Chapter 1 Abstract Vector Space

A *vector space* over a field K is a set V together with 2 binary operations:

$$+ : V \times V \to V \quad \text{(Addition)} \qquad \cdot : K \times V \to V \quad \text{(Scalar Multiplication)}$$

subject to the following 8 rules for all $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ and $c, d \in K$:

(+1 Addition Commutativity)	$\mathbf{u}+\mathbf{v}=\mathbf{v}+\mathbf{u}$	(+2 Addition Associativity)	$(\mathbf{u}+\mathbf{v})+\mathbf{w}=\mathbf{u}+(\mathbf{v}+\mathbf{w})$
(+3 Zero exists)	$\exists 0 \in V : \mathbf{u}+\mathbf{0}=\mathbf{u}$	(+4 Additive Inverse exists)	$\exists \mathbf{u}' \in V : \mathbf{u}+\mathbf{u}'=\mathbf{0} \quad (\mathbf{u}'=-\mathbf{u})$
(·1 Multiplication Associativity)	$(cd)\cdot\mathbf{u}=c\cdot(d\cdot\mathbf{u})$	(·2 Unity)	$1\cdot\mathbf{u}=\mathbf{u}$
(·3 Distributivity 1)	$c\cdot(\mathbf{u}+\mathbf{v})=c\cdot\mathbf{u}+c\cdot\mathbf{v}$	(·4 Distributivity 2)	$(c+d)\cdot\mathbf{u}=c\cdot\mathbf{u}+d\cdot\mathbf{u}$

Elements of a vector space V are called *vectors*.
$\mathbf{0} \in V$ is unique. For all $\mathbf{v}$, $-\mathbf{v}$ is unique. (Uniqueness) $\quad \mathbf{u}+\mathbf{v}=\mathbf{u}+\mathbf{v}' \Leftrightarrow \mathbf{v}=\mathbf{v}'$ (Cancellation Law)
$c\mathbf{u}=\mathbf{0} \Leftrightarrow c=0$ or $\mathbf{u}=\mathbf{0}$.
A subset $U \subset V$ is *subspace* if it is vector space:

(1 Zero exists)	$\mathbf{0} \in U$	(2 Closure under addition)	$\mathbf{u},\mathbf{v} \in U \to \mathbf{u}+\mathbf{v} \in U$
(3 Closure under multiplication)	$\mathbf{u} \in U, c \in K \to c\mathbf{u} \in U$		

Let $\mathcal{S} \subset V$. *Linear combination* of $\mathcal{S}$ is any $c_1\mathbf{v}_1+\cdots+c_n\mathbf{v}_n \in V$ with $c_1,\cdots,c_n \in K, \mathbf{v}_1,\cdots,\mathbf{v}_n \in \mathcal{S}$. Set of all l.c. is $\text{Span}(\mathcal{S}) \in V$
If $\mathcal{S}=\emptyset$, $\text{Span}(\mathcal{S})=\{\mathbf{0}\}$ and $\text{Span}(V)=V$
(Subspace Criterion) $U \subset V$ is subspace of V iff U is non-empty and $c \in K, \mathbf{u},\mathbf{v} \in U \Rightarrow c\mathbf{u}+\mathbf{v} \in U$.
$\{\mathbf{v}_1,\cdots,\mathbf{v}_n\}$ is *linearly dependent* if $c_1\mathbf{v}_1+\cdots+c_n\mathbf{v}_n=\mathbf{0}$ for some $c_i \in K$ possibly not all zero.
$\{\mathbf{v}_1,\cdots,\mathbf{v}_n\}$ is *linearly independent* if $c_1\mathbf{v}_1+\cdots+c_n\mathbf{v}_n=\mathbf{0}$ implies all $c_i=0$.
Ordered set $\mathcal{B} \subset V$ is *basis* for V iff $\mathcal{B}$ is l.i. and $V=\text{Span}(\mathcal{B})$
(Spanning Set Theorem) Let $\mathcal{S}=\{\mathbf{v}_1,\cdots,\mathbf{v}_n\} \subset V$ and $U=\text{Span}(\mathcal{S})$.
$U=\text{Span}(\mathcal{S}\setminus\{\mathbf{v}_k\})$ if $\mathbf{v}_k$ is l.c. of other vectors. If $U \neq \{\mathbf{0}\}$, some subset of $\mathcal{S}$ is basis of U.
(Unique Representation Theorem) Let $\mathcal{B}=\{\mathbf{b}_1,\cdots,\mathbf{b}_n\} \subset V$. Any $\mathbf{v} \in V$ is an unique l.c. of $\mathcal{B}$ iff $\mathcal{B}$ is basis of V.
$[\mathbf{v}]_{\mathcal{B}}=\begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} \in K^n$ is *coordinate vector* of $\mathbf{v}$ relative to $\mathcal{B}$ and $c_1,\cdots,c_n$ are *coordinates* of $\mathbf{v}$ relative to $\mathcal{B}$.
(Replacement Theorem) If $V=\text{Span}(\mathbf{v}_1,\cdots,\mathbf{v}_n)$ and $\mathcal{S}=\{\mathbf{u}_1,\cdots,\mathbf{u}_m\} \subset V$ is l.i. subset, then $|\mathcal{S}|=m \leq n$
All basis of V has same number of elements.
Dimension of V is $\dim V=|\mathcal{B}|=n$. $\dim\{\mathbf{0}\}=0$. V is *finite dimensional* if $\dim V<\infty$. V is *infinite dimensional* if $\dim V=\infty$.
(Basis Extension Theorem) Let $\dim V<\infty$ and $U \subset V$ be subspace. L.i. $\mathcal{S} \subset U$ can be extended to basis of V and $\dim U \leq \dim V$.
If U is subspace of V and $\dim U=\dim V$, then $U=V$.
(Basis Criterion) Let $\dim V=n \geq 1$ and $\mathcal{S} \subset V$ has n elements. If $\mathcal{S}$ is l.i. or $\mathcal{S}$ spans V, then $\mathcal{S}$ is basis of V.
Let $U, W \subset V$ be subspaces. $U+W=\{\mathbf{u}+\mathbf{w} : \mathbf{u} \in U, \mathbf{w} \in W\}$.
$V=U \oplus W$ is *direct sum* of U and W iff $V=U+W$ and $U \cap W=\{\mathbf{0}\}$.
(Uniqueness of direct sum) $V=U \oplus W$ iff every $\mathbf{v} \in V$ can be expressed uniquely as $\mathbf{v}=\mathbf{u}+\mathbf{w}$, $\mathbf{u} \in U, \mathbf{w} \in W$.
(Direct Sum Complement) Assume $\dim V<\infty$ and let $U \subset V$ be subspace.
There exist subspace $W \subset V$ called *direct sum complement* of U such that $V=U \oplus W$.
(Dimension formula) $\dim U+\dim W=\dim(U+W)+\dim(U \cap W)$ and $\dim U+\dim W=\dim(U \oplus W)$
Let U, W be vector spaces. *Product* of U, W is the set $U \times W=\{(\mathbf{u},\mathbf{w}) : \mathbf{u} \in U, \mathbf{w} \in W\}$ with operations:

$$(\mathbf{u},\mathbf{w})+(\mathbf{u}',\mathbf{w}')=(\mathbf{u}+\mathbf{u}',\mathbf{w}+\mathbf{w}') \qquad c\cdot(\mathbf{u},\mathbf{w})=(c\mathbf{u},c\mathbf{w})$$

which makes $U \times W$ a vector space with $(\mathbf{0},\mathbf{0}) \in U \times W$.
$U \times W=\{(\mathbf{u},\mathbf{0}) : \mathbf{u} \in U\} \oplus \{(\mathbf{0},\mathbf{w}) : \mathbf{w} \in W\}$ and $\dim(U \times W)=\dim U+\dim W$
Let $U_1,\cdots,U_n$ be subspace of V. If $V=U_1+\cdots+U_n$ and $U_j \cap \sum_{i \neq j} U_i=\{\mathbf{0}\}$ for all j. Then $V=U_1 \oplus \cdots \oplus U_n$.
(Uniqueness of direct sum) $V=U_1 \oplus \cdots \oplus U_n$ iff every $\mathbf{v} \in V$ can be uniquely written as $\mathbf{v}=\mathbf{u}_1+\cdots+\mathbf{u}_n$ where $\mathbf{u}_i \in U_i$

