MATH251 - Algebra 2

Summary of Results

Winter, 2024
Notes by Louis Meunier
Complete notes

1 Notation

$\mathbb{F}$ denotes an arbitrary field; in section 6 we will restrict $\mathbb{F}$ to either $\mathbb{R}$ or $\mathbb{C}$. Upper case U, V, W will typically denote vector spaces, lower case Greek letters α, β, γ bases, and lower case a, b, c scalars from $\mathbb{F}$. A subscript (eg I_V, $0_{\mathbb{F}}$) denote "where" an element comes from (eg identity on V, zero on $\mathbb{F}$), but will often be omitted.

$M_{m\times n}(\mathbb{F}) := \{m \times n$ matrices with entries in $\mathbb{F}\}$; if $m = n$ we denote $M_n(\mathbb{F})$. $\mathrm{GL}_n(\mathbb{F}) := \{A \in M_n(\mathbb{F}) : A \text{ invertible }\} \subseteq M_n(\mathbb{F})$.

$\mathbb{F}[t]_n := \{a_0 + a_1 t + \cdots + a_n t^n : a_i \in \mathbb{F}\}$.

Important (purely subjectively) results are highlighted with ⋆ for their use in proofs and other results.

2 Vector Spaces, Linear Relations

Definition 1 (Vector Space). A vector space V defined over a field $\mathbb{F}$ is an abelian group with respect to an addition operation $+$ with identity element $0 \equiv 0_V$, and with an additional scalar multiplication from the field such that for $u, v \in V$ and $a, b \in \mathbb{F}$,

1. $1 \cdot v = v$; $1 \in \mathbb{F}$ (identity)
2. $a \cdot (b \cdot v) = (\alpha \cdot \beta)v$ (associativity of multiplication)
3. $(a + b)v = av + bv$ (distribution of scalar addition over scalar multiplication)
4. $a(u + v) = au + av$ (distribution of scalar multiplication over vector addition)

To follow, unless otherwise specified, take V to be an arbitrary vector space.

Proposition 1. $0_{\mathbb{F}} \cdot v = 0_V$; $-1 \cdot v = -v$; $a \cdot 0_V = 0_V$, $a \in \mathbb{F}$.

Definition 2 (Subspace). $W \subseteq V$, such that W nonempty and W closed under vector addition and scalar multiplication.

Definition 3 (Linear Combination, Span, Spanning Sets). A linear combination of vectors $v_i \in S$ for some set $S \subseteq V$ is a summation $a_1 v_1 + \cdots + a_n v_n$ for scalars $a_i \in \mathbb{F}$.

Define $\operatorname{Span}(\{v_1, \ldots, v_n\}) := \{a_1 v_1 + \cdots + a_n v_n : a_i \in \mathbb{F}\}$.

We say a set S spans V if $\operatorname{Span}(S) = V$; we say S minimally spanning if $\nexists v \in S : S \setminus \{v\}$ spanning.

Proposition 2. *For any set $S \subseteq V$,* $\operatorname{Span}(S)$ *is a subspace, and moreover the smallest subspace containing S (ie, any other subspace containing S must also contain* $\operatorname{Span}(S)$*).*

Sketch. Use the linearity definition of $\operatorname{Span}(S)$ on any other subspace containing S. $\square$

Definition 4 (Linear Independence). A set $S \subseteq V$ is linearly independent if there is no nontrivial linear combinations equal to 0_V; conversely, S is linearly dependent if such a linear combination exists. Symbolically, letting $S := \{v_1, \ldots, v_n\}$

$$S \text{ linearly independent} \iff \left(\sum_i a_i v_i = 0 \iff a_i \equiv 0\right)$$

$$S \text{ linearly dependent } \iff \exists a_i's, \text{ not all zero s.t. } \sum_i a_i v_i = 0$$

Remark 1. Recall the a_i's from a field, so they have inverses unless equal to zero. A common proof technique is to assume one is nonzero, hence has an inverse, and derive a contradiction.

Definition 5 (Maximal Independence)**.** A set S maximally independent if it is independent, and $\nexists v \in V$ s.t. $S \cup \{v\}$ still independent.

Theorem 1. *For $S \subseteq V$, S minimally spanning $\iff$ S linearly independent and spanning $\iff$ S maximally linearly independent $\iff$ every $v \in V$ equals a* unique *linear combination of vectors in S.*

Definition 6 (Basis)**.** If any (hence all) of the above requirements holds, we say S a basis for V.

Lemma 1 (Steinitz Substitution)**.** *Let $Y \subseteq V$ be independent and $Z \subseteq V$ (finite) spanning. Then $|Y| \leqslant |Z|$ and $\exists Z' \subseteq Z : |Z'| = |Z| - |Y|$, and $Y \cup Z'$ still spanning.*

Theorem 2. *If V admits a finite basis, any two bases are equinumerous.*

In such a case, we define $\dim(V) := |\beta|$ *for any basis β for V, and put* $\dim(V) = \infty$ *if V does not admit a finite basis.*

Sketch. Immediate corollary of Steinitz Substitution. □

Corollary 1 (⋆)**.** *For V finite dimensional, any independent set I can be completed to a basis β for V such that $I \subseteq \beta$.*

Remark 2. Other than the general definitions and equivalent notions of a basis, this corollary is certainly the most important from this section, and is used extensively in proofs to follow.

