

LINEAR ALGEBRA OUTLINE
UNIVERSITY OF ARIZONA INTEGRATION WORKSHOP, AUGUST 2017

These notes are intended only as a rough summary of the background in linear algebra expected of incoming graduate students. With brevity as our guiding principle, we include virtually no proofs of the results stated in the text below. The reader is encouraged to fill in as many of the arguments and details as possible.

1. LECTURE 1

Throughout these notes, we fix a field k .

1.1. First definitions and properties.

Definition 1.1.1. A **vector space** over k is a set V equipped with maps $+$: $V \times V \rightarrow V$ (addition) and \cdot : $k \times V \rightarrow V$ (scalar multiplication) which satisfy:

- (1) The pair $(V, +)$ forms an abelian group. That is:
 - (a) $+$ is *commutative*: $v + w = w + v$ for all $v, w \in V$.
 - (b) $+$ is *associative*: $u + (v + w) = (u + v) + w$ for all $u, v, w \in V$.
 - (c) There *exists an additive identity* $0 \in V$ such that $0 + v = v$ for all $v \in V$.
 - (d) There *exist additive inverses*: for all $v \in V$, there exists $-v \in V$ such that $v + (-v) = 0$.
- (2) Scalar multiplication gives an action of the group k^\times on V . That is:
 - (a) $a \cdot (b \cdot v) = (ab) \cdot v$ for all $a, b \in k$ and all $v \in V$.
 - (b) $1 \cdot v = v$ for all $v \in V$.
- (3) Scalar multiplication distributes over addition. That is:
 - (a) $a \cdot (v + w) = a \cdot v + a \cdot w$ for all $a \in k$ and all $v, w \in V$.
 - (b) $(a + b) \cdot v = a \cdot v + b \cdot v$ for all $a, b \in k$ and all $v \in V$.

As is customary, we write the scalar product $a \cdot v$ simply as av , we will not mention k when it is clear from context, and we will call elements of the set V *vectors*.

Definition 1.1.2. Let V be a vector space over k . A subset $W \subseteq V$ is called a **subspace** of V if W is also a vector space over k under the addition and scalar multiplication on V .

Remark 1.1.3. It is easy to see that a subset $W \subseteq V$ of a vector space V is a subspace if and only if:

- (1) W is nonempty.
- (2) W is closed under addition: $u + w \in W$ for all $u, w \in W$.
- (3) W is closed under scalar multiplication: $aw \in W$ for all $a \in k$ and all $w \in W$.

Note that the subsets of a fixed vector space V are partially ordered with respect to inclusion.

Date: August 3, 2015.

These notes are heavily influenced by versions from previous years, especially those written by Nick Rogers, Dinesh Thakur, and Doug Ulmer.

1.2. Constructions. We now discuss some fundamental constructions of new vector spaces and subspaces from old ones.

Example 1.2.1. The following are fundamental examples:

- (1) For any k , we have the **zero vector space**, 0 , consisting of the single vector 0 .
- (2) k is a vector space over itself with respect to the canonical addition and multiplication.
- (3) More generally, $k^n := \{(a_1, \dots, a_n) : a_i \in k\}$ is a vector space over k with respect to coordinate-wise addition and scalar multiplication.
- (4) For any vector space V , both V and 0 are subspaces of V .

Definition 1.2.2. Let V be a vector space over k and let $\{W_i\}_{i \in I}$ a collection of subspaces of V .

- (1) The **sum** of the $\{W_i\}$ is

$$\sum_{i \in I} W_i := \left\{ \sum_{\substack{i \in J \\ J \subseteq I \text{ finite}}} w_i : w_i \in W_i \text{ for all } i \right\}.$$

It is a subspace of V , and is the smallest subspace of V containing W_i for all i .

- (2) The **intersection** of the $\{W_i\}$ is the set

$$\bigcap_{i \in I} W_i := \{v \in V : v \in W_i \text{ for all } i\}.$$

It is a subspace of V , and is the largest subspace of V contained in W_i for all i .

Definition 1.2.3. Let I be an index set and $\{V_i\}_{i \in I}$ a collection of vector spaces over k .

- (1) The **direct product** of the V_i is the cartesian product

$$\prod_{i \in I} V_i := \{(v_i) : v_i \in V_i\}$$

with coordinate-wise addition and scalar multiplication.

- (2) The **direct sum** of the V_i is

$$\bigoplus_{i \in I} V_i := \{(v_i) \in \prod_{i \in I} V_i : v_i = 0 \text{ for all but finitely many } i\}$$

with coordinate-wise addition and scalar multiplication.

These are both vector spaces over k , and $\bigoplus_{i \in I} V_i$ is a subspace of $\prod_{i \in I} V_i$ with equality if and only if I is a finite set or $V_i = 0$ for all but finitely many i .

Definition 1.2.4. Let V be a vector space over k and W a subspace of V . The **quotient** of V by W , written V/W , is the vector space whose underlying set is the set $\{v + W : v \in V\}$ of (left) cosets of W in V , with addition and scalar multiplication given by

$$(u + W) + (v + W) := (u + v) + W \quad \text{respectively} \quad a \cdot (v + W) := av + W.$$

One checks that this really is a vector space over k .

