu(\lambda) = 0$

Proof.

(1) If T satisfies $p(t)$ then the degree of $p(t)$ is at least the degree of $\mu(t)$. Therefore, because $\mathbb{F}[x]$ is a Euclidean domain, there exists $q(t)$ and $r(t)$ such that $r(t)$ has degree less than $\mu(t)$ and

$$p(t) = q(t)\mu(t) + r(t) \implies 0 = p(T) = r(T)$$

Because $r(t)$ has degree less than $\mu(t)$, this implies that $r(t) = 0$ and so $p(t) = q(t)\mu(t)$.

(2) *Forwards :* If λ is an eigenvalue of T then

$$\mu(T)v = 0$$
$$\mu(\lambda)v = 0$$

Which implies that $\mu(\lambda) = 0$ because $v \neq 0$.

Backwards : If $\mu(\lambda) = 0$ then $\mu(t) = (t - \lambda)p(t)$ for some $p(t)$. Because $\mu(t)$ has minimal degree, $p(T) \neq 0$. Therefore, there exists some $w \in V$ such that $v = p(T)w \neq 0$. Then

$$\mu(T)w = 0$$
$$(T - \lambda I)p(T)w = 0$$
$$(T - \lambda I)v = 0$$
$$Tv = \lambda v$$

So λ is an eigenvalue of T.

Cayley-Hamilton Theorem

Theorem. (Cayley-Hamilton Theorem) : Let $T : V \to V$ be a linear operator. Then $\chi_T(T) = 0$. In particular, $\mu_T | \chi_t$.

Cyclic Subspace C_v
Let $T : V \to V$ be a linear operator and let there exist $v \in V$. Then

$$C_v = \text{span}\{v, Tv, T^2v, \dots\}$$

is the cyclic subspace generated by T and v.

Theorem. Let V have dimension n and $T : V \to V$ be a linear operator. Then

(i) C_v is T-invariant. That is to say that $T(C_v) \subset C_v$.

(ii) If $v \neq 0$ then there exists a positive integer $k \leq n$ such that $v, Tv, \dots, T^{k-1}v$ is a basis for C_v. Further, the matrix representation of T with respect to this basis is

$$T = \begin{bmatrix} 0 & 0 & 0 & 0 & \dots & -\alpha_0 \\ 1 & 0 & 0 & 0 & \dots & -\alpha_1 \\ 0 & 1 & 0 & 0 & \dots & -\alpha_2 \\ 0 & 0 & 1 & 0 & \dots & -\alpha_3 \\ 0 & 0 & 0 & 1 & \dots & -\alpha_4 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \dots & -\alpha_{k-1} \end{bmatrix}$$

Where $T^k v + \alpha_{k-1}T^{k-1}v + \dots + \alpha_0 v = 0$.

(iii) The characteristic polynomial of $T|_{C_v}$ is $\chi_{T|_{C_v}}(t) = t^k + \alpha_{k-1}t^{k-1} + \dots + \alpha_0 = 0$.

Proof. Left as an exercise for me.

Proof. **(Cayley-Hamilton Theorem) :** Let there exist $v \in V \setminus \{0\}$. Let $v, Tv, \dots, T^{k-1}v$ be a basis for C_v. Extend to a basis $v, Tv, \dots, T^{k-1}v, w_{k+1}, \dots, w_n$ of V. Then T has block matrix form

$$T = \begin{bmatrix} A & B \\ 0 & D \end{bmatrix}$$

Where A has the same form as in the previous theorem. Then $\chi_T(t) = \chi_A(t)\chi_D(t)$ which implies that

$$\chi_T(T)v = \chi_A(T) \underbrace{\chi_D(T)v}_{0} = 0$$

By the previous theorem. As this holds for all $v \in V$, $\chi_T(T) = 0$.

Theorem. (Bon-Soon's Theorem) : Let V be a vector field over $\mathbb{F}$ and let $\dim(V) = n$. Let there exist a linear operator $A : V \to V$. Then

$$\chi_A(t)|\mu_A(t)^n$$

Proof. If we extend $\mathbb{F}$ to a splitting field, then the proof is trivial as any roots of μ_A must be roots of χ_A and vice versa. Excluding the use of a splitting field, the proof is a little bit more involved. I would recommend the lecture notes from Tuesday of week 2's video.

Diagonalizability

Diagonalizable
A linear operator $T : V \to V$ is diagonalizable if there is a basis of V consisting of eigenvectors of T.

Theorem. $T : V \to V$ is diagonalizable if and only if the minimal polynomial factors as

$$\mu_T(t) = (t - \lambda_1) \dots (t - \lambda_k)$$

Where the λ_i are distinct.

Theorem. If $\dim V = n$ and T has n distinct eigenvalues then it is diagonalizable.

Proof. Immediate corollary from previous theorem because $\mu_T | \chi_T$.