Chapter 2 Linear Transformations and Matrices

Linear Transformation $T : V \to W$ is a map such that $T(c\mathbf{u}+\mathbf{v}) = cT(\mathbf{u}) + T(\mathbf{v})$ for all $\mathbf{u}, \mathbf{v} \in V, c \in K$.
V is *domain* and W is *target*. Set of all l.t. $T : V \to W$ is $\mathscr{L}(V, W)$.
Let $S, T \in (V, W)$, $\mathbf{v} \in V$ and $c \in K$. $(S + T)(\mathbf{v}) = S\mathbf{v} + T\mathbf{u} \in W$ and $(cT)(\mathbf{v}) = cT(\mathbf{v}) \in W$. This makes (V, W) a vector space.
Any $T \in (V, W)$ is uniquely determined by image of any basis $\mathcal{B}$ of V.
Zero map $O : V \to W$ is given by $O(\mathbf{v}) = \mathbf{0}$. Identity map $I : V \to V$ is given by $I(\mathbf{v}) = \mathbf{v}$.
$T \in \mathscr{L}(K^n, K^m)$ is represented by $m \times n$ matrix $\mathbf{A}$: $\begin{pmatrix}\mathbf{w}_1 & \dots & \mathbf{w}_n\end{pmatrix}$. i-th column $\mathbf{w}_i$ is image $T(\mathbf{e}_i)$ of standard basis vector $\mathbf{e}_i \in K^n$.
Kernel (null space) $\mathrm{Ker}(T) = \{\mathbf{v} \in V : T(\mathbf{v}) = \mathbf{0}\}$ is subspace of V.
Image (range) $\mathrm{Im}(T) = \{\mathbf{w} \in W : \mathbf{w} = T(\mathbf{v})$ for some $\mathbf{v} \in V\}$ is subspace of W.
If $S \in \mathscr{L}(U, V)$ and $T \in \mathscr{L}(V, W)$, then the *composition* $T \circ S \in \mathscr{L}(U, W)$ is linear. $T^2 = T \circ T$.
T is *one-to-one* (injective) if $T(\mathbf{u}) = T(\mathbf{v}) \to \mathbf{u} = \mathbf{v}$. T is *onto* (surjective) if every $\mathbf{w} \in W$ exists $\mathbf{v} \in V$ such that $T(\mathbf{v}) = \mathbf{w}$.
T is *isomorphism* if both one-to-one and onto. If $T \in \mathscr{L}(V, W)$ exists, V is *isomorphic* to W ($V \simeq W$).
(Proposition 2.10) T is onto $\Leftrightarrow T(V) = W \Leftrightarrow T$ maps spanning set to spanning set.
(Proposition 2.10) T is one-to-one $\Leftrightarrow \mathrm{Ker}(T) = \{\mathbf{0}\} \Leftrightarrow T$ maps from l.i. set to l.i. set $\Leftrightarrow \dim T(U) = \dim U$ for any subspace $U \subset V$.
(Proposition 2.11) T is isomorphism $\to \dim V = \dim W$ and unique $T^{-1} \in \mathscr{L}(W, V)$ exists such that $T^{-1} \circ T = I_V$ and $T \circ T^{-1} = T_W$.
(Proposition 2.12) If S, T are isomorphism, then $S \circ T$ is isomorphism and $(S \circ T)^{-1} = T^{-1} \circ S^{-1}$.
(Theorem 2.13) If $\mathcal{B}$ is basis of V, then coordinate mapping is isomorphism $V \simeq K^n$.

$$\psi_{V,\mathcal{B}} : V \to K^n \qquad\qquad \mathbf{v} \mapsto [\mathbf{v}]_{\mathcal{B}}$$

$[T]^{\mathcal{B}}_{\mathcal{B}'} \in \mathscr{L}(K^n, K^m)$ is matrix representing T with respect to bases $\mathcal{B}$ and $\mathcal{B}'$. $[T]_{\mathcal{B}}$ if $V = W$ and $\mathcal{B} = \mathcal{B}'$.
If $\mathbf{A}$ is $m \times n$ matrix and $\mathbf{B}$ is $n \times r$ matrix, $(\mathbf{AB})_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj}$.
Trace $\mathrm{Tr}(\mathbf{A}) = \sum_{i=1}^{n} a_{ii}$ and $\mathrm{Tr}(\mathbf{AB}) = \mathrm{Tr}(\mathbf{BA})$ (Proposition 2.16)
(Proposition 2.17) $\mathbf{AB} = \begin{pmatrix}\mathbf{Ab}_1 & \cdots & \mathbf{Ab}_r\end{pmatrix}$.
(Proposition 2.18) If $\mathbf{A}$ represent $T \in \mathscr{L}(K^n, K^m)$ and $\mathbf{B}$ represent $S \in \mathscr{L}(K^r, K^n)$, then $\mathbf{AB}$ represent $T \circ S$.
(Proposition 2.19) For $S \in \mathscr{L}(U, V)$ and $T \in \mathscr{L}(V, W)$ with basis $\mathcal{B}, \mathcal{B}', \mathcal{B}''$ of U, V, W respectively, $[T]^{\mathcal{B}'}_{\mathcal{B}''}$ and $[T \circ S]^{\mathcal{B}}_{\mathcal{B}''} = [T]^{\mathcal{B}'}_{\mathcal{B}''} [S]^{\mathcal{B}}_{\mathcal{B}'}$
Kernal of T is *null space* $\mathrm{Nul}\,\mathbf{A}$ which is set of all solutions to $\mathbf{Ax} = \mathbf{0}$.
Image of T is *column space* $\mathrm{Col}\,\mathbf{A}$ which is set of l.c. of columns of $\mathbf{A}$.
Row operations (similar to column operations): $\mathbf{r}_i \leftrightarrow \mathbf{r}_j \qquad \mathbf{r}_i \mapsto \mathbf{r}_i + c\mathbf{r}_j \qquad \mathbf{r}_i \mapsto c\mathbf{r}_i$
(Proposition 2.22) If $\mathbf{A}'$ is $\mathbf{A}$ after row operations, $\mathrm{Nul}\,\mathbf{A}' = \mathrm{Nul}\,\mathbf{A}$. If $\mathbf{A}''$ is $\mathbf{A}$ after column operations, $\mathrm{Col}\,\mathbf{A}'' = \mathrm{Col}\,\mathbf{A}$.
Row echelon form (REF) is matrix with all zero rows at bottom and pivots equals 1 strictly right of pivot of rows about it.
Reduced row echelon form (RREF) is if each column containing pivot has zeros in entries other than pivot.
(Proposition 2.24) Row operations do not change linear dependency of column vectors.
(Corollary 2.25) RREF of matrix $\mathbf{A}$ is unique and uniquely determines spanning set of $\mathrm{Nul}\,\mathbf{A}$ that is RCEF.
Transpose $\mathbf{A}^T$ is $n \times m$ matrix with entries $(\mathbf{A}^T)_{ji} = (\mathbf{A})_{ij}$. $(\mathbf{AB})^T = \mathbf{B}^T\mathbf{A}^T$.
Rank of $\mathbf{A} = \dim \mathrm{Col}\,\mathbf{A}$. Row rank of $\mathbf{A} = \dim \mathrm{Col}\,\mathbf{A}^T$. Nullity of $\mathbf{A} = \dim \mathrm{Nul}\,\mathbf{A}$. Rank of $\mathbf{A}$ = Rank of $\mathbf{A}^T$ (Theorem 2.29)
(Rank-Nullity Theorem) Let $\dim V < \infty$ and $T \in \mathscr{L}(V, W)$. Then $\dim \mathrm{Im}(T) + \dim \mathrm{Ker}(T) = \dim V$.
Square matrix $\mathbf{A}$ is *invertible* if it represents isomorphism T. $\mathbf{A}^{-1}\mathbf{A} = \mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$.
(Invertible Matrix Theorem) $n \times n$ matrix $\mathbf{A}$ is invertible iff any one of statements hold:

(1)	$\mathrm{Col}\,\mathbf{A} = K^n$. ($\mathbf{A}$ is surjective)	(2)	$\mathrm{Nul}\,\mathbf{A} = \{\mathbf{0}\}$. Nullity of $\mathbf{A} = 0$.
(3)	Rank of $\mathbf{A} =$ Rank of $\mathbf{A}^T = n$.	(4)	RREF of $\mathbf{A}$ is $\mathbf{I}$.