1.3. Span and linear independence, bases and dimension.

Definition 1.3.1. Let V be a vector space over k .

- (1) Let $v \in V$ be any vector. The **span** of v is the subspace $k \cdot v := \{av : a \in k\}$ of all k -multiples of v .
- (2) Let $S \subseteq V$ be any subset of V . The **span** of S is the subspace of V

$$\text{span}(S) := \sum_{v \in S} k \cdot v$$

Definition 1.3.2. Let V be a vector space over k .

- (1) Let $v_1, \dots, v_n \in V$ be arbitrary. We say that v_1, v_2, \dots, v_n are **linearly independent** if the only solution to the equation

$$\sum a_i v_i = 0$$

is $a_1 = a_2 = \dots = a_n = 0$.

- (2) Let $S \subseteq V$ be any subset of V . We say that S is a **linearly independent** set of vectors if every finite subset of S is linearly independent. Otherwise we say that S is **linearly dependent**.

Definition 1.3.3. Let V be a vector space over k and S a subset of V . We say that S is a **basis** of V provided:

- (1) S spans V , i.e. $\text{span}(S) = V$.
 (2) S is linearly independent.

Theorem 1.3.4. *The following facts are fundamental:*

- (1) *Every vector space over k has a basis (requires the axiom of choice in general).*
 (2) *Any two bases of a given vector space have the same cardinality.*
 (3) *Any maximal (with respect to inclusion) linearly independent subset of V is a basis.*
 (4) *Any minimal (with respect to inclusion) spanning subset of V is a basis.*

Example 1.3.5. The **standard basis** of $k^n := \bigoplus_{i=1}^n k$ is the basis $\{e_i\}_{i=1}^n$ with $e_i := (0, \dots, 0, 1, 0, \dots, 0)$ the vector whose entries are all 0 except for a 1 in the i th position. Using this, it is easy to see that a choice of basis S of a vector space V is equivalent to an isomorphism $\bigoplus_{s \in S} k \simeq V$.

Definition 1.3.6. Let V be a vector space over k and S a basis of V . The **dimension** of V is the cardinality of S , i.e.: $\dim_k(V) := \#S$.

Remark 1.3.7. As with every definition in these notes, it is possible to give an *intrinsic* definition of $\dim(V)$; i.e. one that does not use bases. It is not, however, possible to avoid the axiom of choice in showing that dimension is well defined in general (one ends up needing the existence of maximal chains of subspaces).

Technique 1.3.8 (Coordinates). If V is a vector space over k and S is a basis of V , then every nonzero $v \in V$ can be written as an *unique* linear combination of elements of S . That is, v determines a unique tuple $\{a_s\}_{s \in S} \in k^S$ such that $a_s = 0$ for all but finitely many $s \in S$ and $v = \sum_{s \in S} a_s s$. Note that this sum makes sense, as only finitely many of its summands are nonzero. We say that $\{a_s\}_{s \in S}$ are the **coordinates of v with respect to the basis S** .

1.4. Linear maps.

Definition 1.4.1. Let V and W be vector spaces over k .

- (1) A **linear map** or a **linear transformation** is a map of sets $T : V \rightarrow W$ which respects addition and scalar multiplication:
 (a) $T(u + v) = T(u) + T(v)$ for all $u, v \in V$.
 (b) $T(av) = aT(v)$ for all $v \in V$ and all $a \in k$.
 (2) Let $T : V \rightarrow W$ be a linear map.
 (a) The **kernel of T** is

$$\ker(T) := \{v \in V : T(v) = 0\}.$$

It is a subspace of V .

- (b) The **image of T** is

$$\text{im}(T) := \{w \in W : w = T(v) \text{ for some } v \in V\}.$$

It is a subspace of W .

Definition 1.4.2. Let V and W be vector spaces over k . We define

$$\text{Hom}(V, W) := \{\text{linear maps } T : V \rightarrow W\}.$$

This set is naturally a vector space over k with addition with addition $(S + T)(v) := S(v) + T(v)$ and scalar multiplication $(aT)(v) := aT(v)$. In the special case that $W = k$, we obtain the **dual** of V given by $V^* := \text{Hom}(V, k)$. If S is a basis of V , and $s \in S$, the **dual functional** $s^* : V \rightarrow k$ is the unique k -linear map with $s^*(t) = \delta_{st}$ for $t \in S$ where δ is the Dirac delta function. The set $S^* := \{s^* : s \in S\}$ is a linearly independent subset of V^* which need not be a basis of V^* ; when it is a basis (e.g. when V has finite dimension) it is called the **dual basis**.

Theorem 1.4.3. Let V be a vector space over k . Then there is a canonical linear map

$$(1.4.1) \quad V \rightarrow V^{**} \quad \text{given by} \quad v \mapsto (f \mapsto f(v))$$

which is injective. When V has finite dimension n , then V^* also has dimension n , and the map (1.4.1) is a canonical isomorphism of V with V^{**} .

Definition 1.4.4. Let $T : V \rightarrow W$ be a linear map of vector spaces over k .

- (1) The **rank** of T is $\text{rk}(T) := \dim_k \text{im}(T)$.
- (2) The **nullity** of T is $\text{null}(T) := \dim_k \text{ker}(T)$.