Theorem. Let V be a vector space over $\mathbb{C}$. Then a linear operator $T : V \to V$ is diagonalizable if and only if for every eigenvalue, the geometric and algebraic multiplicities are the same.

Jordan Canonical Form

For this section, all vector spaces are over $\mathbb{C}$.

Theorem. (Jordan-Chevally Decomposition) : Let $L : V \to V$ be a linear operator. Then $L = S + N$ where S is diagonalizable, N is nilpotent ($N^k = 0$), and $[N, S] = 0$.

Theorem. (Jordan Canonical Form) : Let $L : V \to V$ be a linear operator where V is a complex vector space. Then we can find L-invariant subspaces $M_1, \dots, M_s$ such that

$$V = M_1 \oplus \dots \oplus M_s$$

and there is a basis for M_j such that the matrix representation of $L|_{M_j}$ is a Jordan block.

Jordan Block
A block of the form

$$\begin{bmatrix} \lambda & 1 & 0 & \dots & 0 \\ 0 & \lambda & 1 & \dots & 0 \\ 0 & 0 & \lambda & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \lambda \end{bmatrix}$$

Where λ is the eigenvalue associated with the block.

This decomposition is unique, up to reordering the blocks.

How to Calculate the Jordan Form of a Matrix

Let A be a $n \times n$ complex matrix. This can be done with any matrix such that the characteristic polynomial splits, but we will only use this for complex matrices.

(1) Find the distinct eigenvalues $\lambda_1, \ldots, \lambda_k$ of A.

(2) Pick some eigenvalue λ and note the algebraic multiplicity, $\mathrm{am}(\lambda)$.

> **Theorem.** The algebraic multiplicity of an eigenvalue is the total size of Jordan blocks involving said eigenvalue.

(3) Compute

$$\begin{aligned} \mathrm{nullity}(A - \lambda I) &= n_1 \\ \mathrm{nullity}(A - \lambda I)^2 &= n_2 \\ \mathrm{nullity}(A - \lambda I)^3 &= n_3 \\ &\vdots \\ \mathrm{nullity}(A - \lambda I)^m &= n_m = \mathrm{am}(\lambda) \end{aligned}$$

Then there are

$$\begin{aligned} n_1 &\text{ Jordan blocks} \\ n_2 - n_1 &\text{ Jordan blocks of size } \geq 2 \\ n_3 - n_2 - n_1 &\text{ Jordan blocks of size } \geq 3 \\ &\vdots \\ n_m - n_{m-1} - \cdots - n_1 &\text{ Jordan blocks of size } \geq m \text{ (largest size)} \end{aligned}$$

From this, we find that

> **Theorem.** The number of Jordan blocks corresponding to an eigenvalue is equivalent to the geometric multiplicity of λ, which is $\mathrm{nullity}(A - \lambda I)$.

Additionally,

> **Theorem.** The power of $(t - \lambda)$ in $\mu_A(t)$, the minimal polynomials of A, is the biggest block with eigenvalue λ.

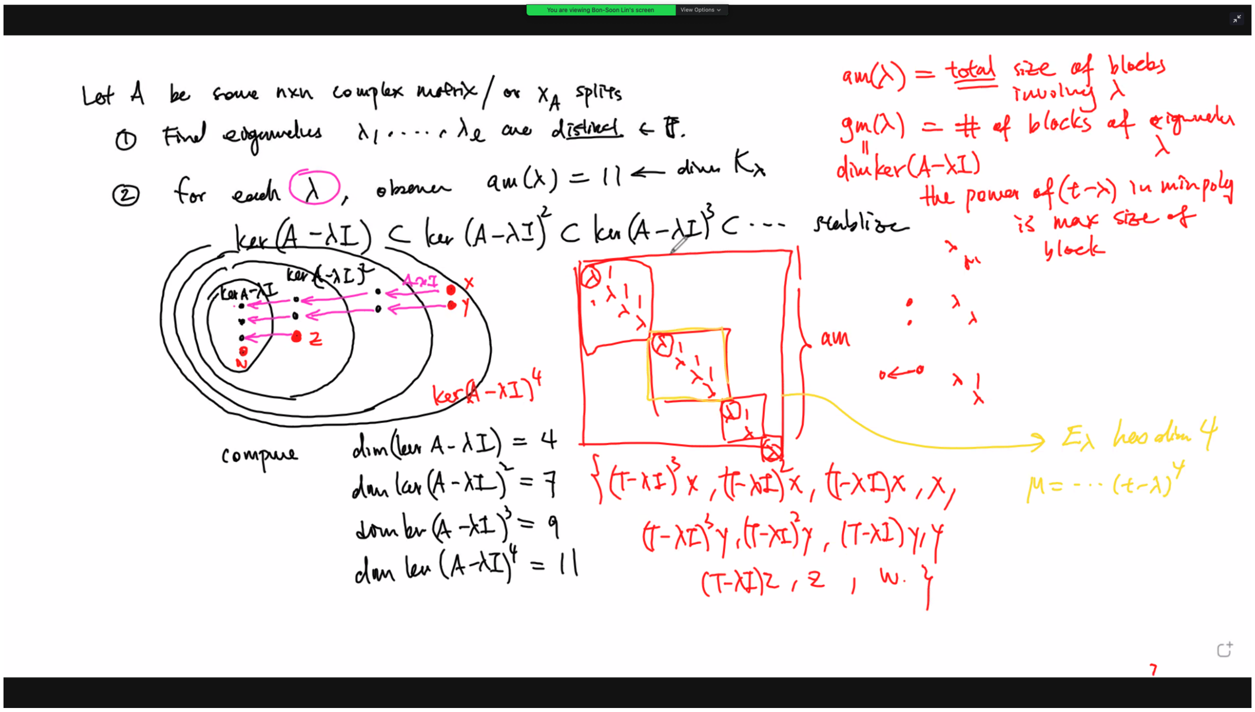
(4) Now we will find the basis that puts our matrix into this form. We will give a procedure to do it for one eigenvalue, which must be repeated for all other blocks.