3 Linear Transformations

Throughout this section, assume V, W are vector spaces and T, S linear transformations unless specified otherwise.

Definition 7 (Linear Transformation)**.** A function $T : V \longrightarrow W$ is a linear transformation if it respects the vector space structures, namely $T(av_1 + v_2) = aT(v_1) + T(v_2)$ for any $a \in \mathbb{F}$, $v_1, v_2 \in V$.

We let $I_V : V \longrightarrow V, v \mapsto v$ be the identity transformation. We sometimes call a transformation from a vector space to itself a linear operator.

Proposition 3. $T(0) = 0$

Theorem 3 (★)**.** *Linear transformations are completely determined by their effects on a basis; if $T_0(v_i) = T_1(v_i)$ for every $v_i \in \beta$ for a basis β of V, then $T_0 \equiv T_1$.*

Sketch. Define a transformation as mapping $v := a_1v_1 + \cdots + a_nv_n \mapsto a_1w_1 + \cdots + a_nw_n$ for arbitrary $w_i \in W$. Show that this is linear, and uniquely determined. □

Definition 8 (Isomorphism)**.** An isomorphism of vector spaces V, W is a linear transformation $T : V \longrightarrow W$ that admits a linear inverse T^{-1}. We write $V \cong W$ in this case.

Proposition 4. *T isomorphism $\iff$ T linear and bijection.*

Theorem 4 (★)**.** *If $\dim(V) = n$, $V \cong \mathbb{F}^n$. Moreover, every n-dimensional vector spaces are isomorphic.*

Sketch. Define a transformation that maps $v_i \mapsto e_i$ where v_i basis vectors for V and e_i basis vectors for $\mathbb{F}^n$. Show that this is a linear bijection. □

Definition 9 (Kernel, Image)**.** For $T : V \longrightarrow W$, and put

$$\mathrm{Ker}(T) := \{v \in V : T(v) = 0\} = T^{-1}\{0\} \subseteq V$$
$$\mathrm{Im}(T) := \{T(v) : v \in V\} = T(V) \subseteq W$$

Proposition 5. $\mathrm{Ker}(T), \mathrm{Im}(T)$ *subspaces of* V, W *resp; hence, put* $\mathrm{nullity}(T) := \dim(\mathrm{Ker}(T)), \mathrm{rank}(T) := \dim(\mathrm{Im}(T))$.

Proposition 6. *For* $T : V \to W$ *and* β *a basis for* V, $T(\beta)$ *spans* $\mathrm{Im}(W)$*; hence,* $T(\beta)$ *spans* $W \iff T$ *surjective.*

Proposition 7 ($\star$)**.** *Let* $T : V \to W$*;* T *injective* $\iff \mathrm{Ker}(T) = \{0\}$ *(or, "is trivial")* $\iff$ $T(\beta)$ *independent for any* β*-basis for* $V \iff T(\beta)$ *independent for some* β*-basis for* V.

Remark 3. The second criterion in particular gives a usually quicker way to check injectivity.

Theorem 5 ($\star$ Dimension Theorem)**.** *For* $\dim(V) < \infty$, $\mathrm{nullity}(T) + \mathrm{rank}(T) = \dim(V)$

Sketch. Direct proof follows by constructing a basis for $\mathrm{Ker}(T)$, completing it to a basis for V, taking $T(\beta)$ and noticing the number of redundant vectors.

Alternatively, the first isomorphism theorem gives that $V/\mathrm{Ker}(T) \cong \mathrm{Im}(T)$ and thus $\dim(V/\mathrm{Ker}(T)) = \dim(V) - \dim(\mathrm{Ker}(T)) = \dim(\mathrm{Im}(T))$ where the second equality needs some proof. □

Corollary 2. *Let* $\dim(V) = \dim(W) = n$*. Then* $T : V \to W$ *injective* $\iff$ *surjective* $\iff$ $\mathrm{rank}(T) = n$.

Theorem 6 (First Isomorphism Theorem)**.** $V/\mathrm{Ker}(t) \cong \mathrm{Im}(T)$

Definition 10 (Homomorphism Space)**.** Put $\mathrm{Hom}(V, W) := \{T : V \to W\}$ for T linear. This is a vector space under the natural operations endowed by the linearity of the transforms themselves, ie $(aT_1 + T_2)(v) := a \cdot T_1(v) + T_2(v)$.

Theorem 7. *Let* β, γ *be bases for* V, W *resp. Then* $\{T_{v,w} : v \in \beta, w \in \gamma\}$ *where*

$$T_{v,w}(v') = \begin{cases} w & v' = v \\ 0 & v' \neq v \end{cases}$$

a basis for $\mathrm{Hom}(V, W)$.

Corollary 3. $\dim(\mathrm{Hom}(V, W)) = \dim(V) \cdot \dim(W)$

Sketch. A counting game. $\square$

For any discussion of linear transformations represented with matrices, assume V, W finite dimensional.

Definition 11 (★ Matrix representation of a linear operator)**.** Let $\dim(V) = n, \dim(W) = m$. For a basis $\beta := \{v_1, \ldots, v_n\}$ of V and $\gamma := \{w_1, \ldots, w_m\}$ and $T : V \to W$, put

$$[T]_\beta^\gamma := \begin{pmatrix} | & & | \\ [T(v_1)]_\gamma & \cdots & [T(v_n)]_\gamma \\ | & & | \end{pmatrix} \in M_{m\times n}(\mathbb{F}),$$

where, if $T(v_i) = a_1 w_1 + \cdots + a_m w_m$, we put $[T(v_i)]_\gamma = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$. We call this the coordinate vector of $T(v_i)$ in base γ.