Theorem 1.4.5. Let V and W be vector spaces over k and $T : V \rightarrow W$ a linear map. There is a canonical isomorphism of k -vector spaces

$$V / \text{ker}(T) \simeq \text{im}(T).$$

In particular, if V and W are finite dimensional then

$$\text{rk}(T) + \text{null}(T) = \dim(V)$$

Technique 1.4.6 (Matrix of a linear map). Let V and W be vector spaces over k of finite dimension and $T : V \rightarrow W$ a linear transformation. Let $\mathbf{v} := \{v_1, \dots, v_n\}$ and $\mathbf{w} := \{w_1, \dots, w_m\}$ be bases of V and of W , respectively. By Technique 1.3.8, there are unique $a_{ij} \in k$ such that

$$T(v_j) = \sum_{i=1}^m a_{ij} w_i \quad \text{for } 1 \leq j \leq n.$$

The **matrix of T with respect to the bases \mathbf{v} , \mathbf{w}** is the $m \times n$ matrix ${}_{\mathbf{w}}[T]_{\mathbf{v}} := (a_{ij})$. If $v \in V$ has coordinates (b_1, \dots, b_n) with respect to the basis \mathbf{v} , then the coordinates of $T(v)$ with respect to \mathbf{w} are given by the matrix product

$$(1.4.2) \quad \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

Conversely, if M is an arbitrary $m \times n$ -matrix, then M may be viewed as the linear transformation $M : k^n \rightarrow k^m$ whose associated matrix in the standard bases on source and target is M . In other words, choosing bases for V and W gives a bijection

$$\text{Hom}(V, W) \simeq \text{Mat}_{m \times n}(k).$$

We will frequently utilize this perspective, and treat $m \times n$ matrices as linear transformations $k^n \rightarrow k^m$.

Remark 1.4.7. The set $\text{Mat}_{m \times n}(k)$ is a k -vector space via addition and scalar multiplication of matrices. The bijection $\text{Hom}(V, W) \simeq \text{Mat}_{m \times n}(k)$ is readily checked to respect the natural vector space structure on both sides. In particular, $\text{Hom}(V, W)$ has dimension $\dim(V) \cdot \dim(W)$.

Technique 1.4.8 (Change of basis). As a special case of Technique 1.4.6, we have the following. Let V be a finite-dimensional vector space over k with fixed basis $\mathbf{e} = \{e_i\}_{i=1}^n$. Let $\mathbf{v} = \{v_i\}_{i=1}^n$ be another basis of V . The **change of basis matrix from \mathbf{e} to \mathbf{v}** is the $n \times n$ matrix $A = (a_{ij}) := {}_{\mathbf{v}}[\text{id}_V]_{\mathbf{e}}$.

- (1) It follows easily from definitions that if $B := {}_{\mathbf{e}}[\text{id}_V]_{\mathbf{v}}$ is the change of basis matrix from \mathbf{v} to \mathbf{e} then

$$AB = BA = \text{id}_n,$$

with id_n the $n \times n$ identity matrix. In particular, A is invertible and $B = A^{-1}$.

- (2) If $v \in V$ has coordinates (b_1, \dots, b_n) with respect to \mathbf{e} , then the coordinates of v with respect to \mathbf{v} are given by the matrix product as in (1.4.2).
- (3) Let W be another finite dimensional vector space with fixed basis $\mathbf{f} = \{f_i\}_{i=1}^m$. Let $\mathbf{w} = \{w_i\}_{i=1}^m$ be another basis of W and let $T : V \rightarrow W$ be a linear transformation. If $A := {}_{\mathbf{v}}[\text{id}_V]_{\mathbf{e}}$ and $B := {}_{\mathbf{w}}[\text{id}_W]_{\mathbf{f}}$ are the change of basis matrices, then the equality of matrices

$${}_{\mathbf{w}}[T]_{\mathbf{v}} = B({}_{\mathbf{f}}[T]_{\mathbf{e}})A^{-1}$$

holds.

1.5. Systems of equations and Gaussian elimination.

Technique 1.5.1 (Elementary row and column operations). Let M be any $m \times n$ matrix over k .

- (1) The following operations on M are called **elementary row operations**

- (a) Swap rows i and j : $R_i \leftrightarrow R_j$.
- (b) Multiply row i by a *nonzero* scalar: $R_i \rightarrow aR_i$.
- (c) Add *any* scalar multiple of row j to row i : $R_i \rightarrow R_i + aR_j$.

If E is any one of these operations (so that $E(M)$ is the matrix obtained from M by performing the row operation E to M), then a little thought with change of basis matrices (or with explicit matrix multiplication) shows that the matrix equation

$$E(M) = E(\text{id}_m)M$$

holds. It is clear that each of these operations can be undone by an operation of the same type, so that $E(\text{id}_m)$ is an invertible matrix. In particular, **any (finite) succession of elementary row operations corresponds to left multiplication by an invertible matrix**. In fact, the converse is true as well.

- (2) The following operations on M are called **elementary column operations**

- (a) Swap columns i and j : $C_i \leftrightarrow C_j$.
- (b) Multiply column i by a *nonzero* scalar: $C_i \rightarrow aC_i$.
- (c) Add *any* scalar multiple of column j to column i : $C_i \rightarrow C_i + aC_j$.