Find x such that $(A - \lambda I)^{\mathrm{am}(\lambda)} x = 0$ but $(A - \lambda I)^{\mathrm{am}(\lambda)-1} x \neq 0$. Then add $x, (A - \lambda I)x, \ldots, (A - \lambda I)^{\mathrm{am}(\lambda)} x$ to our set of basis vectors.

Repeat this process for the next smallest block. However, it is crucial that the first vector chosen is linearly independent from all previous vectors.

See Bonsoon's notes below for extra clarification.

The vectors in this order will produce the classic Jordan canonical form.


Let A be some n×n complex matrix / or χ_A splits
① Find eigenvalues λ_1, ..., λ_ℓ are distinct ∈ F.
② for each λ, observe am(λ) = 11 ← dim K_λ
ker(A − λI) ⊂ ker(A−λI)^2 ⊂ ker(A−λI)^3 ⊂ ··· stabilize
compare dim(ker A − λI) = 4
dim ker(A − λI)^2 = 7
dim ker(A − λI)^3 = 9
dim ker(A − λI)^4 = 11
am(λ) = total size of blocks involving λ
gm(λ) = # of blocks of eigenvalue λ
dim ker(A−λI)
the power of (t−λ) in minpoly is max size of block
am
E_λ has dim 4
μ = ··· (t−λ)^4
{(T−λI)^3 x, (T−λI)^2 x, (T−λI)x, x,
(T−λI)^3 y, (T−λI)^2 y, (T−λI)y, y
(T−λI)z, z, w.}

Inner Product Space

Unless otherwise stated, all vector spaces are finite dimensional spaces over $\mathbb{C}$.

(Hermitian) Inner Product
An inner product is a map

$$\langle *, * \rangle : V \times V \to \mathbb{C}$$

such that for all $x, y \in V$,

(i) *Positive Definiteness :* $\langle x, x \rangle \geq 0$ with equality if and only if $x = 0$.

(ii) *Skew-Symmetry :* $\langle x, y \rangle = \overline{\langle y, x \rangle}$

(iii) *Linearity in First Argument :* $x \mapsto \langle x, y \rangle$ is linear.

From now on, V is a complex, finite dimensional, inner product space.

Norm Induced by Inner Product
We define the canonical norm on an inner product space to be

$$\|x\| = \sqrt{\langle x, x \rangle}$$

Orthogonal
Two vectors $x, y \in V$ are orthogonal if $\langle x, y \rangle = 0$.

Theorem. (Cauchy Schwartz and Others) : For all $x, y \in V$,

(i) (This one is Cauchy Schwartz)

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

(ii)

$$\|x + y\| \leq \|x\| + \|y\|$$

Theorem. (Pyhtagoras' Inequality) : If $x, y \in V$ are orthogonal then

$$\|x + y\|^2 = \|x\|^2 + \|y\|^2$$

Proof.

$$\begin{aligned}
\|x + y\|^2 &= \langle x + y, x + y \rangle \\
&= \langle x, x \rangle + \langle x, y \rangle + \langle y, x \rangle + \langle y, y \rangle \\
&= \|x\|^2 + \|y\|^2
\end{aligned}$$

Orthogonal Projections

Orthogonal Projection
Let there exist $y \in V \setminus \{0\}$. The orthogonal projection of $x \in V$ in the direction of y is

$$\operatorname{proj}_y(x) = \langle x, \frac{y}{\|y\|} \rangle \frac{y}{\|y\|} = \frac{\langle x, y \rangle}{\|y\|^2} y$$

Theorem. Let there exist $y \in V \setminus \{0\}$. Then

(i) $\operatorname{proj}_y : V \to V$ is a projection

(ii) For any $x \in V$, $\operatorname{proj}_y(x)$ and $x - \operatorname{proj}_y(x)$ are orthogonal

(iii) For any $x \in V$,
$$\| \operatorname{proj}_y(x)\| \leq \|x\|$$

Orthogonal/Orthonormal Set
A set of vectors $e_1, \ldots, e_n \in V$ are said to be orthogonal if $\langle e_i, e_j \rangle = 0$ if $i \neq j$.
The set is said to be orthonormal if $\langle e_i, e_j \rangle = \delta_{ij}$.