Proposition 8. *Let $n = \dim(V)$ and let $I_\beta : V \to \mathbb{F}^n$, $v \mapsto [v]_\beta$. This is an isomorphism.*

Theorem 8 (★)**.** *Let $T : V \to W$, β, γ bases for V, W respectively. The following diagram commutes:*

$$\begin{tikzcd} \bullet V \arrow[r, "T"] \arrow[d, "I_\beta"'] & \bullet W \arrow[d, "I_\gamma"] \\ \bullet \mathbb{F}^n \arrow[r, dashed, "L_{[T]_\beta^\gamma}"'] & \bullet \mathbb{F}^m \end{tikzcd}$$

ie $I_\gamma \circ T = L_{[T]_\beta^\gamma} \circ I_\beta$, where $L_A(v) := A \cdot v$.

Moreover, $\mathrm{Hom}(V, W) \to M_{m\times n}(\mathbb{F}), T \mapsto [T]_\beta^\gamma$ *an isomorphism.*

Remark 4. This theorem is quite powerful (and has a pretty diagram): any $m \times n$ matrix corresponds to a linear transformation between n- and m-dimensional spaces, and conversely, any such linear transformation can be represented as a matrix. It also allows us to "be a little clever" with our definitions of matrix operations.

Definition 12. For $A \in M_{m\times n}, B \in M_{\ell\times m}(\mathbb{F})$, define $B \cdot A := [L_B \circ L_A]$.

Corollary 4. *Matrix multiplication associative.*

Sketch. Indeed, as function composition is. $\square$

Corollary 5. *For $T : V \to W$, $S : W \to U$ and bases α, β, γ for V, W, U resp., $[S \circ T]_\alpha^\gamma = [S]_\beta^\gamma \cdot [T]_\alpha^\beta$*

Corollary 6. *For $A \in M_n(\mathbb{F})$, L_A invertible $\iff$ A invertible in which case $L_A^{-1} = L_{A^{-1}}$.*

Definition 13 (T-invariant subspace)**.** Let $T : V \to V$; $W \subseteq V$ T-invariant if $T(W) \subseteq W$.

Proposition 9. $\mathrm{Im}(T^n)$ *T-invariant for any $n \in \mathbb{N}$ ie* $V \supseteq \mathrm{Im}(T) \supseteq \mathrm{Im}(T^2) \supseteq \cdots \supseteq \mathrm{Im}(T^n) \supseteq \cdots$.

Similarly, $\mathrm{Ker}(T^n)$ *T-invariant for any $n \in \mathbb{N}$, ie* $\{0\} \subseteq \mathrm{Ker}(T) \subseteq \mathrm{Ker}(T^2) \subseteq \cdots \subseteq \mathrm{Ker}(T^n) \subseteq \cdots$.

Definition 14 (Nilpotent)**.** $T : V \to V$ nilpotent if $T^n = 0$ for some $n \in \mathbb{N}$.

Proposition 10. *If $T : V \to V$ nilpotent, $T^{\dim(V)} = 0$.*

Sketch. Nilpotent $\implies$ $\exists k : T^k = 0$. If $k \leqslant \dim(V)$ this is clear. If $k > \dim(V)$, use proposition 9. $\square$

Definition 15 (Direct Sum)**.** For $W_0, W_1 \subseteq V$, we write $V = W_0 \oplus W_1$ if $W_0 \cap W_1 = \{0_V\}$ and $V = W_0 + W_1$, and say V the direct sum of W_0, W_1.

Theorem 9 (Fitting's Lemma)**.** *For V finite dimensional and a linear transformation $T : V \to V$, we can decompose $V = U \oplus W$ such that U, W T-invariant, T_U nilpotent and T_W an isomorphism.*

Sketch. Using proposition 9 and the finite dimensions, remark that $\exists N$ such that $W := \mathrm{Im}(T^N) = \mathrm{Im}(T^{N+1})$ and $U := \mathrm{Ker}(T^N) = \mathrm{Ker}(T^{N+1})$. Proceed. $\square$

Definition 16 (Dual Space)**.** Let $V^* := \mathrm{Hom}(V, \mathbb{F})$.

Proposition 11. *For V finite dimensional,* $\dim(V^*) = \dim(V)$*; moreover $V^* \cong V$.*

Sketch. Follows directly from the more general corollary 3, or, more instructively, by considering the dual basis: □

Proposition 12. *Let V finite dimensional. For a basis $\beta := \{v_1, \dots, v_n\}$ for V, the dual basis $\beta^* := \{f_1, \dots, f_n\}$, where $f_i(v_j) := \delta_{ij} := \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$ a basis for V^*.*

Definition 17. For each $x \in V$, define $\hat{x} \in V^{**}$ by $\hat{x} : V^* \to \mathbb{F}$, $\hat{x}(f) := f(x)$.

For $S \subseteq V$, put $\hat{S} := \{\hat{x} : x \in S\}$.

Theorem 10 (★)**.** *$x \mapsto \hat{x}$, $V \mapsto V^{**}$ a linear injection, and in particular, an isomorphism if V finite dimensional.*

Moreover, $V^{**} = \hat{V}$.