(Change of Basis Formula) There exists $n \times n$ $\mathbf{P}^{\mathcal{B}}_{\mathcal{B}'}$ such that $[\mathbf{v}]_{\mathcal{B}'} = \mathbf{P}^{\mathcal{B}}_{\mathcal{B}'} [\mathbf{v}]_{\mathcal{B}}$ where $\mathbf{P}^{\mathcal{B}}_{\mathcal{B}'} = [I]^{\mathcal{B}}_{\mathcal{B}'}$.
$\mathbf{P}^{\mathcal{B}}_{\mathcal{B}'}$ is *change-of-coordinate matrix* from $\mathcal{B}$ to $\mathcal{B}'$. It is invertible with inverse given by $\mathbf{P}^{\mathcal{B}'}_{\mathcal{B}} = (\mathbf{P}^{\mathcal{B}}_{\mathcal{B}'})^{-1}$. (Proposition 2.34)
(Proposition 2.35) For any bases $\mathcal{B}, \mathcal{B}', \mathcal{B}''$ of V, $\mathbf{P}^{\mathcal{B}'}_{\mathcal{B}''}\mathbf{P}^{\mathcal{B}}_{\mathcal{B}'} = \mathbf{P}^{\mathcal{B}}_{\mathcal{B}''}$. $\mathbf{P}^{\mathcal{B}}_{\mathcal{B}'} = (\mathbf{P}^{\mathcal{B}'}_{\mathcal{E}})^{-1}\mathbf{P}^{\mathcal{B}}_{\mathcal{E}}$.
(Theorem 2.36) Let $T \in \mathscr{L}(V, V)$ and $[T]_{\mathcal{B}} = \mathbf{A}, [T]_{\mathcal{B}'} = \mathbf{B}, \mathbf{P} = \mathbf{P}^{\mathcal{B}}_{\mathcal{B}'}$. Then $\mathbf{B} = \mathbf{PAP}^{-1}$ and $\mathbf{A}$ is similar to $\mathbf{B}$. ($\mathbf{A} \sim \mathbf{B}$)
(Proposition 5.13) If $\mathbf{A} = \mathbf{PBP}^{-1}$, $q(\mathbf{A}) = \mathbf{P}q(\mathbf{B})\mathbf{P}^{-1}$ for any polynomial $q(t)$.
(Proposition 2.38) Two polynomials in $K_n[t]$ are same if they agree on $n + 1$ distinct points.
Given $n + 1$ distinct points $t_0, \cdots, t_n \in K$ ($t_i \neq t_j$ for $i \neq j$), the evaluation map is

$$T : K_n[t] \to K^{n+1} \qquad\qquad p(t) \mapsto \begin{pmatrix}p(t_0) & \dots & p(t_n)\end{pmatrix}^T$$

(Proposition 2.40) Choosing $\mathcal{E} = \{1, t, \cdots, t^n\}$ of $K_n[t]$ and $\mathcal{E}' = \{\mathbf{e}_0, \cdots, \mathbf{e}_n\}$ of K^{n+1}.
T is represented by Vandermonde matrix: $[T]^{\mathcal{E}}_{\mathcal{E}'} \begin{pmatrix}1 & \dots & t_0^n \\ \vdots & \ddots & \vdots \\ 1 & \dots & t_n^n\end{pmatrix}$ and is invertible,
(Proposition 2.41) For $k = 0, \cdots, n$, image of polynomial (of degree n) $p_k(t) = \prod_{\substack{0 \le j \le n \\ j \neq k}} \frac{t - t_j}{t_k - t_j}$ under T is $\mathbf{e}_k \in K^{n+1}$.
Image of polynomial $p(t) = \sum_{i=0}^{n} \lambda_i p_i(t)$ under T is *Lagrange interpolation polynomial* $\begin{pmatrix}\lambda_0 & \dots & \lambda_n\end{pmatrix}^T \in K^{n+1}$.

Chapter 3 Determinants

Determinant is a function $\det : M_{n\times n}(K) \to K$ which satisfies:

1. Multilinear: $\det\begin{pmatrix}\dots & \mathbf{w} & c\mathbf{u}+\mathbf{v} & \dots\end{pmatrix} = c\det\begin{pmatrix}\dots & \mathbf{w} & \mathbf{u} & \dots\end{pmatrix} + \det\begin{pmatrix}\dots & \mathbf{w} & \mathbf{v} & \dots\end{pmatrix}$
2. Alternating: $\det\begin{pmatrix}\dots & \mathbf{w} & \mathbf{u} & \mathbf{v} & \dots\end{pmatrix} = -\det\begin{pmatrix}\dots & \mathbf{w} & \mathbf{v} & \mathbf{u} & \dots\end{pmatrix}$
3. $\det(\mathbf{I}) = 1$

(Proposition 3.2) Determinant of triangular (or diagonal) matrix is product of its diagonal entries.
(Leibniz Expansion / Expansion by Permutations) $\det \mathbf{A} = \sum_{\sigma\in\mathcal{S}_n} \epsilon(\sigma) a_{\sigma(1)1}\cdots a_{\sigma(n)n}$ where $\mathcal{S}_n$: set of all permutations of $\{1,\cdots,n\}$
Sign of permutation $\epsilon(\sigma)$: $+1$ if $\sigma\in\mathcal{S}_n$ can be obtained by even number of transpositions, -1 if odd number of transpositions.
Any function D satisfying (1),(2) of definition of determinant is scalar multiple $c\cdot\det$ where $c = D(\mathbf{I})$.
$\det(\mathbf{AB}) = \det(\mathbf{A})\det(\mathbf{B})$. $\mathbf{A}$ is invertible iff $\det(\mathbf{A}) \neq 0$. $\det(\mathbf{A}^{-1}) = \det(\mathbf{A})^{-1}$. $\mathbf{A}\sim\mathbf{B} \longrightarrow \det(\mathbf{A}) = \det(\mathbf{B})$.
If $\dim V < \infty$, determinant of $T \in \mathscr{L}(V,V)$ is $\det(T) = \det(\mathbf{A})$ for any square matrix $\mathbf{A}$ representing T.
$\det\left(\begin{array}{c|c}\mathbf{A}_{k\times k} & \mathbf{B}_{k\times l}\\ \hline \mathbf{O}_{l\times k} & \mathbf{C}_{l\times l}\end{array}\right) = \det\left(\begin{array}{c|c}\mathbf{A}_{k\times k} & \mathbf{O}_{k\times l}\\ \hline \mathbf{B}_{l\times k} & \mathbf{C}_{l\times l}\end{array}\right) = \det(\mathbf{A}_{k\times k})\det(\mathbf{C}_{l\times l})$.
(Proposition 3.9) Column operations are same as multiplying on the right by elementary matrices (1 in diagonal, 0 in others)

1. $\mathbf{E} = \begin{pmatrix}\ddots & & & \\ & 1 & c & \\ & 0 & 1 & \\ & & & \ddots\end{pmatrix}$: Adding multiple of i-th column to j-th column. $\det(\mathbf{E}) = 1$.
2. $\mathbf{E} = \begin{pmatrix}\ddots & & \\ & c & \\ & & \ddots\end{pmatrix}$: Scalar multiplying i-th column bu c. $\det(\mathbf{E}) = c$.
3. $\mathbf{E} = \begin{pmatrix}\ddots & & & \\ & 0 & 1 & \\ & 1 & 0 & \\ & & & \ddots\end{pmatrix}$: Interchanging two columns. $\det(\mathbf{E}) = -1$.