If E is any one of these operations then as above one has the matrix equation

$$E(M) = ME(\text{id}_n)$$

holds. In particular, **any (finite) succession of elementary column operations corresponds to right multiplication by an invertible matrix**, and the converse holds as well.

Technique 1.5.2 (Gaussian Elimination). We say that a matrix M is in **row-echelon form** provided the following hold:

- All nonzero rows are above any row with all zero entries.
- The **leading coefficient** (or **pivot**) of any row (i.e. the first nonzero entry) is to the left of the pivot in every row beneath it.

We say that M is in **reduced row-echelon form** if, in addition, we have:

- Every pivot of M is equal to 1 and all entries above (and below) any pivot are equal to 0.

Let $M = (a_{ij})$ be an $m \times n$ matrix with entries in k . The following procedure (called Gaussian Elimination) uses only elementary row operations to bring M to reduced row-echelon form:

- (1) Find the first column of M (say column j) containing a nonzero entry in some row (row i).
- (2) Swap row i and row 1.
- (3) Multiply row 1 by $1/a_{ij}$. The first nonzero entry in this row is now 1.
- (4) Subtract the appropriate multiple of row 1 from every row below it to obtain all zeroes in column j below the first row.
- (5) While $m > 1$, repeat the above procedure with M replaced by the $(m - 1) \times n$ submatrix of M obtained from deleting the first row and m replaced by $m - 1$.
- (6) For each pivot of the new matrix, subtract the appropriate multiple of the row containing that pivot from all rows above it so that the pivot is the only nonzero entry in its column.

There is of course a variant of Gaussian elimination for columns; we leave this to the reader to work out as an exercise.

Technique 1.5.3. Let $T : V \rightarrow W$ be a linear transformation of finite dimensional vector spaces and let \mathbf{v} and \mathbf{w} be bases of V and W , respectively. Set $M := {}_{\mathbf{w}}[T]_{\mathbf{v}}$. It follows from Technique 1.5.1 that:

- (1) Any elementary row operation on M is equivalent to left multiplying M by a change of basis matrix on V . In particular, **elementary row operations do not change** $\ker(T)$.
- (2) Any elementary column operation on M is equivalent to right multiplying M by a change of basis matrix on W . In particular, **elementary column operations do not change** $\operatorname{im}(T)$.

In particular, if M' is the reduced row-echelon form of M , then $\ker(M) = \ker(M')$ and we conclude that **the nullity of M is the number of zero-rows of M'** . Thanks to Theorem 1.4.5, **the rank of M is the number of leading entries of M'** (all of which are 1).

Technique 1.5.4 (Solving systems of linear equations). A **system of m linear equations in n unknowns** is any collection of equations

$$(1.5.1) \quad \begin{array}{cccc} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n & = & b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n & = & b_2 \\ \vdots & & \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n & = & b_m \end{array}$$

in which the a_{ij} and b_i are fixed scalars and the x_j are unknowns. A **solution** of (1.5.1) is any vector $(x_1, \dots, x_n) \in k^n$ satisfying every one of these equations. Let $A = (a_{ij}) \in \operatorname{Mat}_{m \times n}(k)$ and let $b := (b_1, \dots, b_m) \in k^m$ and $\mathbf{x} := (x_1, \dots, x_n)$, viewed as column vectors. Then the system (1.5.1) is equivalent to the matrix equation $A\mathbf{x} = b$. This may be solved by forming the **augmented matrix**

$$M := \left(\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array} \right)$$

and applying Gaussian elimination to bring it to reduced row-echelon form. The resulting augmented matrix corresponds to an equivalent system of equations whose solutions are easily determined inductively, starting with the equation corresponding to the bottom row and proceeding upwards via “back-substitution”. In the special case where $b = 0$, this allows us to determine a basis of $\ker(A)$ for any $m \times n$ matrix A .

2. LECTURE 2

In this section, we assume that k is algebraically closed, as it simplifies our treatment of eigenvalues.

2.1. Determinants and characteristic polynomial.

Definition 2.1.1. Let n be a positive integer. A **determinant function** is a map $D : \text{Mat}_{n \times n}(k) \rightarrow k$ satisfying

- (1) D is multilinear. That is, for each i with $1 \leq i \leq n$, the function D is a linear map on the i -th row when the other $n - 1$ rows are fixed.
- (2) $D(A) = 0$ whenever A has two equal rows.
- (3) $D(\text{id}_n) = 1$.

Theorem 2.1.2. For any positive integer n , a determinant function exists and is unique. It is given by

$$(2.1.1) \quad \det((a_{ij})) := \sum_{\sigma \in S_n} \text{sgn}(\sigma) \prod_{i=1}^n a_{i\sigma(i)}$$

Here, the sum is over all permutations σ of the set $\{1, 2, \dots, n\}$ and $\text{sgn}(\sigma)$ is the sign of σ .¹ In addition to the characterizing properties listed above, \det satisfies:

- (1) $\det(AB) = \det(A) \det(B)$.
- (2) $\det(A) = \det(A^t)$, where A^t is the transpose of A .
- (3) A is an invertible matrix if and only if $\det(A) \in k^\times$.
- (4) If $A, B \in \text{Mat}_{n \times n}(k)$ with A invertible, then $\det(ABA^{-1}) = \det(B)$.