Sketch. Isomorphism also follows directly from $V^{**} \cong V^*$ (being the dual of the dual) and $\cong$ being an equivalence relation. □

Definition 18 (Annihilator)**.** For $S \subseteq V$ a set, $S^\perp := \{f \in V^* : f|_S = 0\}$.

Proposition 13. *$S^\perp$ a subspace of V^*, $S_1 \subseteq S_2 \subseteq V \implies S_1^\perp \supseteq S_2^\perp$.*

Theorem 11. *If V finite dimensional and $U \subseteq V$ a subspace, $(U^\perp)^\perp = \hat{U}$.*

Definition 19 (Transpose)**.** For $T : V \to W$, define $T^t : W^* \to V^*$, $g \mapsto g \circ T$, ie $T^t(g)(v) = g(T(v))$.

Proposition 14. *(1) T^t linear, (2) $\mathrm{Ker}(T^t) = (\mathrm{Im}(T))^\perp$, (3) $\mathrm{Im}(T^t) = (\mathrm{Ker}(T))^\perp$, and (4) if V, W finite and β, γ bases resp, then $([T]_\beta^\gamma)^t = [T^t]_{\gamma^*}^{\beta^*}$, where A^t represents the typical matrix transpose.*

Sketch. Remark that (1), (2), (3) hold for infinite dimensional spaces; (2) is fairly clear, but the converse direction of (3) is a little tricky. (4) is just a pain notationally. □

Theorem 12. *Let V finite dimensional and $U \subseteq V$ a subspace. Then (1) $\dim(U^\perp) = \dim(V) - \dim(U)$ and (2) $(V/U)^* \cong U^\perp$ by the map $f \mapsto f_U, f_U : V \to \mathbb{F}, v \mapsto f(v + U)$.*

Sketch. For (1), construct a basis for U, complete it, then take the basis and "stare". □

Corollary 7. *T^t injective $\iff$ T surjective; if V, W finite dimensional, T^t surjective $\iff$ T injective.*

Definition 20 (Matrix Rank, C-Rank, R-Rank)**.** For $A \in M_{m\times n}(\mathbb{F})$, define $\operatorname{rank}(A) := \operatorname{rank}(L_A)$, $\text{c-rank}(A) :=$ size of maximally independent subset of columns $\{A^{(1)}, \dots, A^{(n)}\}$, and $\text{r-rank}(A) :=$ the same definition but for rows.

Proposition 15. $\operatorname{rank}(A) = \text{c-rank}(A) = \text{r-rank}(A)$

Sketch. First equality should be clear; second follows either from remarking that $\operatorname{rank}(A) = \operatorname{rank}(A^t) = \text{r-rank}(A)$, or by using tools of the next section. □

4 Elementary Matrices; Determinant

Proposition 16. *For $A \in M_{m\times n}(\mathbb{F}), b \in \operatorname{Im}(L_A)$, the set of solutions to $A\vec{x} = \vec{b}$ is precisely the* coset *$\vec{v} + \operatorname{Ker}(L_A)$ where $\vec{v} \in \mathbb{F}^n$ such that $A\vec{v} = \vec{b}$.*

Proposition 17. *If $m < n$ and $A \in M_{m\times n}(\mathbb{F})$, there is always a nontrivial solution to $A\vec{x} = \vec{0}$.*

Definition 21 (Elementary Row/Column Operations)**.** For $A \in M_{m\times n}(\mathbb{F})$, an elementary row (column) operation is one of

1. interchanging two rows (columns) of A
2. multiplying a row (column) by a nonzero scalar
3. adding a scalar multiple of one row (column) to another.

Remark each operation is invertible.

Definition 22 (Elementary Matrix)**.** An elementary matrix $E \in M_n(\mathbb{F})$ is one obtained from I_n by a elementary row/column operation.

Proposition 18. *Elementary matrices are invertible.*

Proposition 19. *Let $T : V \to W, S : W \to W$ and $R :\to V$ where V, W finite dimensional, and S, R invertible. Then* $\operatorname{rank}(S \circ T) = \operatorname{rank}(T) = \operatorname{rank}(T \circ R)$.

In the language of matrices, if $A \in M_{m\times n}(\mathbb{F}), P \in \mathrm{GL}_m(\mathbb{F}), Q \in \mathrm{GL}_n(\mathbb{F})$, then $\operatorname{rank}(PA) = \operatorname{rank}(A) = \operatorname{rank}(AQ)$.

Proposition 20. *For any two bases α, β for V, there exists a $Q \in \mathrm{GL}_n(\mathbb{F})$ such that $[T]_\alpha Q = Q[T]_\beta$.*

Conversely, for any $Q \in \mathrm{GL}_n(\mathbb{F})$, there exists bases α, β for V such that $Q = [I]_\alpha^\beta$.

Corollary 8 (⋆)**.** *Elementary matrices preserve rank.*

Sketch. Elementary matrices are invertible by proposition 18, so directly apply proposition 19. □

Theorem 13 (Diagonal Matrix Form)**.** *Every matrix $A \in M_n(\mathbb{F})$ can be transformed into a matrix*

$$\begin{bmatrix} I_r & \mathbf{0}_{r\times(n-r)} \\ \mathbf{0}_{(n-r)\times(r)} & \mathbf{0}_{(n-r)\times(n-r)} \end{bmatrix}$$

via row, column operations. Moreover, $\operatorname{rank}(A) = r$.