Theorem. Let $e_1, \ldots, e_n$ be orthonormal. Then

(1) The set $\{e_1, \ldots, e_n\}$ is linearly independent.

(2) If $x \in \operatorname{span}\{e_1, \ldots, e_n\}$ then
$$x = \sum_{j=1}^{n} \langle x, e_j \rangle e_j$$
and
$$\|x\|^2 = \sum_{j=1}^{n} |\langle x, e_j \rangle|^2$$

Proof.

(1) If $a_1 e_1 + a_2 e_2 + \cdots + a_n e_n$ then
$$a_j = \langle e_j, a_1 e_1 + a_2 e_2 + \cdots + a_n e_n \rangle = 0$$

(2) Direct computation.

Theorem. (Gram-Schmidt Process) : For any linearly independent set $\{v_1, \ldots, v_n\}$, there exists an orthonormal set $\{e_1, \ldots, e_n\}$ such that for all $m \leq n$, $\operatorname{span}\{v_1, \ldots, v_m\} = \operatorname{span}\{e_1, \ldots, e_m\}$.

Proof. Proceed by induction on m. For $m = 1$ take

$$e_1 = \frac{v_1}{\|v_i\|}$$

Assume that we have constructed $e_1, \ldots, e_{m-1}$. Then let

$$w_m = v_m - \sum_{i=1}^{m-1} \langle v_m, e_j \rangle e_j$$

and finally

$$e_m = \frac{w_m}{\|w_m\|}$$

Orthogonal Complement

Orthogonal Complement
Let $M \subset V$ be a subspace. Then the orthogonal complement to M is

$$M^\perp = \{x \in V : \langle x, z \rangle = 0 \text{ for all } z \in M\}$$

Orthogonal Projection
Let $M \subset V$ be a subspace with orthonormal basis $e_1, \ldots, e_n$. Then the orthogonal projection of $x \in V$ onto M is

$$\text{proj}_M(x) = \sum_{j=1}^{n} \langle x, e_j \rangle e_j$$

The matrix form of this is

$$\text{proj}_M = \sum_{j=1}^{n} e_j e_j^*$$

Theorem. Let $M \subset V$ be a subspace. Then

(i) The map $\text{proj}_M : V \to V$ is a projection onto M.

(ii) For all $x \in V$, $\text{proj}_M(x) \in M$ and $x - \text{proj}_M(x) \in M^\perp$.

(iii) $V = M \oplus M^\perp$

(iv) $(M^\perp)^\perp = M$

Theorem. Let $M \subset V$ be a subspace. Then for all $x \in V$, $\text{proj}_M(x)$ is the vector in M that is closest to x.

Proof. $\min_{m\in M}\|x-m\| = \min_{m\in M}\|(x-\text{proj}_M(x))+\text{proj}_M(x)-m\| = \min_{m\in M}\|(x-\text{proj}_M(x))\|+\|\text{proj}_M(x)-m\|$ Which is minimal if and only if $m=\text{proj}_M x$.

Adjoint

Theorem. Let $L : V \to V$ be a linear map between finite-dimensional inner product spaces over $\mathbb{C}$. Then there exists a unique linear map $L^* : W \to V$ such that for all $v \in V$ and for all $w \in W$,

$$\langle Lv, w\rangle_W = \langle v, L^*w\rangle$$

Proof. The proof of this really just comes from rearranging things. Viewing them in matrix notation though,

$$\langle Lv, w\rangle = v^*L^*w = \langle v, L^*w\rangle$$

Adjoint
The linear map L^* is called the adjoint of L.

Theorem. (The Fredholm Alternative) : Let $L : V \to W$ be a linear map between complex inner product spaces. Then

$$\ker L = \text{im}(L^*)^\perp$$
$$\ker L^* = \text{im}(L)^\perp$$
$$(\ker L)^\perp = \text{im}(L^*)$$
$$(\ker L^*)^\perp = \text{im}(L)$$

Theorem. (The Polarization Identity) : For $x, y \in V$,

$$\langle x, y\rangle = \frac{1}{4}\left(\|x+y\|^2 - \|x-y\|^2 - i\|x-iy\|^2 + i\|x+iy\|^2\right)$$

Isometry
A linear operator $L : V \to W$ between inner product spaces is said to be an isometry if, for every $x \in V$,

$$\|Lx\|_W = \|x\|_V$$

Orthogonal Matrix
An isometry in $\mathbb{R}^n$.