Technique 2.1.3. There are a number of useful computational techniques for computing determinants, the most common of which is *Laplace's Formula*. Let $A = (a_{ij})$ be an $n \times n$ matrix and define M_{ij} to be the (i, j) -**minor** of A , i.e. the determinant of the $(n - 1) \times (n - 1)$ submatrix of A obtained by deleting the i -th row and j -th column. By definition, the (i, j) -**cofactor** of A is $(-1)^{i+j} M_{ij}$. Then for any fixed i (respectively j) one has

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

These formulae allow one to inductively compute the determinant of any square matrix.

Theorem-Definition 2.1.4. Let V be a finite dimensional vector space over k and $T : V \rightarrow V$ a linear map. For each basis \mathbf{e} of V , the number $\det(\mathbf{e}[T]_{\mathbf{e}})$ is independent of \mathbf{e} ; we call this number the **determinant of T** and denote it $\det(T)$.

Remark 2.1.5. There is of course an *intrinsic* way to define the determinant of a linear map which avoids the crutch of bases and matrices. This method involves exterior algebra, which we do not discuss in these notes. However, we *highly encourage* the reader to study tensor and wedge products. In what follows, we will phrase things in terms of linear maps; the corresponding definitions and theorems for matrices follow immediately from the discussion in [Technique 1.4.6](#).

In what follows, we fix an n -dimensional vector space V and a linear map $T : V \rightarrow V$.

¹One possible definition is that $\text{sgn}(\sigma)$ is $(-1)^{d(\sigma)}$, where $d(\sigma)$ is the number of transpositions in (any) product decomposition of σ into transpositions.

Definition 2.1.6. The **characteristic polynomial** of T is

$$\text{char}_T(x) := \det(T - x \cdot \text{id}).$$

Using the definition (2.1.1) of the determinant and its basic properties, it is easy to see that the following hold:

- (1) $\text{char}_T(x)$ is a monic degree n polynomial in x with coefficients in k .
- (2) The constant term of $\text{char}_T(x)$ is $(-1)^n \det(T)$.

Theorem 2.1.7 (Cayley–Hamilton). *Every linear map satisfies its characteristic polynomial: $\text{char}_T(T) = 0$.*

It follows from the Cayley–Hamilton Theorem that there *exists* a polynomial (of degree n , in fact) satisfied by T , and hence that the set of *all* monic polynomials satisfied by T is nonempty. The well-ordering principle (using the degree of a polynomial) then guarantees the existence of a monic polynomial of least degree satisfied by T . Using the fact that $k[x]$ is a principal ideal domain and monicity, one shows that any two such polynomials are equal:

Definition 2.1.8. The **minimal polynomial** of T , written $\text{min}_T(x)$, is the monic polynomial of least degree satisfied by T .

Theorem 2.1.9. *Let $p(x) \in k[x]$ be any polynomial with $p(T) = 0$. Then min_T divides p . In particular, min_T divides char_T . Furthermore, char_T divides some power of min_T ; in other words, the roots of char_T are precisely the roots of min_T (although the multiplicities may be different).*

2.2. Eigenvalues and eigenvectors.

Definition 2.2.1. A scalar $\lambda \in k$ is said to be an **eigenvalue** of T if there exists a *nonzero* vector $v \in V$ such that

$$(2.2.1) \quad Tv = \lambda v. \quad \text{or equivalently} \quad (T - \lambda)v = 0$$

Any nonzero vector v satisfying (2.2.1) is called an **eigenvector** of T **corresponding to** λ . For an eigenvalue λ , the associated **eigenspace** is the subspace V_λ of V consisting of all vectors satisfying (2.2.1). A **generalized eigenvector corresponding to** λ is any $v \in V$ satisfying

$$(2.2.2) \quad (T - \lambda)^j v = 0 \quad \text{for some } j > 0.$$

For an eigenvalue λ of T , the associated **generalized eigenspace** is the set U_λ of all $v \in V$ which satisfy (2.2.2). It is easy to see that U_λ is a subspace of V . The **geometric multiplicity** $m_{\text{geom}}(\lambda)$ of an eigenvalue λ of T is by definition $\dim(V_\lambda)$. The **algebraic multiplicity** $m_{\text{alg}}(\lambda)$ of an eigenvalue λ is its multiplicity as a root of $\text{char}_T(x)$. We define eigenvalues, (generalized) eigenvectors, and eigenspaces of $n \times n$ matrices to be the corresponding objects of the associated linear maps on k^n .