Sketch. By induction. Not very enlightening proof. □

Corollary 9. *For each $A \in M_n(\mathbb{F})$, there exist $P, Q \in \mathrm{GL}_n(\mathbb{F})$ such that $B := PAQ$ of the form above.*

Corollary 10. *Every invertible matrix a product of elementary matrices.*

Definition 23 ((r)ref)**.** A matrix is said to be in row echelon form (ref) if

1. All zero rows are at the bottom, ie each nonzero row is above each zero row;
2. The first nonzero entry (called a pivot) of each row is the only nonzero entry in its column;
3. The pivot of each row appears to the right of the pivot of the previous row.

If all pivots are 1, then we say that B is in reduced row echelon form (rref).

Theorem 14. *There exist a sequence of row operations 1., 3., to bring any matrix to ref; there exists a sequence of row operations of type 2. to bring a ref matrix to rref. We call such operations "Gaussian elimination".*

Theorem 15. *Applying Gaussian elimination to the augmented matrix* $(A|b) \to (\tilde{A}|\tilde{b})$ *in rref, then* $Ax = b$ *has a solution* $\iff \operatorname{rank}(\tilde{A}|\tilde{b}) = \operatorname{rank}(\tilde{A}) = \sharp$ *non-zero rows of* $\tilde{A}$.

Corollary 11. $Ax = b \iff$ *if* $(A|b)$ *in ref, there is no pivot in the last column.*

Lemma 2. *Let* B *be the rref of* $A \in M_{m\times n}(\mathbb{F})$*. Then, (1)* $\sharp$ *non-zero rows of* $B = \operatorname{rank}(B) = \operatorname{rank}(A) =: r$*, (2) for each* $i = 1, \dots, r$*, denoting* j_i *the pivot of the ith row, then* $B^{(j_i)} = e_i \in \mathbb{F}^m$*; moreover,* $\{B^{(j_1)}, \dots, B^{(j_r)}\}$ *linearly independent, and (3) each column of* B *without a pivot is in the span of the previous columns.*

Corollary 12. *The rref of a matrix is unique.*

Remark 5. See here for a "thorough" derivation of the determinant. It won't be repeated here.

Definition 24 (Multilinear)**.** We say a function $\delta : M_n(\mathbb{F}) \to \mathbb{F}$ is multilinear if it is linear in every row ie

$$\delta\begin{pmatrix} \vec{v}_1 \\ \vdots \\ c\vec{x} + \vec{y} \\ \vdots \\ \vec{v}_n \end{pmatrix} = c \cdot \delta\begin{pmatrix} \vec{v}_1 \\ \vdots \\ c\vec{x} \\ \vdots \\ \vec{v}_n \end{pmatrix} + \delta\begin{pmatrix} \vec{v}_1 \\ \vdots \\ \vec{y} \\ \vdots \\ \vec{v}_n \end{pmatrix}$$

Proposition 21. *For* $\delta : M_n(\mathbb{F}) \to \mathbb{F}$*, if* A *has a zero row, then* $\delta(A) = 0$.

Definition 25 (Alternating)**.** A multilinear form $\delta : M_n(\mathbb{F}) \to \mathbb{F}$ called alternating if $\delta(A) = 0$ for any matrix A with two equal rows.

Proposition 22. *Let* $\delta : M_n(\mathbb{F}) \to \mathbb{F}$ *be alternating and multilinear; then if* B *obtained from* A *by swapping two rows* $\delta(B) = -\delta(A)$.

Proposition 23. *A multilinear $\delta : M_n(\mathbb{F}) \to \mathbb{F}$ is alternating iff $\delta(A) = 0$ for every matrix A with two equal consecutive rows.*

Proposition 24. *If $\delta : M_n(\mathbb{F}) \to \mathbb{F}$ be an alternating multilinear form. Then for $A \in M_n(\mathbb{F})$,*

$$\delta(A) = \sum_{\pi \in S_n} A_{1\pi(1)} A_{2\pi(2)} \cdots A_{n\pi(n)} \delta(\pi I),$$

where $\pi I_n := \begin{pmatrix} - & e_{\pi(1)} & - \\ & \vdots & \\ - & e_{\pi(n)} & - \end{pmatrix}$.

Definition 26 (sgn)**.** Denote $\operatorname{sgn}(\pi) := (-1)^{\sharp\pi}$ where $\sharp\pi :=$ parity of $\pi \equiv$ number of inversions by π.

Corollary 13. *If $\delta : M_n(\mathbb{F}) \to \mathbb{F}$ be an alternative multilinear form. Then for $A \in M_n(\mathbb{F})$,*

$$\delta(A) = \sum_{\pi \in S_n} sgn(\pi) A_{1\pi(1)} A_{2\pi(2)} \cdots A_{n\pi(n)} \delta(I).$$

Moreover, δ uniquely determined by its value on I_n.

Definition 27 (★ Determinant)**.** Let $\delta : M_n(\mathbb{F}) \to \mathbb{F}$ be the unique normalized ($\delta(I_n) = 1$) alternating multilinear form, ie $\det(A) := \sum_{\pi \in S_n} \operatorname{sgn}(\pi) A_{1\pi(1)} \cdots A_{n\pi(n)}$.