(Proposition 3.10) Row operations are same as multiplying on the left by elementary matrices.
(Theorem 3.11) For any square matrix $\mathbf{A}$, $\overline{\det(\mathbf{A})} = \det(\mathbf{A}^*)$.
(Theorem 3.12) $\det(\mathbf{A})$ is *n-dimensional signed volume* of parallelepiped $\mathbf{P} = \{c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n : 0 \le c_i \le 1\} \subset \mathbb{R}^n$.
It is spanned by columns of $\mathbf{A}$. *Volume* $= |\det(\mathbf{A})|$
Let $\mathcal{B} = \{\mathbf{v}_1,\cdots,\mathbf{v}_n\} \subset \mathbb{R}^n$ be a basis. $\mathcal{B}$ is *positively oriented* if $\det\begin{pmatrix}\mathbf{v}_1 & \dots & \mathbf{v}_n\end{pmatrix} > 0$. Otherwise, $\mathcal{B}$ is *negatively oriented.*
(Laplace Expansion / Expansion by Cofactors) For any i, j,
$C_{ij} = (-1)^{i+j}\det(\mathbf{A}_{ij})$ are *cofactors.* $\mathbf{A}_{ij}$ is submatrix from $\mathbf{A}$ by deleting i-th row and j-th column.

1. Expansion by rows: $\det(\mathbf{A}) = a_{i1}C_{i1} + \cdots + a_{in}C_{in}$
2. Expansion by columns: $\det(\mathbf{A}) = a_{1j}C_{1j} + \cdots + a_{nj}C_{nj}$

(Cramer's Rule) If $\mathbf{A}$ is invertible, solution to $\mathbf{Ax} = \mathbf{b}$ is given by $\mathbf{x} = \begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}$ where $x_i = \frac{\det(\mathbf{A}_i)}{\det(\mathbf{A})}$.
$\mathbf{A}_i$ is obtained by replacing i-th column of $\mathbf{A}$ with $\mathbf{b}$.
(Theorem 3.16) If $\mathbf{A}$ is invertible, cofactor matrix $\mathbf{C} = (C_{ij})$ and *adjugate matrix* $\operatorname{adj}(\mathbf{A}) = \mathbf{C}^T$. $\mathbf{A}^{-1} = \frac{\operatorname{adj}(\mathbf{A})}{\det(\mathbf{A})}$.

Chapter 4 Inner Product Spaces

Real dot product of $\mathbf{u} = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}, \mathbf{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \in \mathbb{C}^n$ is defined to be $\mathbf{u} \cdot \mathbf{v} = \mathbf{v}{*}\mathbf{u} = \overline{\mathbf{u}^*\mathbf{v}} = \sum_{i=1}^n u_i \overline{v_i} \in \mathbb{C}$.
(Proposition 4.2) For any $\mathbf{A} = (a_{ij}) \in M_{m\times n}(\mathbb{R})$, then $a_{ij} = \mathbf{e}_i' \cdot \mathbf{A}\mathbf{e}_j$.
$\{\mathbf{e}_1, \cdots, \mathbf{e}_n\}$ is standard basis for $\mathbb{R}^n$ and $\{\mathbf{e}_1', \cdots, \mathbf{e}_m'\}$ is standard basis for $\mathbb{R}^m$.
If $\mathbf{A} = \begin{pmatrix} \mathbf{a}_1^T \\ \vdots \\ \mathbf{a}_m^T \end{pmatrix}$ is $m \times k$ real matrix and $\mathbf{B} = \begin{pmatrix} \mathbf{b}_1 & \ldots & \mathbf{b}_n \end{pmatrix}$ is $k \times n$ real matrix, then $\mathbf{AB} = \begin{pmatrix} \mathbf{a}_1 \cdot \mathbf{b}_1 & \ldots & \mathbf{a}_1 \cdot \mathbf{b}_n \\ \vdots & \ddots & \vdots \\ \mathbf{a}_n \cdot \mathbf{b}_1 & \ldots & \mathbf{a}_n \cdot \mathbf{b}_n \end{pmatrix}$
Inner product is binary operator $\langle \ , \ \rangle : V \times V \to \mathbb{R}(\mathbb{C})$ satisfying for any $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ and $c \in \mathbb{R}(\mathbb{C})$:

1. Commutativity: $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle \, (\overline{\langle \mathbf{v}, \mathbf{u} \rangle})$
2. Linearity: $\langle c\mathbf{u} + \mathbf{v}, \mathbf{w} \rangle = c \langle \mathbf{u}, \mathbf{w} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle$
3. Positivity: $\langle \mathbf{u}, \mathbf{u} \rangle \geq 0$ and $\langle \mathbf{u}, \mathbf{u} \rangle = 0 \Leftrightarrow \mathbf{u} = \mathbf{0}$

Euclidean space is $V = \mathbb{R}^n$ equipped with dot product as inner product.
Norm (length) of $\mathbf{v} \in V$ is non-negative scalar $\|\mathbf{v}\| = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle} \in \mathbb{R}_{\geq 0}$.
Unit vector is vector $\mathbf{u} \in V$ with $\|\mathbf{u}\| = 1$. Given $\mathbf{0} \neq \mathbf{v} \in V$, vector $\frac{\mathbf{v}}{\|\mathbf{v}\|}$ has unit length and is the *normalization* of $\mathbf{v}$.
Distance between $\mathbf{u}, \mathbf{v} \in V$ $\text{dist}(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\|$.
(Law of Cosine) The angle $0 \leq \theta \leq \pi$ between non-zero $\mathbf{u}, \mathbf{v} \in V$ is defined by $\langle \mathbf{u}, \mathbf{v} \rangle = \|\mathbf{u}\|\|\mathbf{v}\| \cos\theta$.
Two vectors $\mathbf{u}, \mathbf{v} \in V$ are *orthogonal* (perpendicular) to each other if $\langle \mathbf{u}, \mathbf{v} \rangle = 0$.
(Pythagorean Theorem) Let $\mathbf{u}, \mathbf{v} \in V$, If $\langle \mathbf{u}, \mathbf{v} \rangle = 0$, then $\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$.
(Cauchy-Schwarz Inequality) For all $\mathbf{u}, \mathbf{v} \in V$, $|\langle \mathbf{u}, \mathbf{v} \rangle| \leq \|\mathbf{u}\|\|\mathbf{v}\|$.
(Triangle Inequality) $\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$.
Let $\mathcal{S} = \{\mathbf{u}_1, \cdots, \mathbf{u}_r\} \subset V$ be finite set. $\mathcal{S}$ is *orthogonal set* if $\langle \mathbf{u}_i, \mathbf{u}_j \rangle = 0$ for all $i \neq j$.
If in addition $\mathcal{S}$ is basis of V, $\mathcal{S}$ is *orthogonal basis* for V. If in addition all vectors in $\mathcal{S}$ has unit norm, $\mathcal{S}$ is *orthonormal basis* for V.
Standard basis $\{\mathbf{e}_1, \cdots, \mathbf{e}_n\}$ for $\mathbb{R}^n$ is orthonormal basis $\mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$ where δ_{ij} is Kronecker delta.
(Proposition 4.13) Let $\mathcal{B} = \{\mathbf{b}_1, \cdots, \mathbf{b}_n\}$ be o.g. basis for V. Then coordinate mapping with respect to $\mathcal{B}$ is

$$\psi_{V,\mathcal{B}} : V \to \mathbb{R}^n \qquad \mathbf{v} \mapsto [\mathbf{v}]_{\mathcal{B}} = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} \qquad \text{where } c_i = \frac{\langle \mathbf{v}, \mathbf{b}_i \rangle}{\langle \mathbf{b}_i, \mathbf{b}_i \rangle}, \qquad i = 1, \cdots, n$$