Theorem 2.2.2. *Let V and T be as above.*

- (1) *The eigenvalues of T are precisely the roots of $\text{char}_T(x)$ (equivalently $\text{min}_T(x)$).*
- (2) *$m_{\text{geom}}(\lambda) = \dim(V_\lambda) \leq \dim_k(U_\lambda) = m_{\text{alg}}(\lambda)$ with equality for all λ if and only if T is diagonalizable.²*
- (3) *Each U_λ is stable under T ; i.e. the restriction $T|_{U_\lambda}$ of T to U_λ has image in U_λ .*
- (4) *If $\lambda \neq \lambda'$ are distinct eigenvalues of T then $U_\lambda \cap U_{\lambda'} = 0$.*
- (5) *Let $\lambda_1, \dots, \lambda_r$ be the distinct eigenvalues of T . The inclusion mapping $\sum_{i=1}^r U_{\lambda_i} \rightarrow V$ induces an isomorphism*

$$\bigoplus_{i=1}^r U_{\lambda_i} \simeq V$$

²See Technique 2.2.5.

Technique 2.2.3 (Finding eigenvalues and vectors). Let $M \in \text{Mat}_{n \times n}(k)$. To find the eigenvalues of M , we simply find all roots of $\text{char}_M(x)$. For each eigenvalue λ , we can then determine a basis of the space of eigenvectors with eigenvalue λ by using Technique 1.5.4 to find a basis of $\ker(M - \lambda \cdot \text{id}_n)$.

Definition 2.2.4. We say that T is:

- (1) **nilpotent** if $T^s = 0$ for some positive integer s , or equivalently:
 - (a) In Theorem 2.2.2 (5) we have $r = 1$ and $\lambda_1 = 0$.
 - (b) \min_T and char_T are both powers of x .
- (2) **semisimple** if any T -stable subspace $W \subseteq V$ admits a T -stable complement, or equivalently
 - (a) \min_T is square-free.
 - (b) V admits a basis of eigenvectors.
 - (c) Every generalized eigenvector is an eigenvector.
 - (d) There exists a basis of V in which the matrix of T is diagonal.

Technique 2.2.5 (Diagonalizing). Let $M \in \text{Mat}_{n \times n}(k)$. We describe an algorithm to determine whether or not M is diagonalizable, and when it is, to find a matrix P with $P^{-1}MP$ a diagonal matrix.

- (1) Compute $\text{char}_M(x)$.
- (2) For each root λ of $\text{char}_M(x)$, use Technique 1.5.4 to determine a basis B_λ of $V_\lambda = \ker(M - \lambda \cdot \text{id}_n)$.
- (3) If $\sum_\lambda \#B_\lambda < n$, then M is not diagonalizable.
- (4) Otherwise, let $B = \bigcup_\lambda B_\lambda$; it is a basis of k^n consisting of eigenvectors of M . Let P be the $n \times n$ matrix whose columns are the elements of B ; it is an invertible matrix satisfying

$$P^{-1}MP = D$$

where D is a diagonal matrix whose diagonal entries are the eigenvalues of M (with multiplicity).

2.3. Jordan canonical form.

Definition 2.3.1. The **Jordan block** of size s associated to $\lambda \in k$ is the $s \times s$ matrix

$$J_{\lambda,s} := \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \cdots & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & \cdots & & & \lambda \end{pmatrix}$$

with λ 's on the diagonal, 1's on the super-diagonal, and 0's elsewhere.

By studying the nilpotent operators $(T - \lambda)|_{U_\lambda}$, one proves:

Theorem 2.3.2 (Jordan Canonical Form). *Let $T : V \rightarrow V$ be a linear map on an n -dimensional vector space. There exists a basis of V with respect to which the matrix of T is in the block diagonal form (called **Jordan form**)*

$$(2.3.1) \quad \begin{pmatrix} J_{\lambda_1,s_1} & & & \\ & J_{\lambda_2,s_2} & & \\ & & \ddots & \\ & & & J_{\lambda_t,s_t} \end{pmatrix}$$

Here, $\lambda_1, \dots, \lambda_t$ are (not necessarily distinct) eigenvalues of T . For any eigenvalue λ , the dimension of U_λ is equal to the sum of the sizes all Jordan blocks above that are associated to λ . The Jordan form of T is uniquely determined by T , up to permutations of the Jordan blocks. Said differently, there exists an invertible matrix $P \in \text{Mat}_{n \times n}(k)$ such that PMP^{-1} is equal to a matrix of the form (2.3.1).

Theorem 2.3.3 (Abstract Jordan form). *Let $T : V \rightarrow V$ be a linear map on a finite dimensional vector space. Then T can be written as a sum $T = S + N$ with $S : V \rightarrow V$ semisimple and $N : V \rightarrow V$ nilpotent, with $SN = NS$.*

Technique 2.3.4 (Computing Jordan form). Let $M \in \text{Mat}_{n \times n}(k)$. To find the Jordan form of M , we proceed as follows:

- (1) Compute $\text{char}_M(x)$.
- (2) Determine all distinct roots $\{\lambda_1, \dots, \lambda_r\}$ of $\text{char}_M(x)$ and their algebraic multiplicities $\{m_1, \dots, m_r\}$.
- (3) For each i , compute $n_j := \text{null}((M - \lambda_i \cdot \text{id}_n)^j)$ for $1 \leq j \leq m_i/2$ using Gaussian elimination.
- (4) For each partition $p := (a_1, \dots, a_v)$ of m_i into nonnegative integers, and each j above, compute

$$z_j := \sum_{i: a_i \geq j} j + \sum_{i: a_i < j} a_i.$$

- (5) For each i , there will be a unique partition (a_1, \dots, a_v) with $n_j = z_j$ for all j . The Jordan blocks in the Jordan form of M corresponding to λ_i are then J_{λ_i, a_ℓ} for $1 \leq \ell \leq v$.