Lemma 3. *Let $\delta : M_n(\mathbb{F}) \to \mathbb{F}$ be an alternating multilinear form. Then for any $A \in M_n(\mathbb{F})$ and an elementary matrix E, then $\delta(EA) = c \cdot \delta(A)$ for some non-zero scalar c.*

In particular, if E swaps 2 rows, then $c = -1$; if E multiplies a row by a scalar c, $c = c$; if E adds a scalar multiple of one row to another, $c = 1$.

Theorem 16. *For $A \in M_n(\mathbb{F})$, $\det(A) = 0 \iff A$ noninvertible.*

Sketch. Follows from lemma 3 by writing $A' = E_1 \cdots E_k A$ where A' in rref and applying det. □

Theorem 17. $\det(AB) = \det(A)\det(B)$ *for any $A, B \in M_n(\mathbb{F})$.*

Sketch. Consider two cases, where A either invertible or not. In the former, write A as a product of elementary matrices and apply lemma 3. ◻

Corollary 14. $\det(A^{-1}) = (\det(A))^{-1}$ *for any* $A \in \mathrm{GL}_n(\mathbb{F})$*.*

Corollary 15. $\det(A^t) = \det(A)$ *for any* $A \in M_n(\mathbb{F})$*.*

5 Diagonalization

Motivation to keep in mind: linear transformations are icky. How can we represent them more simply on particular subspaces? Namely, scalar multiplication is the simplest linear transformation (verify that is indeed linear) - can we pick subspaces such that T becomes scalar multiplication on these subspaces?

Definition 28 (Linearly Independent Subspaces)**.** For $V_1, \dots, V_k \subseteq V$, we say $\{V_1, \dots, V_k\}$ linearly independent if $V_i \cap \sum_{j \neq i} V_j = \{0_V\}$ and call $V_1 \oplus \cdots \oplus V_k$ a direct sum.

Definition 29 (Diagnolizable)**.** We say $T : V \to V$ is diagnolizable if there exists V_i's such that $V = \bigoplus_{i=1}^{k} V_i$ and $T|_{V_i}$ is multiplication by a fixed scalar $\lambda_i \in \mathbb{F}$.

Definition 30 (Eigenvalue/vector)**.** For a linear operator $T : V \to V$ and $\lambda \in \mathbb{F}$, we call λ an eigenvalue if there exists a nonzero vector v such that $T(v) = \lambda v$; we call such a v an eigenvector.

Remark 6. v must be nonzero! This is important for proofs to go forward.

Definition 31 (Eigenspace)**.** For an eigenvalue λ of $T : V \to V$, let $\mathrm{Eig}_V(\lambda) \coloneqq \{v \in V : Tv = \lambda v\}$ be the eigenspace of T corresponding to λ.

Proposition 25. *$Eig_V(\lambda)$ a subspace of V.*

Proposition 26. *Trace and determinant are conjugation-invariant; ie for $A, B \in M_n(\mathbb{F})$, if there exists $Q \in \mathrm{GL}_n(\mathbb{F})$ such that $AQ = QB$,* $\mathrm{tr}(A) = \mathrm{tr}(B)$ *and* $\det(A) = \det(B)$.

Definition 32 (Trace, Determinant of Transformation)**.** For $T : V \to V$ where V finite dimensional, put $\mathrm{tr}(T) := \mathrm{tr}(T) := \mathrm{tr}\big([T]_\beta\big)$ and $\det(T) := \det\big([T]_\beta\big)$ for some/any basis for V.

Remark 7. This is well-defined; $[T]_\alpha, [T]_\beta$ are conjugate for any two bases α, β.

Proposition 27 ($\star$)**.** *T diagonalizable $\iff$ there exists a basis β for V such that $[T]_\beta^\beta$ $\iff$ there is a basis for V consisting of eigenvectors for T*

Proposition 28. *A diagonalizable iff $\exists Q \in \mathrm{GL}_n(\mathbb{F})$ such that $Q^{-1}AQ$ diagonal, with the columns of Q eigenvectors of A.*

Proposition 29. *(1) $v \in V$ an eigenvector of T with eigenvalue λ $\iff \in \mathrm{Ker}(\lambda I - T)$, (2) $\lambda \in \mathbb{F}$ an eigenvalue $\iff$ $\lambda I - T$ not invertible $\iff$ $\det(\lambda I - T) = 0$.*

Definition 33 (Characteristic polynomial)**.** For $T : V \to V$, put $p_T(t) = \det(tI_V - T)$. For $A \in M_n(\mathbb{F})$, put $p_A(t) := \det(tI_n - A)$.

Proposition 30 ($\star$)**.** *$p_T(t) = t^n - \mathrm{tr}(T)t^{n-1} + \cdots + (-1)^n \det(T)$, ie p_T a polynomial of degree n and $\cdots$ some polynomials of degree $n-2$.*

Corollary 16. *$T : V \to V$ has at most n distinct eigenvalues.*

Proposition 31. *For eigenvalues $\lambda_1, \ldots, \lambda_k$ and corresponding eigenvectors $v_1, \ldots, v_k$, $\{v_1, \ldots, v_k\}$ linearly independent. Moreover, the eigenspaces $Eig_T(\lambda_i)$ are linearly independent.*

Definition 34 (Geometric, Algebraic Multiplicity)**.** For an eigenvalue λ of $T : V \to V$, put

$$m_g(\lambda) := \dim(\mathrm{Eig}_T(\lambda))$$

and call it the geometric multiplicity of λ, and

$$m_a(\lambda) := \max\{k \geqslant 1 : (t - \lambda)^k | p_T(t)\}$$

and call it the algebraic multiplicity of T.