(Corollary 4.14) For $\mathbf{A} = (a_{ij}) = [T]_{\mathcal{B}'}^{\mathcal{B}}$ with respect to basis $\mathcal{B} = \{\mathbf{v}_1, \cdots, \mathbf{v}_n\} \subset V$ and o.n. basis $\mathcal{B}' = \{\mathbf{w}_1, \cdots, \mathbf{w}_n\} \subset W$, then $a_{ij} = \langle \mathbf{w}, T\mathbf{v}_j \rangle$.
Orthogonal complement of U is subset $U^\perp = \{\mathbf{v} \in V : \langle \mathbf{v}, \mathbf{u} \rangle = 0 \text{ for any } \mathbf{u} \in U\}$.
(Proposition 4.16) $U^\perp \subset V$ is subspace. $U \subset (U^\perp)^\perp$. $\mathbf{u} \in U^\perp$ iff $\langle \mathbf{u}, \mathbf{v} \rangle = 0$ for all $\mathbf{v} \in U$.
(Proposition 4.17) Let $\mathbf{A} \in M_{n\times n}(\mathbb{R})$. $(\text{Col}\,\mathbf{A}^T)^\perp = \text{Nul}\,\mathbf{A}$, $(\text{Col}\,\mathbf{A})^\perp = \text{Nul}\,\mathbf{A}^T$ with respect to standard dot product on $\mathbb{R}^n$ and $\mathbb{R}^m$.
(Proposition 4.18) O.g. projection of $\mathbf{b}$ onto $\mathbf{u}$ is given by $\text{Proj}_{\mathbf{u}}(\mathbf{b}) = \langle \mathbf{b}, \mathbf{e} \rangle \mathbf{e} = \frac{\langle \mathbf{b}, \mathbf{u} \rangle}{\langle \mathbf{u}, \mathbf{u} \rangle} \mathbf{u}$ where $\mathbf{e} = \frac{\mathbf{u}}{\|\mathbf{u}\|}$ is normalization of $\mathbf{u}$.
(Gram-Schmidt Process) Let $\dim V = n$ and $\{\mathbf{x}_1, \cdots, \mathbf{x}_n\}$ be basis for V.

$$\mathbf{u}_1 = \mathbf{x}_1 \qquad \mathbf{u}_2 = \mathbf{x}_2 - \text{Proj}_{\mathbf{u}_1}(\mathbf{x}_2) \qquad \mathbf{u}_k = \mathbf{x}_k - \left(\sum_{i=0}^{k-1} \text{Proj}_{\mathbf{u}_i}(\mathbf{x}_k) \right)$$

Then $\mathcal{B} = \{\mathbf{u}_1, \cdots, \mathbf{u}_n\}$ is o.g. basis for V. For all $1 \leq k \leq n$, $\text{Span}(\mathbf{u}_1, \cdots, \mathbf{u}_k) = \text{Span}(\mathbf{x}_1, \cdots, \mathbf{x}_k)$
(Corollary 4.20) Any finite dimensional inner product space has an o.n. basis.
(Orthogonal Decomposition Theorem) Let $\dim V < \infty$ and $U \subset V$ be subspace.
Each $\mathbf{v} \in V$ can be written uniquely in form $\mathbf{v} = \mathbf{v}_{||} + \mathbf{v}_\perp$ where $\mathbf{v}_{||} \in U$ and $\mathbf{v}_\perp \in U^\perp$. Therefore, $V = U \oplus U^\perp$ and $U \cap U^\perp = \{\mathbf{0}\}$.
(Corollary 4.22) If $\dim V < \infty$ and $U \subset V$ be a subspace. $(U^\perp)^\perp = U$. (Not true if $\dim V = \infty$)
(Proposition 4.23) If $\{\mathbf{u}_1, \cdots, \mathbf{u}_r\}$ is orthonormal basis for $U \subset \mathbb{R}^n$ with respect to dot product, $\text{Proj}_U(\mathbf{x}) = (\mathbf{x}\cdot\mathbf{u}_1)\mathbf{u}_1 + \cdots + (\mathbf{x}\cdot\mathbf{u}_r)\mathbf{u}_r$.
If $\mathbf{P} = \begin{pmatrix} \mathbf{u}_1 & \ldots & \mathbf{u}_r \end{pmatrix}$ is $n \times r$ matrix, then $\text{Proj}_U(\mathbf{x}) = \mathbf{P}\mathbf{P}^*\mathbf{x}$
Projection matrix is $n \times n$ matrix $\mathbf{M}$ such that $\mathbf{M}^2 = \mathbf{M}$. It is *orthogonal projection matrix* if in addition $\mathbf{M}^* = \mathbf{M}$.
(Best Approximation Theorem) Let U be subspace of $\mathbb{R}^n$ and $\mathbf{x} \in \mathbb{R}^n$. Then $\|\mathbf{x} - \text{Proj}_U(\mathbf{x})\| \leq \|\mathbf{x} - \mathbf{u}\|$ for any $\mathbf{u} \in U$.
$\text{Proj}_U(\mathbf{x}) \in U$ is closest point in U to $\mathbf{x}$.
(Riesz Representation Theorem) If $\dim V < \infty$ and $T \in \mathscr{L}(V, \mathbb{R})$, then there exists **unique** vector $\mathbf{u} \in V$ such that $T(\mathbf{v}) = \langle \mathbf{u}, \mathbf{v} \rangle$.

Chapter 5 Spectral Theory

$T \in \mathscr{L}(V,V)$ is l.i. and is represented by $\mathbf{A} \in M_{n\times n}(K)$ if $\dim V < \infty$. $\mathbf{D}$ is diagonal matrix consisting of corresponding eigenvalues.

Eigenvector of T is non-zero vector $\mathbf{u} \in V$ such that $T\mathbf{u} = \lambda\mathbf{u}$ for some scalar $\lambda \in K$. λ is *eigenvalue*.
Space $V_\lambda = \{\mathbf{u} : T\mathbf{u} = \lambda\mathbf{u}\} \subset V$ is *eigenspace* of λ. If finite set of eigenvectors form a basis of V, it is *eigenbasis*.
(Proposition 5.2) $V_\lambda = \mathrm{Ker}(T - \lambda I)$ is non-zero subspace of V. ($\mathbf{0} \in V_\lambda$)
(Proposition 5.3) If $\dim V < \infty$, λ is eigenvalue of T iff $\det(T - \lambda I) = 0$.
(Proposition 5.4) λ of triangular matrix are given by entries on main diagonal.
(Theorem 5.5) Eigenvector set $\{\mathbf{v}_1, \cdots, \mathbf{v}_r\}$ is l.i. if it corresponds to *distinct* $\lambda_1, \cdots, \lambda_r$.
$p(\lambda) = \det(T - \lambda I)$ is *characteristic polynomial* over K of degree n. Eigenvalues are roots of $p(\lambda) = 0$.
(Proposition 5.7) For $p(\lambda)$, top term is λ^n with coefficient $(-1)^n$, coefficient of λ^{n-1} $\mathrm{Tr}(\mathbf{A})$, constant term is $\det(\mathbf{A})$.
Geometric multiplicity of λ_i is $\dim V_{\lambda_i}$. *Algebraic multiplicity* of λ_i is number of factors $(\lambda - \lambda_i)$ in $p(\lambda)$.
(Proposition 5.9) If $K = \mathbb{C}$, a.m. of all eigenvalues add up to $\dim_{\mathbb{C}} V = n$. (Every complex $\mathbf{A}$ has n eigenvalues)
(Corollary 5.10) $\mathrm{Tr}(\mathbf{A})$ is sum of all complex eigenvalues. $\det(\mathbf{A})$ is product of all complex eigenvalues.
(Proposition 5.11) If $\mathbf{A} \sim \mathbf{B}$, they have same determinant, trace, charpoly, eigenvalues, a.m. and g.m..
$\mathbf{A}$ is *diagonalizable* iff $\mathbf{A} \sim \mathbf{D}$ ($\mathbf{A} = \mathbf{P}\mathbf{D}\mathbf{P}^{-1}$)
(Diagonalization Theorem) $\mathbf{A}$ is diagonalizable iff $\mathbf{A}$ has n linearly independent eigenvectors.
Columns of $\mathbf{P}$ are eigenvectors of $\mathbf{A}$ and $\mathbf{D}$ is diagonal matrix consisting of corresponding eigenvalues.
T is diagonalizable of V has an eigenbasis of T.
(Corollary 5.16) T is diagonalizable iff $V = V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_r}$ with **distinct** $\lambda_1, \cdots, \lambda_r \in K$.
(Corollary 5.17) T has n different eigenvalues $\longrightarrow$ diagonalizable.
(Corollary 5.18) T is diagonalizable $\longrightarrow$ a.m. $=$ g.m. for λ_i. a.m. $=$ g.m. for λ_i and a.m. add up to $\dim V = n \longrightarrow T$ is diagonalizable.
$\mathbf{A}$ is *symmetric* if $\mathbf{A}^T = \mathbf{A}$.
(Proposition 5.20) $\mathbf{A}$ is symmetric iff $\mathbf{u} \cdot \mathbf{A}\mathbf{v} = \mathbf{A}\mathbf{u} \cdot \mathbf{v}$.
(Theorem 5.21) If $\mathbf{A}$ is symmetric, then different eigenspaces are orthogonal to each other.
Matrix $\mathbf{A}$ is *orthogonally diagonalizable* if $\mathbf{A} = \mathbf{P}\mathbf{D}\mathbf{P}^{-1} = \mathbf{P}\mathbf{D}\mathbf{P}^T$ for some o.g. matrix $\mathbf{P}$ and $\mathbf{D}$.
(Theorem 5.23) If $\mathbf{A}$ is a real symmetric matrix, then all eigenvalues are real. (charpoly has n real roots)
(Theorem 5.24) $\mathbf{A}$ is symmetric iff it is o.g. diagonalizable.