Remark 2.3.5. For each distinct eigen value λ_i of M , let s_i be the size of the largest Jordan block associated to λ_i . Then $\sum s_i$ is equal to the degree of min_M . This fact is often useful in speeding up the computation of the Jordan form of M .

Remark 2.3.6. Using the structure theory of modules over the PID $k[x]$ gives an immediate proof of the existence and uniqueness of Jordan canonical form. It also gives (via Smith normal form) an alternative algorithm for computing the Jordan canonical form of a matrix.

3. LECTURE 3

3.1. Bilinear forms.

Definition 3.1.1. Let V be a vector space over k . A **bilinear form on V** is a map $f : V \times V \rightarrow k$ satisfying

- (1) $f(au + v, w) = af(u, w) + f(v, w)$
- (2) $f(u, av + w) = af(u, v) + f(u, w)$

for all $u, v, w \in V$ and any $a \in k$. That is, f is a linear map in each argument when the other argument is held fixed. We will sometimes write $\langle u, v \rangle$ for $f(u, v)$ if f is clear from context. We denote by $\text{Bilin}(V, V)$ the set of all bilinear forms on V .

It is easy to see that any scalar multiple of a bilinear form or any sum of two bilinear forms is again a bilinear form. This gives $\text{Bilin}(V, V)$ the structure of a vector space over k . Moreover, any bilinear form f on V gives rise to *two* canonical linear maps

$$\begin{aligned} L_f : V &\rightarrow V^* & \text{given by} & & L_f(v)(w) &:= f(v, w) \\ R_f : V &\rightarrow V^* & \text{given by} & & R_f(v)(w) &:= f(w, v) \end{aligned}$$

Definition 3.1.2. Let $f : V \times V \rightarrow k$ be a bilinear form. We say that f is:

- (1) **Left (respectively right) nondegenerate** if L_f (respectively R_f) is injective.
- (2) **Nondegenerate** if it is both left and right nondegenerate.
- (3) **Symmetric** if $f(u, v) = f(v, u)$ for all $u, v \in V$ (equivalently $L_f = R_f$).
- (4) **Skew symmetric** if $f(u, v) = -f(v, u)$ for all $u, v \in V$ (equivalently $L_f = -R_f$).
- (5) **Alternating** if $f(v, v) = 0$ for all $v \in V$.
- (6) **Reflexive** if $f(u, v) = 0 \iff f(v, u) = 0$ for all $u, v \in V$.

Remark 3.1.3. Every alternating form f satisfies $f(u, v) = -f(v, u)$ and hence is skew-symmetric. If $\text{char}(k) \neq 2$ then the converse holds and every skew-symmetric form is alternating. In characteristic 2, the notions of skew-symmetric and symmetric coincide, and there are alternating forms which are not (skew-) symmetric.

Theorem 3.1.4. *Let $f : V \times V \rightarrow V$ be a bilinear form. If $\text{char}(k) \neq 2$ then there exist unique bilinear forms f^\pm on V with f^+ symmetric and f^- alternating (equivalently skew-symmetric) such that $f = f^+ + f^-$.*

It is easy to see that any alternating or symmetric bilinear form is reflexive. In fact, the converse holds:

Theorem 3.1.5. *A reflexive bilinear form on a vector space V is either symmetric or alternating.*

Definition 3.1.6. Let f be a reflexive (i.e. symmetric or alternating) bilinear form on a vector space V .

- (1) We say that $u, v \in V$ are **orthogonal** with respect to f if $f(u, v) = 0$. Thanks to the reflexivity of f , orthogonality is a symmetric condition; i.e. u and v are orthogonal if and only if v and u are orthogonal.
- (2) A vector $v \in V$ is **isotropic** with respect to f if $f(v, v) = 0$.
- (3) A vector $v \in V$ is **anisotropic** with respect to f if $f(v, v) \neq 0$.
- (4) If $W \subseteq V$ is a subspace of V , we define the **orthogonal complement of W in V** to be

$$W^\perp := \{v \in V : f(v, w) = 0 \text{ for all } w \in W\}.$$

- (5) The **radical** of a subspace $W \subseteq V$ with respect to f is

$$\text{rad}(W) := W \cap W^\perp.$$

- (6) A subspace $W \subseteq V$ is

- (a) **Isotropic** with respect to f if $\text{rad}(W) = W$.
- (b) **Coisotropic** with respect to f if $\text{rad}(W) = W^\perp$.
- (c) **Lagrangian** with respect to f if it is isotropic and coisotropic, or equivalently if $W = W^\perp$.
- (d) **Nondegenerate** with respect to f if $\text{rad}(W) = 0$.