Unitary Matrix
An isometry in $\mathbb{C}^n$.

Special Types of Matrices

Self-Adjoint
$L = L^*$

Skew-Adjoint
$L = -L^*$

Normal
$[L, L^*] = 0$.

Spectral Theorem Fist Go

Theorem. If $A \in M_n(\mathbb{R})$ is symmetric and has largest eigenvalue $\lambda \in \mathbb{R}$ then for any $v \in \mathbb{R}^n$,

$$\langle Av, v\rangle \leq \lambda \|v\|^2$$

with equality if and only if $Av = \lambda v$.

Proof. Follows by definitions from the spectral theorem

Theorem. (The Spectral Theorem) : If $A \in M_n(\mathbb{R})$ is symmetric then there is an orthonormal basis $e_1, \ldots, e_n$ of $\mathbb{R}^n$ of eigenvectors of A.

The Spectral Theorem

Theorem. (The Spectral Theorem) : Let $L : V \to V$ be a self-adjoint linear operator on a real or complex inner product space. Then there exists an orthonormal bassi of V of eigenvectors of L.
Corollary : Let $A \in M_n(\mathbb{R})$ be symmetric. Then there exists an orthogonal matrix O such that $O^T AO$ is diagonal.
Corollary : Let $A \in M_n(\mathbb{C})$ be Hermitian. Then there exists a unitary matrix U such that U^*AU is diagonal.

Proof. Usually, the spectral theorem is proven in four parts, outlined by this warm-up exercise.

Warm-up Problem 1 (S08-12, S10-2)**.** Let A be an $n \times n$ real symmetric matrix.

(i) Prove that there exists $x \in \mathbb{S}^{n-1} = \{x \in \mathbb{R}^n : \|x\| = 1\}$ so that

$$\langle Ax, x\rangle = \sup\left\{\langle Ay, y\rangle : y \in \mathbb{S}^{n-1}\right\}.$$

(ii) Prove that if $\langle x, y\rangle = 0$ then $\langle Ax, y\rangle = 0$.

(iii) Use (ii) to prove that x is an eigenvector for A.

(iv) *(The Spectral Theorem)* Use induction to prove that $\mathbb{R}^n$ has an orthonormal basis of eigenvectors for A.

(i) The function $y \mapsto \langle Ay, y\rangle$ is continuous on $\mathbb{S}^{n-1}$, which is compact. Therefore, the function attains its maximum at some $x \in \mathbb{S}^{n-1}$.

(ii) If this property holds for unit vectors y, then it easily generalizes to all vectors y. Therefore, assume that y is a unit vector. Define a vector v_t such that

$$v_t = \cos(t)x + \sin(t)y$$

Because $\langle x, y\rangle = 0$, we know that v_t is a unit vector. Define a function f such that

$$f(t) = \langle Av_t, v_t\rangle$$

Expanding this,

$$\begin{aligned} f(t) &= \langle A(\cos(t)x + \sin(t)y), \cos(t)x + \sin(t)y\rangle \\ &= \cos^2(t)\langle Ax, x\rangle + \sin^2(t)\langle Ay, y\rangle + 2\sin(t)\cos(t)\langle Ax, y\rangle \end{aligned}$$

As shown by the expansion, $f(t)$ is differentiable in t and has derivative,

$$f'(t) = 2\cos(t)\sin(t)\langle Ax, x\rangle + 2\sin(t)\cos(t)\langle Ay, y\rangle + 2(\cos^2(t) - \sin^2(t))\langle Ax, y\rangle$$

As found in part 1, f is maximal when $v_t = x$. Therefore, $f(t)$ has a maximum at $x = 0$ and so $f'(0) = 0$. This implies that

$$\begin{aligned} f'(0) &= 0 \\ 2\cos(0)\sin(0)\langle Ax, x\rangle + 2\sin(0)\cos(0)\langle Ay, y\rangle + (2\cos^2(0) - 2\sin^2(0))\langle Ax, y\rangle &= 0 \\ \langle Ax, y\rangle &= 0 \end{aligned}$$

(iii) For $n = 1$, the result is clear from the previous work.

Assume that the result holds for $n-1$. By part (iii), we can find an unitary eigenvector x of A. Complete x to an orthonormal basis $x, e_2, \ldots, e_n$. By part (ii), because $\langle x, e_i\rangle = 0$ for all i, $\langle Ax, e_i\rangle = \langle x, Ae_i\rangle = 0$ for all i. Therefore $\text{span}\{x\}^\perp$ is A-invariant. Consider the operator $A|_{\text{span}\{x\}^\perp}$. By the inductive assumption, we can find an orthonormal basis of eigenvectors of $A|_{\text{span}\{x\}^\perp}$. This completes the proof.