Chapter 6 Positive Definite Matrices and SVD

Assume $\dim V < \infty$.
Bilinear form on V is a real-valued function $f(\mathbf{x}, \mathbf{y})$ in two variables that is linear in both arguments $\mathbf{x}, \mathbf{y} \in V$.
Any bilinear form on $\mathbb{R}^n$ is of form $f(\mathbf{x}, \mathbf{y}) = \mathbf{x}^*\mathbf{A}\mathbf{y}$ for some matrix $\mathbf{A}$.
Let $\mathbf{A}$ be real symmetric. It is *positive (semi)definite* if $\mathbf{x}^*\mathbf{A}\mathbf{x} = \mathbf{A}\mathbf{x} \cdot \mathbf{x} > (\geq) 0$ for all non-zero $\mathbf{x} \in \mathbb{R}^n$.
Any inner product on $\mathbb{R}^n$ is of form $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^*\mathbf{A}\mathbf{y}$ for some p.d. matrix $\mathbf{A}$.
Symmetric matrix $\mathbf{A}$ is +ve.(semi)d. iff $\lambda_{\mathbf{i}} > (\geq) 0$ for all i. ($\rightarrow \det(\mathbf{A}) > (\geq) 0$)
(Theorem 6.7) Let $\mathbf{A}$ be +ve.(semi)d. matrix. There exists **unique** +ve.(semi)d. matrix $\mathbf{B}$ such that $\mathbf{B}^2 = \mathbf{A}$.
$\mathbf{B} = \sqrt{\mathbf{A}}$ is *square root* of $\mathbf{A}$. If $\mathbf{A} = \mathbf{P}\mathbf{D}\mathbf{P}^*$, $\sqrt{\mathbf{A}} = \mathbf{P}\mathbf{D}^{\frac{1}{2}}\mathbf{P}^*$. $\sqrt{\mathbf{A}} = q(\mathbf{A})$ for some polynomial $q(t)$.
If $\mathbf{B}$ commutes with p.(semi)d. $\mathbf{A}$, then it commutes with $\sqrt{\mathbf{A}}$.
Let $\mathbf{A}$ be any $m \times n$ matrix. $\mathbf{A}^*\mathbf{A}$ is +ve.(semi)d.. *Absolute value* of $\mathbf{A}$ $|\mathbf{A}| = \sqrt{\mathbf{A}^*\mathbf{A}}$.
Any +ve.(semi)d. matrix is of form $\mathbf{A}^*\mathbf{A}$ for some $\mathbf{A}$.
Singular values of $\mathbf{A}$ is eigenvalues σ_i of $|\mathbf{A}|$.
If $\mathbf{A}$ rank r, then we have r non-zero singular values arranged in descending order. ($\sigma_1 \geq \cdots \geq \sigma_r > 0$).
(Singular Value Decomposition) Let $\mathbf{A}$ be $m \times n$ matrix with rank r. $\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^*$. $\mathbf{U}, \mathbf{V}$ are $m \times m, n \times n$ o.g. matrix respectively.
$\mathbf{\Sigma}$ is $m \times n$ quasi-diagonal matrix consisting of r non-zero singular values $\sigma_1, \cdots, \sigma_r$ of $\mathbf{A}$.
For any real square matrix $\mathbf{A}$, we have polar decomposition $\mathbf{A} = \mathbf{P}\mathbf{H}$ where $\mathbf{P} = \mathbf{U}\mathbf{V}^*$ is o.g. and $\mathbf{H} = \mathbf{V}\mathbf{\Sigma}\mathbf{V}^*$ is +ve.(semi)d..
Columns $\{\mathbf{u}_1, \cdots, \mathbf{u}_r\}$ is o.n. basis of $\mathrm{Col}\,\mathbf{A}$. Columns $\{\mathbf{u}_{r+1}, \cdots, \mathbf{u}_m\}$ is o.n. basis of $\mathrm{Nul}(\mathbf{A}^*)$.
Columns $\{\mathbf{v}_1, \cdots, \mathbf{v}_r\}$ is o.n. basis of $\mathrm{Row}\,\mathbf{A} = \mathrm{Col}\,\mathbf{A}^*$. $\{\mathbf{v}_{r+1}, \cdots, \mathbf{v}_n\}$ is o.n. basis of $\mathrm{Nul}\,\mathbf{A}$.
Let $\mathbf{U}_r, \mathbf{V}_r$ be submatrix consisting of first r columns, $\mathbf{A} = \begin{pmatrix} \mathbf{U}_r & * \end{pmatrix} \begin{pmatrix} \mathbf{D} & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{pmatrix} \begin{pmatrix} \mathbf{V}_r^* \\ * \end{pmatrix} = \mathbf{U}_r\mathbf{D}\mathbf{V}_r^*$.
Pseudo-inverse of $\mathbf{A}$ is $\mathbf{A}^+ = \mathbf{V}_r\mathbf{D}^{-1}\mathbf{U}_r^*$. Given equation $\mathbf{A}\mathbf{x} = \mathbf{b}$, least square solution is given by $\hat{\mathbf{x}} = \mathbf{A}^+\mathbf{b} = \mathbf{V}_r\mathbf{D}^{-1}\mathbf{U}_r^*\mathbf{b}$.
Matrix norm $\|\mathbf{A}\| = \max_{|\mathbf{v}|=1}\|\mathbf{A}\mathbf{v}\|$.