Proposition 32. *If $T : V \to V$ has eigenvalues $\lambda_1, \ldots, \lambda_k$, $\sum_{i=1}^k m_g(\lambda_i) \leqslant n$; moreover, $\sum_{i=1}^k m_g(\lambda_i) = n \iff T$ diagonalizable.*

Proposition 33. *$m_g(\lambda) \leqslant m_a(\lambda)$ for any λ.*

Sketch. To prove this, you need to use the fact that the characteristic polynomial of T restricted to any T-invariant subspace of V divides the characteristic polynomial of T. □

Definition 35. A polynomial $p(t) \in \mathbb{F}[t]$ splits over $\mathbb{F}$ if $p(t) = a(t - r_1) \cdots (t - r_n)$ for some $a \in \mathbb{F}, r_i \in \mathbb{F}$.

Remark 8. For an eigenvalue λ of $T : V \to V$, $\sum_{i=1}^k m_a(\lambda_i) = n$

Theorem 18 (★ Main Criterion of Diagonalizability)**.** *T diagonalizable iff $p_T(t)$ splits and $m_g(\lambda) = m_a(\lambda)$ for each eigenvalue λ of T.*

Definition 36 (T-cyclic subspace)**.** For $T : V \to V$ and any $v \in V$, the T-cyclic subspace generated by v is the space $\operatorname{Span}(\{T^n(v) : v \in \mathbb{N}\})$.

Lemma 4. *For V finite dimensional, let $v \in V$ and $W :=$ T-cyclic subspace generated by v. Then (1) $\{v, T(v), \ldots, T^{k-1}(v)\}$ is a basis for W where $k := \dim(W)$ and (2) if $T^k(v) = a_0 v + a_1 T(v) + \cdots + a_{k-1} T^{k-1}(v)$, then $p_{T_W}(t) = t^k - a_{k-1}t^{k-1} - \cdots - a_1 t - a_0$.*

Sketch. For (2), write down $[T_W]_\beta$ where β as in part (1). □

Theorem 19 (★ Cayley-Hamilton)**.** *T satisfies its own characteristic polynomial, namely $p_T(T) \equiv 0$.*

6 Inner Product Spaces

All vector spaces in this section should be assumed to be inner product spaces, and all fields $\mathbb{F} \in \{\mathbb{C}, \mathbb{R}\}$.

Definition 37 (Inner Product)**.** A function $\langle ., . \rangle : V \times V \to \mathbb{F}$ is called an inner product if for $u, v, w \in V$, $\alpha \in \mathbb{F}$,

- $\langle v+w, u\rangle = \langle v, u\rangle + \langle w, u\rangle$

- $\langle \alpha u, v\rangle = \alpha\langle u, v\rangle$

- $\langle u, v\rangle = \overline{\langle v, u\rangle}$

- $\langle u, u\rangle \geqslant 0$ and $\langle u, u\rangle = 0 \iff u = 0$.

We call V equipped with such a function an inner product space. Given an inner product, we can define an associated norm $||v|| := \sqrt{\langle v, v\rangle}, v \in V$, and call vectors u such that $||u||$ unit; any vector can be "normalized" to a unit by $\tilde{v} := ||v||^{-1} \cdot v$.

Remark 9. Requirement 3 also gives us that $\langle u, u\rangle$ always real.

Proposition 34 (Properties of Inner Products). *For $u, v, w \in V$, $\alpha \in \mathbb{F}$, $\langle u, v+w\rangle = \langle u, v\rangle + \langle u, w\rangle$, $\langle u, \alpha v\rangle = \overline{\alpha}\langle u, v\rangle$, $||\alpha v|| = |\alpha|\,||v||$, and $\langle v, 0_V\rangle = \langle 0_V, v\rangle = 0$.*

Definition 38 (Orthogonal). $u, v \in V$ orthogonal if $\langle u, v\rangle = 0$; we write $u \perp v$.

We say a set $S \subseteq V$ orthogonal if vectors in S are pair-wise orthogonal, and if in addition each are units, we say S orthonormal.

We say a set $S \subseteq V$ orthogonal to a vector $v \in V$ if $v \perp s \, \forall s \in S$.

Theorem 20 (Pythagorean). *If $u \perp v$, then $||u||^2 + ||v||^2 = ||u+v||^2$; in particular $||u||, ||v|| \leqslant ||u+v||$.*

Definition 39. For u a unit, put $\mathrm{proj}_u(v) := \langle v, u\rangle \cdot u$.

Proposition 35. *For any $v \in V$, u-unit, $v - \mathrm{proj}_u(v) \perp u$.*

Proposition 36. *For any $x, y \in V$, $|\langle x, y\rangle| \leqslant ||x||||y||$ and $||x+y|| \leqslant ||x|| + ||y||$.*

Proposition 37. *Sets of orthonormal vectors are linearly independent. In particular, if $\dim(V) = n$ and $\beta := \{u_1, \ldots, u_n\}$ an orthonormal set, β forms a basis for V, and for any $v \in \beta$,*

$$v = \langle v, u_1\rangle u_1 + \cdots + \langle v, u_n\rangle u_n = \mathrm{proj}_{u_1}(v) + \cdots + \mathrm{proj}_{u_n}(v).$$

Proposition 38. $v \perp V \iff v = 0_V$.