Theorem. (The Min-Max Theorem) : Let $L : V \to V$ bee a self-adjoint linear operator on a real or complex inner product space of dimension n. Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of L repeated according to algebraic multiplicity. Then

$$\lambda_k = \min_{U \subset V \text{ subspace} \dim U = k} \left[\sup_{\substack{x \in U \\ \|x\|=1}} \langle Lx, x\rangle \right] \tag{2}$$

Theorem. (The Spectral Theorem for Skew-Adjoint): Let $L : V \to V$ be a skew adjoint linear operator on a **complex** inner product space. Then there is an orthonormal basis of V of eigenvectors of L.

Proof. Same as the spectral theorem but replace L with iL because iL is symmetric.

Theorem. (The Spectral Theorem for Normal Operators): Let $L : V \to V$ be a linear operator on a **complex** inner product space. Then L is normal if and only if there is an orthonormal basis of V of eigenvectors of L.

Proof. Decompose $L = \underbrace{\frac{1}{2}(L + L^*)}_{A} + i \underbrace{\frac{1}{2i}(L - L^*)}_{B}$. Then A, B are self-adjoint and $[A, B] = 0$.
Then it can be shown that A and B are simultaneously diagonalizable.
Backwards direction follows easily.

Theorem. (Spectral Resolution): Let $L : V \to V$ be a normal linear operator on a complex inner product space with distinct eigenvalues $\lambda_1, \ldots, \lambda_k$ and corresponding eigenspaces $E_j = \ker(\lambda_j I - L)$. Then

$$\begin{aligned} I_V &= \sum_{j=1}^{k} \text{proj}_{E_j} \\ L &= \sum_{j=1}^{k} \lambda_j \, \text{proj}_{E_j} \end{aligned} \tag{3}$$

Real Canonical Form

Theorem. (Real Canonical Form) : Let $L : V \to V$ be a normal linear operator on a **real** inneere product space V of dimension n. Then there is an orthonormal basis $e_1, \dots, e_k, x_1, y_1, \dots, x_\ell, y_\ell$ where $k + 2\ell = n$ and

$$\begin{aligned} Le_j &= \lambda_j e_j \\ Lx_j &= \alpha_j x_j + \beta_j y_j \\ Ly_j &= -\beta_j x_j + \alpha_j y_j \end{aligned}$$

With $\lambda_j, \alpha_j, \beta_j \in \mathbb{R}$. Then the matrix of L with respect to this basis is

$$L = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & \lambda_2 & \dots & 0 & 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & \lambda_k & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & \alpha_1 & -\beta_1 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & \beta_1 & \alpha_1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & \alpha_\ell & -\beta_\ell \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & \beta_\ell & \alpha_\ell \end{bmatrix}$$

Note : If considered in $\mathbb{C}$, then L has eigenvalues $\lambda_i \in \mathbb{R}$ and $\alpha_j + \beta_j i, \alpha_j - \beta_j i \in \mathbb{C}$.

Quadratic Forms

Bilinear Form
Let V be a real vector space. A bilinear form on V is a map $B : V \times V \to \mathbb{R}$ which is linear in both variables. We say B is symmetric if for all $x, y \in V$, $B(x, y) = B(y, x)$.

Sesquilinear Form
Let V be a complex vector space. A sesquilinear form on V is a map $B : V \times V \to \mathbb{R}$ which is linear in the first variable and conjugate linear in the second. We say B is Hermitian if for all $x, y \in V$, $B(x, y) = \overline{B(y, x)}$.

Theorem. (Characterization of Symmetric/Hermitian Forms) : All these form things can also be expressed in Matrix form just by $A_{ij} = B[e_j, e_i]$. But also

(1) B is a symmetric bilinear form on $\mathbb{R}^n$ if and only if there is a symmetric matrix $A \in M_n(\mathbb{R})$ such that $B[x, y] = \langle Ax, y \rangle = y^T Ax$.

(2) B is a Hermitian sesquilinear form on $\mathbb{C}^n$ if and only if there is a Hermitian matrix $A \in M_n(\mathbb{C})$ such that $B[x, y] = \langle Ax, y \rangle = y^* Ax$.

Proof. Follows from $A_{ij} = B[e_j, e_i]$. Cause obviously.

Positive/Negative (Semi)Definite
A self-adjoint matrix $A \in M_n(\mathbb{F})$ with $\mathbb{F} = \mathbb{C}$ or $\mathbb{R}$ is said to be positive/negative definite if $\langle Ax, x\rangle$ is strictly positive/negative whenever $x \neq 0$.
A self-adjoint matrix $A \in M_n(\mathbb{F})$ with $\mathbb{F} = \mathbb{C}$ or $\mathbb{R}$ is said to be positive/negative semidefinite if $\langle Ax, x\rangle$ is non-negative/non-positive for all $x \in \mathbb{F}$.
The same terminology applies to the corresponding bilinear form.