Chapter 7 Complex Matrices

Let $T \in \mathscr{L}(V,W)$. *Adjoint* of T is l.t. $T^* \in \mathscr{L}(W,V)$ satisfying $\langle T\mathbf{v}, \mathbf{w}\rangle_W = \langle \mathbf{v}, T^*\mathbf{w}\rangle$ for all $\mathbf{v} \in V, \mathbf{w} \in W$.
If $\dim V < \infty$, then T^* exists and is unique.
If $S \in \mathscr{L}(U,V)$ and $T, T' \in \mathscr{L}(V,W)$ such that adjoints exist, $(cT + T' + T^*)^* = cT^* + T'^* + T$. $(T \circ S) = S^* \circ T^*$.
In case of $\mathbb{C}^n$ with *standard dot product*, if T is represented by $\mathbf{A}$, then T^* is represented by conjugate transpose $\mathbf{A}^* = \overline{\mathbf{A}}^T$.
$((a_{ij})^* = (\overline{a_{ji}}))$. $(\mathbf{AB})^* = \mathbf{B}^*\mathbf{A}^*$.
T is *self-adjoint* if $T = T^*$. ($\langle T\mathbf{v}, \mathbf{w}\rangle = \langle \mathbf{v}, T\mathbf{w}\rangle$ for all $\mathbf{v}, \mathbf{w} \in V$).
If T is self-adjoint, different eigenspaces are orthogonal to each other, all λ are real, and T has o.n. eigenbasis if $\dim V < \infty$.
$\mathbf{U}$ is *unitary matrix* if is square matrix and consists of o.n. columns under complex dot product. $\mathbf{U}^{-1} = \mathbf{U}^*$. $|\det(\mathbf{U})| = 1$.
Set of unitary matrices is $U(n)$. $\mathbf{I}_{n\times n} \in U(n)$. If $\mathbf{U}, \mathbf{V} \in U(n)$, $\mathbf{U}^{-1}, \mathbf{UV} \in U(n)$.
If $\lambda \in \mathbb{C}$ is eigenvalue of $\mathbf{U}$, then $|\lambda| = 1$.
$\mathbf{A}$ is *unitarily equivalent* to $\mathbf{B}$ if there exists unitary $\mathbf{U}$ such that $\mathbf{A} = \mathbf{UBU}^*$.
(Schur's Lemma) Any complex square matrix $\mathbf{A}$ is unitarily equivalent to an upper triangular matrix.
Steps:

1. Pick an eigenvector $\mathbf{v} \in \mathbb{C}^n$ with eigenvalue λ. By Gram-Schmidt Process, extend to o.n. basis $\mathcal{B} = \{\mathbf{v}, \mathbf{u}_2, \cdots, \mathbf{u}_n\}$ of $\mathbb{C}^n$.
2. Matrix $\mathbf{U}'$ with column $\mathcal{B}$ is unitary. We have block form $\mathbf{U}'^*\mathbf{A}\mathbf{U}' = \left(\begin{array}{c|c} \lambda & * \\ \hline 0 & \mathbf{A}_{n-1} \end{array}\right)$

   Repeat this process by picking another eigenvector until an upper triangular matrix $\mathbf{T}$ is obtained.

If two **real** matrices $\mathbf{A} \sim \mathbf{B}$ in $K = \mathbb{C}$, then $\mathbf{A} \sim \mathbf{B}$ in $K = \mathbb{R}$.
If $\mathbf{A} = \mathbf{UBU}^*$ for some $\mathbf{U} \in U(n)$, then $\mathbf{A} = \mathbf{PBP}^T$ for some $\mathbf{P} \in O(n)$.
Note: If $\det(\mathbf{X} + i\mathbf{Y}) \neq 0$, then $\det(\mathbf{X} + \lambda\mathbf{Y})$ is non-zero polynomial.
$\mathbf{A}$ is *unitarily diagonalizable* if it is unitarily equivalent to diagonal matrix $\mathbf{A} = \mathbf{UDU}^{-1} = \mathbf{UDU}^*$.
$\mathbf{A}$ is *Hermitian* if $\mathbf{A} = \mathbf{A}^*$.
(Spectral Theorem for Hermitian Matrices)
Hermitian $\implies$ eigenvectors from different eigenspaces are o.g. + real λ + unitarily diagonalizable.
Unitarily diagonalizable + all real $\lambda \implies \mathbf{A}$ is Hermitian.
L.i. $T \in \mathscr{L}(V,V)$ is *normal* if $TT^* = T{*}T$. ($\mathbf{AA}^* = \mathbf{A}^*\mathbf{A}$)
E.g. Hermitian matrix, unitary matrix, skew-hermitian matrix, complex symmetric and orthogonal matrices.
Let $T \in \mathscr{L}(V,V)$ be normal operator. For all $\mathbf{v} \in V$, $\|T\mathbf{v}\| = \|T^*\mathbf{v}\|$. If λ is eigenvalue of T, $\overline{\lambda}$ is eigenvalue of T^*.
Eigenvectors corresponding to different eigenvalues are orthogonal to each other.
Upper triangular matrix is normal iff it is diagonal.
Square matrix $\mathbf{A}$ is unitarily diagonalizable iff it is normal. (For $\dim V < \infty$, V has o.n. eigenbasis of $T \in \mathscr{L}(V,V)$ iff it is normal).

Chapter 8 Invariant Subspaces

Let $T \in \mathscr{L}(V,V)$, $U \subset V$ be T-invariant subspace.

A subspace U is *T-invariant subspace* if $T(U) \subset U$. If U is T-invariant, we have *restriction* $T|_U : U \longrightarrow U$.
Let $\mathbf{v} \in V$. $U_\mathbf{v} = \mathrm{Span}(\mathbf{v}, T\mathbf{v}, T^2\mathbf{v}, \cdots)$ is the *cyclic subspace* of T generated by *cyclic vector* $\mathbf{v} \in V$.
Cyclic subspace $U_\mathbf{v}$ of T is T-invariant subspace and is spanned by first r elements where $r = \dim U_\mathbf{v}$ if $\dim V < \infty$.
Assume V is direct sum of T-invariant subspaces $V = U_1 \oplus \cdots \oplus U_k$. Let $\mathbf{v} = \mathbf{u}_1 + \cdots + \mathbf{u}_k$ be unique decomposition.
Then $T(\mathbf{v}) = T_1(\mathbf{u}_1) + \cdots + T_k(\mathbf{u}_k)$ where $T_i = T|_{U_i}$. We write $T = T_1 \oplus \cdots \oplus T_k$.
Conversely, V is direct sum of T-invariant subspaces $U_i = \mathrm{Dom}(T_i)$.

If T_i is reprsented by matrix square $\mathbf{A}_i$, then T is represented by $\mathbf{A} = \left(\begin{array}{c|c|c} \mathbf{A}_1 & & \mathbf{O} \\ \hline & \ddots & \\ \hline \mathbf{O} & & \mathbf{A}_k \end{array}\right)$.

Let $V = U \oplus W$. square matrix $\mathbf{A}_U$ represents $T|_U$, then T is represented by $\mathbf{A} = \left(\begin{array}{c|c} \mathbf{A}_U & * \\ \hline \mathbf{O} & * \end{array}\right)$.
Let $V = U \oplus U^\perp$. $U^\perp$ is T^*-invariant. If $\dim V < \infty$ and T is normal, $U^\perp$ is T-invariant.
If T is normal and U is T^*-invariant, then $(T|_U) = T^*|_U \in \mathscr{L}(U,U)$ and $T|_U$ is normal.
Charpoly $p|_U(\lambda)$ of $T|_U$ divides $p(\lambda)$. ($p(\lambda) = p|_U(\lambda)q(\lambda)$ for some polynomial $q(\lambda)$)
(Cayley-Hamilton Theorem) If $p(\lambda)$ is charpoly of T, then $p(T) = O$.
Minimal polynomial $m(\lambda)$ is **unique** polynomial such that $m(T) = O$ with leading coefficient 1.
If $p(\lambda)$ be such that $p(T) = O$, then $m(\lambda)$ divides $p(\lambda)$ ($p(\lambda) = m(\lambda)q(\lambda)$ for some polynomial $q(\lambda)$)
Set of roots of $m(\lambda)$ consist of all eigenvalues of T.
(Primary Decomposition Theorem)
If $m(\lambda) = p_1(\lambda)\cdots p_k(\lambda)$ where $p_i(\lambda)$ and $p_j(\lambda)$ are relatively prime for $i \neq j$, then $V = \mathrm{Ker}(p_1(T)) \oplus \cdots \oplus \mathrm{Ker}(p_k(T))$.
T is diagonalizable iff $m(\lambda)$ only has **distinct** linear factors.
Let V be inner product space. T is diagonalizable $\implies T|_U$ is also diagonalizable. T is normal $\implies T|_U$ is also normal.
(Spectral Theorem of Commuting Operators)
Let $\{T_i\}_{i=1}^k \subset \mathscr{L}(V,V)$ be set ($\dim V \leq \infty$) of diagonalizable l.t.. $T_iT_j = T_jT_i$ for all i,j iff they can be simultaneously diagonalized.
If each T_i is normal, they can be simultaneously unitarily diagonalized.
(Spectral Theorem of Commuting Matrices) Same with Spectral Theorem of Commuting Operators but $\mathbf{A}_i$ represents T_i for all i.