Theorem 21 (Gram-Schmidt)**.** *Every finite-dimensional vector space has an orthonormal basis. One can be constructed "inductively" by starting with a basis $\beta := \{v_1, \ldots, v_n\}$ for V.*

- *(Base) set $u_1 := ||v_1||^{-1}v_1$; put $\alpha := \{u_1\}$.*
- *(Step) given $\alpha := \{u_1, \ldots, u_{k-1}\}$ a set of orthonormal vectors, set*

$$\tilde{u}_k := v_k - \mathrm{proj}_\alpha(v_k) = v_k - \sum_{i=1}^{k-1}\langle v_k, u_i\rangle u_i.$$

  and normalize $u_k := ||\tilde{u}_k||^{-1} \cdot u_k$, and let $\alpha := \alpha \cup \{u_k\}$.
- *Repeat (Step) until $k = n$.*

Definition 40 (Orthogonal Complement)**.** For $S \subseteq V$, put $S^\perp := \{v \in V : v \perp S\}$. Remark that $S^\perp$ a subspace regardless if S is.

Theorem 22. *Let $W \subseteq V$ be a finite dimensional subspace.*

(a) *For $v \in V$, there exists a unique decomposition $v = w + w_\perp$ such that $w \in W, w_\perp \in W^\perp$. We put $\mathrm{proj}_W(v) := w$.*

(b) *$V = W \oplus W^\perp$.*

Corollary 17. *If $\alpha \neq \beta$ two different orthonormal bases for W, $\mathrm{proj}_\alpha(v) = \mathrm{proj}_\beta(v) \,\forall\, v \in V$.*

Theorem 23. *Putting $d(x, y) := ||x - y||, x, y \in V$ and letting $W \subseteq V$-finite subspace, then $d(v, \mathrm{proj}_W(v)) \leqslant d(v, w)$ for any $w \in W$, that is, $\mathrm{proj}_W(v)$ is the closest vector to V in W; it is also unique.*

Corollary 18. *For $W \subseteq V$-finite subspace, $(W^\perp)^\perp = W$.*

For the remainder of the notes, assume V finite dimensional.

Theorem 24 (Riesz Representation). *For V-finite dimensional, then for every $f \in V^*$ there exists a unique $w \in V$ such that $f = f_w$ where $f_w(v) := \langle v, w \rangle, v \in V$. Ie, $w \mapsto f_w$ a linear isomorphism between $V \mapsto V^*$.*

Remark 10. Its helpful to recall what exactly w looks like; namely, if $\{u_1, \ldots, u_n\}$ an orthonormal basis for V, then $w = \overline{f(u_1)}u_1 + \cdots + \overline{f(u_n)}u_n$.

Theorem 25 (Adjoint). *Let $T : V \to V$, then, there exists a unique $T^* : V \to V$ called the adjoint of T such that $\langle Tv, w \rangle = \langle v, T^*w \rangle$ for any $v, w \in V$.*

Remark 11. This proof relies heavily on Riesz.

Proposition 39. *For $T : V \to V$ and β orthonormal basis for V, $[T^*]_\beta = [T]_\beta^*$ (where $A^* := \overline{A^t}$ for $A \in M_n(\mathbb{F})$).*

Proposition 40 (Adjoint Properties).

(a) $T \mapsto T^* : \hom(V, V) \to \hom(V, V)$ *conjugate linear.*

(b) $(T_1 \circ T_2)^* = T_2^* \circ T_1^*$.

(c) $I_V^* = I_V$.

(d) $(T^*)^* = T$.

(e) *T invertible $\implies$ T^* invertible with* $(T^*)^{-1} = (T^{-1})^*$.

Proposition 41 (Kernel, Image of Adjoint). $\mathrm{Im}(T^*)^\perp = \mathrm{Ker}(T)$ *and* $\mathrm{Ker}(T^*) = \mathrm{Im}(T)^\perp$. *Thus,* $\mathrm{rank}(T) = \mathrm{rank}(T^*), \mathrm{nullity}(T) = \mathrm{nullity}(T^*)$.

Remark 12. To prove the second equality, apply the first to T^{**}.

Corollary 19. *λ an eigenvalue of T iff $\overline{\lambda}$ an eigenvalue of T^*.*

Lemma 5 (Schur's). *Let $T : V \to V$ such that $p_T(t)$ splits. Then there is an orthonormal basis β for V such that $[T]_\beta$ upper triangular.*

Definition 41 (Normality)**.** We call $T : V \to V$ normal if $T \circ T^* = T^* \circ T$ (T, T^* commute) and self-adjoint $T = T^*$.

Remark self-adjoint $\implies$ normal, but not the converse; discussion of normal operators applies to self-adjoint.

Proposition 42 (Properties of Normal Operators)**.** *For $T : V \to V$,*

(a) $||Tv|| = ||T^*v||$.

(b) $T - aI_V$ *is normal; moreover* $p(T)$ *for any polynomial* p *normal.*

(c) v *an eigenvector of* T *corresponding to an eigenvalue* λ *iff* v *an eigenvector of* T^* *corresponding to* $\overline{\lambda}$.

(d) For distinct $\lambda_1 \neq \lambda_2$ *eigenvalues* $Eig_T(\lambda_1) \perp Eig_T(\lambda_2)$.