Signature

Theorem. (Digaonalization of Symmetric/Hermitiaion Bilinear/Sesquilinear Form) : Let B be a symmetric bilinear form on $\mathbb{R}^n$. Then there exists an orthogonal matrix $O \in M_n(\mathbb{R})$ such that if $x, y \in \mathbb{R}^n$ and $x = Ox'$, $y = Oy'$ then

$$B[x, y] = \sum_{i=1}^{n} \lambda_i x_i' y_i'$$

where $\lambda_i \in \mathbb{R}$.
Similarly, if B is a Hermitian sesquilinear form on $\mathbb{R}^n$. Then there exists a unitary matrix $U \in M_n(\mathbb{C})$ such that if $x, y \in \mathbb{C}^n$ and $x = Ux'$, $y = Uy'$ then

$$B[x, y] = \sum_{i=1}^{n} \lambda_i x_i' \overline{y_i'}$$

where $\lambda_i \in \mathbb{R}$.

Proof. Spectral Theorem

Signature
Taking λ_j as in the preceeding theorem, we define the signature of B to consist of $n^+ =$ number of positive λ_j's, $n_- =$ number of negative λ_j's and n_0 =number of zero λ_j's.

Theorem. (Decomposition by Signature) : Let B be a symmetric bilinear form on $\mathbb{R}^n$ with signature (n_+, n_-, n_0). Then there exists invertible $P \in M_n(\mathbb{R})$ suhc that if $x, y \in \mathbb{R}^n$ with $x = Px'$ and $y = Py'$ then

$$B[x, y] = \sum_{j=1}^{n_+} x_j' y_j' - \sum_{j=n_+ + 1}^{n_+ + n_-} x_j' y_j'$$

A similar result with obvious modifications occurs in the complex case.

Proof. Spectral theorem and scaling

Theorem. (Sylvester's Law of Inertia) : Let B be a symmetric bilinear form on $\mathbb{R}^n$ with signature (n_+, n_-, n_0). Let $v_1, \dots, v_n$ be a basis of $\mathbb{R}^n$ such that $B[v_i, v_j] = 0$ for all $i \neq j$. Then, up to reordering, $B[v_j, v_j]$ is positive for $j = 1, \dots, n_+$, negative for $j = n_+ + 1, \dots, n_- + n_+$. and zero otherwise.
An essentially identical statement holds in the complex case.

Quadratic Forms

Quadratic Form
A quadratic form on $\mathbb{R}^n$ is a map $Q : \mathbb{R}^n \to \mathbb{R}$ defined by

$$Q(x) = B[x, x]$$

Where B is a symmetric bilinear form on $\mathbb{R}^n$. Notions of positive/negative (semi-)definiteness are inherited from B.

Characterizations of Quadratic Forms
We say that a quadratic form Q on $\mathbb{R}^n$ is

- Elliptic if $n_+ = n$ or $n_- = n$.
- Hyperbolic if $n_0 = 0$ but $n_+ \neq n$ and $n_- \neq n$.
- Parabolic if $n_0 \neq 0$.

Descartes' Rule of Signs

Theorem. (Descartes' Rule of Signs): Let $p(t) = t^n + a_{n-1}t^{n-1} + \cdots + a_0 = (t - \lambda_1)\dots(t - \lambda_n)$ where $a_j, \lambda_j \in \mathbb{R}$. Then

(I) Zero is a root of p if and only if $a_0 = 0$.

(II) All roots of p are negative if and only if $a_j > 0$.

(III) All roots of p are positive if and only if $a_{n-1} < 0$ and the remaining a_j's alternate signs. As in $a_{n-2} > 0$, $a_{n-3} < 1, \dots$.

(IV)

Matrix Norms

Norm
Let V be a vector space over $\mathbb{F} = \mathbb{R}, \mathbb{C}$. A norm is a map $\| * \| : V \to [0, \infty)$ such that

(i) $\|v\| = 0$ if and only if $v = 0$

(ii) If $\lambda \in \mathbb{F}$ and $v \in V$ then

$$\|\lambda v\| = |\lambda| \|v\|$$

(iii) If $v, w \in V$ then

$$\|v + w\| \leq \|v\| + \|w\|$$

Note : If $\langle *, * \rangle$ is an inner product on V then $\|v\| = \sqrt{\langle v, v \rangle}$ is a norm.

Operator Norm
Let V, W be normed vector spaces and $L : V \to W$ be a linear map. The operator norm of L is

$$\|L\| = \sup\{\|Lx\|_W : x \in V \text{ and } \|x\|_V = 1\}$$

If it exists.

Theorem. Let V, W be finite dimensional normed vector spaces and $L : V \to W$ be a linear map. Then

(1) $\|L\| < \infty$

(2) For all $x \in V$,

$$\|Lx\|_W \leq \|L\| \|x\|_V$$

(3) The operator norm is a norm on $\hom(V, W)$.