Chapter 9 Canonical Form

Let $\dim V < \infty$, $T \in \mathscr{L}(V,V)$ and $S = T - \lambda I$.
Generalized eigenspace is invariant subspace of form $\operatorname{Ker}(T-\lambda I)^m$ for some $\lambda \in \mathbb{C}, m \in \mathbb{N}$.
T is *nilpotent* if $T^m = O$ for some positive integer m.
Let $\mathbf{v} \in V$ be vector such that $S^k\mathbf{v} = \mathbf{0}$ but $S^{k-1}\mathbf{v} \neq \mathbf{0}$. Let $U_\mathbf{v}$ be S-invariant cyclic subspace.
$\dim U_\mathbf{v} = k$ with basis $\{S^{k-1}\mathbf{v}, \cdots, S\mathbf{v}, \mathbf{v}\}$ called *Jordan chain* of size k. For $r \le k$, $\operatorname{Ker}(S^r) \cap U_\mathbf{v}$ is spanned by first r basis vectors.

Matrix representing $T = S + \lambda I$ restricted to $U_\mathbf{v}$ is given by *Jordan block* of size $k \times k$ and eigenvalue λ: $\mathbf{J}_\lambda^{(k)} = \begin{pmatrix} \lambda & 1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \dots & \dots & 0 & \lambda \end{pmatrix}$.

V admits a basis of Jordan chains $\mathcal{B} = \{S^{k_1-1}\mathbf{v}_1, \cdots, \mathbf{v}_1\} \cup \{S^{k_2-1}\mathbf{v}_2, \cdots, \mathbf{v}_2\} \cup \cdots \cup \{S^{k_N-1}\mathbf{v}_N, \cdots, \mathbf{v}_N\}$ for some integers N.
In terms of T-invariant cyclic subspaces $V = U_{\mathbf{v}_1} \oplus \cdots \oplus U_{\mathbf{v}_N}$ where $T|_{U_{\mathbf{v}_i}}$ is represented by Jordan block $\mathbf{J}_\lambda^{(k_i)}$.
Combination of Jordan blocks is uniquely determined by T.
Steps to find Jordan basis $\mathbf{P}$:

1. We first find Jordan basis of $\operatorname{Im}(S) : \mathcal{B}_\mathbf{u} = \bigcup_{i=1}^N \{S^{k_i-1}\mathbf{u}_i. \cdots, \mathbf{u}_i\}$ for some $\mathbf{u}_i \in \operatorname{Im}(S)$.
2. Since there exists $\mathbf{v}_i \in V$ such that $\mathbf{u}_i = S\mathbf{v}_i$, We have collection of Jordan chains $\mathcal{B}_\mathbf{v} = \bigcup_{i=1}^N \{S^{k_i}\mathbf{v}_i. \cdots, \mathbf{v}_i\}$ after adding $\mathbf{v}_i$.
3. Extend $\mathcal{B}_\mathbf{v}$ to basis of $\mathcal{B}$ of V by adding some vectors $\mathbf{w}_1, \cdots, \mathbf{w}_k \in V$.
4. There exists $\mathbf{w}_i' \in \operatorname{Span}(\mathcal{B}_\mathbf{v})$ such that $S\mathbf{w}_i = S\mathbf{w}_i'$.
   Modifying $\mathbf{w}_i \mapsto \mathbf{w}_i - \mathbf{w}_i'$, we have Jordan chain basis $\mathcal{B} = \bigcup_{i=1}^N \{S^{k_i}\mathbf{v}_i. \cdots, \mathbf{v}_i\} \cup \{\mathbf{w}_1\} \cup \cdots \cup \{\mathbf{w}_k\}$ of V.

(Jordan Canonical Form) There exist basis of V such that T is represented by $\mathbf{J} = \mathbf{J}_{\lambda_1}^{(k_1)} \oplus \cdots \oplus \mathbf{J}_{\lambda_N}^{(k_N)}$ where λ_i is eigenvalues of T.
Any $\mathbf{A} \in M_{n\times n}(\mathbb{C})$ is similar to *Jordan Canonical Form* $\mathbf{J}$. Decomposition is unique up to permuting order of Jordan blocks. Note: Orthonormal basis may not exist.
Eigenvalues $\lambda_1, \cdots, \lambda_k$ are entries of diagonal.
Charpoly $p(\lambda) = (\lambda - \lambda_1)^{n_1} \cdots (\lambda - \lambda_k)^{n_k}$ and a.m. n_i is number of occurences of λ_i on diagonal.
Minipoly $m(\lambda) = (\lambda - \lambda_1)^{m_1} \cdots (\lambda - \lambda_k)^{m_k}$ where m_i is size of largest λ_i-block in $\mathbf{A}$. G.m. of λ_i is number of λ_i-blocks in $\mathbf{A}$.
If each eigenvalue corresponds to **unique** block, steps to find $\mathbf{P}$ (Do not work if there are multiples Jordan blocks of same eigenvalue):

1. For each eigenvectors $\mathbf{v}_1 = \mathbf{v}$ with eigenvalue λ, solve λ_i such that $(T - \lambda I)\mathbf{v}_i = \mathbf{v}_{i-1}$ until no solution found.
2. Collection $\{\mathbf{v}_1, \cdots, \mathbf{v}_k\}$ is basis corresponding to Jordan block $\mathbf{J}_\lambda^{(k)}$. Repeat with all eigenvectors.

We can use JCF to prove

1. Cayley-Hamilton Theorem holds
2. $\mathbf{A}$ is diagonalizable iff $m(\lambda)$ only has linear factors
3. $\mathbf{A}$ is similar to upper triangulat matrix and $\mathbf{A}^T$

Let $\mathcal{B} = \{\mathbf{v}, T\mathbf{v}, \cdots, T^{r-1}\mathbf{v}\}$ be basis of $U_\mathbf{v}$. $T|_{U_\mathbf{v}}$ is represented by *companion matrix* $\mathbf{C}(g) = \begin{pmatrix} 0 & \dots & \dots & 0 & -a_0 \\ 1 & \ddots & & \vdots & \vdots \\ 0 & \ddots & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & -a_{r-2} \\ 0 & \dots & \dots & 1 & -a_{r-1} \end{pmatrix}$ of $g(\lambda)$.

$g(\lambda) = p|_{U_\mathbf{v}}(\lambda) = \lambda^r + a_{r-1}\lambda^{r-1} + \cdots + a_0$ is charpoly and minipoly of $T|_{U_\mathbf{v}}$.
(Rational Canonical Form)
V can be decomposed into T-invariant cyclic subspaces $V = U_{\mathbf{v}_1} \oplus \cdots \oplus U_{\mathbf{v}_k}$ such that if $g_i(\lambda)$ is charpoly and minipoly of $T|_{U_{\mathbf{v}_i}}$: $g_i(\lambda)$ divides $g_{i+1}(\lambda)$, $p(\lambda) = g_1(\lambda) \cdots g_k(\lambda)$ and $m(\lambda) = g_k(\lambda)$.

Collection of invariant factors $\{g_1(\lambda), \cdots, g_k(\lambda)\}$ is uniquely determined by T. T is represented by $\mathbf{A} = \left(\begin{array}{c|c|c} \mathbf{C}(g_1) & & \mathbf{O} \\ \hline & \ddots & \\ \hline \mathbf{O} & & \mathbf{C}(g_k) \end{array}\right)$