Theorem. ($\hom(V, W)$ is Complete) : If V, W are finite dimensional normed vector spaces and $(L_n)_{n \geq 1} \subset \hom(V, W)$ is Cauchy w.r.t the operator norm, then there exists a unique $L \in \hom(V, W)$ such that $L_n \to L$ in operator norm.

Theorem. Let $A \in M_n(\mathbb{C})$ satisfy $\|A\| < 1$. Then $I - A$ is invertible and $\sum_{j=0}^{\infty} A^j$ converges to $(I - A)^{-1}$.

Singular Value Decomposition

Theorem. (Singular Value Decomposition) : Let $L : V \to W$ be a linear operator between finite dimensional inner product spaces. Then there exists an orthonormal basis $e_1, \ldots, e_m$ for V such that for some $\sigma_1, \ldots, \sigma_m \geq 0$ we have

$$\langle Le_i, Le_j \rangle = \sigma_j^2 \delta_{ij}$$

Moreover, there exists an orthonormal basis $f_1, \dots, f_n$ of W such that the matrix of L with respect to these bases is

$$\begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_m \\ 0 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 0 \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} \sigma_1 & & & 0 & \dots & 0 \\ & \ddots & & 0 & \dots & 0 \\ & & \sigma_m & 0 & \dots & 0 \end{bmatrix}$$

If $n \geq m$ or $n \leq m$ respectively where $\sigma_j = 0$ for $j > \min\{n, m\}$.

Singular Values
Taking σ_j as in the preceding section ordered such that $\sigma_j = 0$ if $j > \min\{n, m\}$ we say $\sigma_1, \dots, \sigma_{\min\{n,m\}}$ are the singular values of L. They are the square roots of the eigenvalues of L^*L.

Theorem. Let $L : V \to W$ be a linear map between finite dimensional inner product spaces. Then

$$\|L\| = \max\{\sigma : \sigma \text{ is a singular value of } L\}$$

p-Schatten norms
For $1 \leq p \leq \infty$, the p-Schatten norm of a linear operator $L : V \to W$ with singular values $\sigma_1, \dots, \sigma_k$ is

$$\begin{aligned} \|L\|_p &= \|(\sigma_1, \dots, \sigma_k)\|_{\ell^p} \\ &= \begin{cases} \left(\sum_{j=1}^k \sigma_j^p\right)^{1/p}, & \text{if } 1 \leq p < \infty \\ \max_j \sigma_j, & \text{if } p = \infty \end{cases} \end{aligned}$$

Frobenius Norm
If $A \in M_{m\times n}(\mathbb{C})$ then the Frobenius norm, which is equivalent to the 2-Schatten norm, is

$$\|A\|_2^2 = \operatorname{tr}(A^*A) = \sum_{i=1}^m \sum_{j=1}^n |a_{ij}|^2$$

Theorem. IF $L : V \to W$ is a linear operator and $1 \leq p \leq q \leq \infty$, then

$$\|L\|_q \leq \|L\|_p$$

Functions of Matrices

Matrix Exponential
Let $A \in M_n(\mathbb{C})$. The matrix exponential is the series

$$e^A = \sum_{j=0}^{\infty} \frac{1}{j!} A^j$$

Theorem. Let $0 \in U \subset \mathbb{C}$ be an open connected set and $f : U \to \mathbb{C}$ be analytic. Suppose f has Taylor series

$$f(z) = \sum_{j=0}^{\infty} a_j z^j$$

at 0, with radius of convergence $R \in (0, \infty]$. Then, for any matrix $A \in M_n(\mathbb{C})$ with operator norm $\|A\| < R$, the series

$$f(A) = \sum_{j=0}^{\infty} a_j A^j$$

converges in operator norm.

Theorem. If $A, B \in M_n(\mathbb{C})$ with $[A, B] = 0$ then

$$e^A e^B = e^{A+B}$$

Theorem. If $A = P\Lambda P^{-1}$ for $P \in M_n(\mathbb{C})$ invertible and

$$\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$$

Then

$$e^A = P\text{diag}(e^{\lambda_1}, \dots, e^{\lambda_n})P^{-1}$$

Function of Matrix
Lt $A = PAP^{-1}$ for invertible $P \in M_n(\mathbb{C})$ and $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$. Let $f : \{\lambda_1, \dots, \lambda_n\} \to \mathbb{C}$. Then we may define

$$f(A) = P\text{diag}(f(\lambda_1), \dots, f(\lambda_n))P^{-1}$$

Theorem. Let $A \in M_n(\mathbb{C})$ and define

$$f : \mathbb{R} \to M_n(\mathbb{C})$$

by

$$f(t) = e^{tA}$$

Then $f'(t) = Ae^{tA}$ exists for all t and $x(t) = f(t)x_0$ is the solution to the ODE

$$x' = Ax$$

$$x(0) = x_0$$

Where $x_0 \in \mathbb{C}^n$.