Theorem 3.1.7. *Let f be a reflexive bilinear form on a vector space V and $W \subseteq V$ a nondegenerate subspace of V . Then $V = W \oplus W^\perp$.*

Definition 3.1.8 (Quadratic form associated to a symmetric bilinear form). Let f be a symmetric bilinear form on a vector space V . The **quadratic form** associated to f is the map

$$q := q_f : V \rightarrow k \quad \text{given by} \quad q_f(v) = f(v, v).$$

If $\text{char}(k) \neq 2$ then q_f determines f via the identity

$$f(u, v) = \frac{1}{2}(q_f(u + v) - q_f(u) - q_f(v))$$

Technique 3.1.9 (Matrix of a bilinear form). Let V be a vector space over k of dimension n and f a bilinear form on V . If $\mathbf{e} = \{e_1, \dots, e_n\}$ is any basis of V , then the **matrix of f with respect to \mathbf{e}** is the matrix $[f]_{\mathbf{e}} := (f(e_i, e_j))$.

Since f is bilinear, the matrix $[f]_{\mathbf{e}}$ determines f . More precisely, if $u, v \in V$ have coordinates $X := (u_1, \dots, u_n)$ and $Y := (v_1, \dots, v_n)$ with respect to \mathbf{e} , then one checks that we have

$$f(u, v) = X[f]_{\mathbf{e}}Y^t.$$

In fact, for any basis \mathbf{e} of V , the resulting map of sets

$$\text{Bilin}(V, V) \rightarrow \text{Mat}_{n \times n}(k) \quad \text{given by} \quad f \mapsto [f]_{\mathbf{e}}$$

is an isomorphism of vector spaces over k .

Technique 3.1.10 (Change of basis for a bilinear form). Let V be an n -dimensional vector space over k and f a bilinear form on V . Let \mathbf{e} and \mathbf{e}' be two bases of V and denote by $P := \mathbf{e}[\text{id}_V]_{\mathbf{e}'}$ the change of basis matrix from \mathbf{e} -coordinates to \mathbf{e}' -coordinates. Then

$$[f]_{\mathbf{e}'} = P^t [f]_{\mathbf{e}} P$$

Let V be a finite dimensional vector space and f a bilinear form on V . The map $L_f : V \rightarrow V^*$ induces, by pullback, a linear mapping

$$(3.1.1) \quad V^{**} \rightarrow V^* \quad \text{given by} \quad \theta \mapsto \theta \circ L_f.$$

On the other hand, as V is of finite dimension we have a canonical isomorphism $V \simeq V^{**}$ given as in Theorem dualisom. We thus obtain a map

$$L_f^* : V \simeq V^{**} \xrightarrow{(3.1.1)} V^*$$

Theorem 3.1.11. *Let V and f be as above. Then $L_f^* = R_f$. In particular, $\text{rk}(R_f) = \text{rk}(L_f)$.*

Definition 3.1.12. Let f be a bilinear form on a finite dimensional vector space V . The **rank** of f is the number $\text{rk}(f) := \text{rk}(R_f) = \text{rk}(L_f)$.

Corollary 3.1.13. *Let f be a bilinear form on an n -dimensional vector space V . The following are equivalent:*

- (1) f is left nondegenerate.
- (2) $L_f : V \rightarrow V^*$ is an isomorphism.
- (3) f is right nondegenerate.
- (4) $R_f : V \rightarrow V^*$ is an isomorphism.
- (5) f is nondegenerate.
- (6) $\text{rk}(f) = n$
- (7) The matrix $[f]_{\mathbf{e}}$ of f in any basis \mathbf{e} has rank n .

3.2. The cases $k = \mathbf{R}$ and $k = \mathbf{C}$. In the special case that $k = \mathbf{R}$ (the real numbers) or $k = \mathbf{C}$ (complex numbers), we may use the ordering on \mathbf{R} combined with the absolute value on \mathbf{C} to further develop the theory of pairings on vector spaces. We begin with a slight variant on the notion of symmetric bilinear form which is adapted to the case when the base field is \mathbf{C} .

Definition 3.2.1. Let V be a vector space over \mathbf{C} . An **Hermitian form** on V is a map $f : V \times V \rightarrow V$ satisfying

- (1) $f(au + v, w) = af(u, w) + f(v, w)$
- (2) $f(u, v) = \overline{f(v, u)}$

for all $u, v, w \in V$ and all $a \in \mathbf{C}$. These conditions force

- (3) $f(v, au + w) = \overline{a}f(v, u) + f(v, w)$
- (4) $f(v, v) \in \mathbf{R}$

for all $u, v, w \in V$ and $a \in \mathbf{C}$.

Technique 3.2.2 (Matrix of an Hermitian form). Let V be a vector space over \mathbf{C} of dimension n and f an Hermitian form on V . If $\mathbf{e} = \{e_1, \dots, e_n\}$ is any basis of V , then the **matrix of f with respect to \mathbf{e}** is the matrix $[f]_{\mathbf{e}} := (f(e_i, e_j))$. If $u, v \in V$ have coordinates $X := (u_1, \dots, u_n)$ and $Y := (v_1, \dots, v_n)$ with respect to \mathbf{e} , then one checks that we have

$$f(u, v) = X [f]_{\mathbf{e}} \overline{Y}^t.$$

We have the following general structure theorem:

Theorem 3.2.3. *Let V be a finite dimensional vector space over k .*

- (1) *If k has characteristic zero and f is a symmetric bilinear form on V of rank r then there exists a basis of V with respect to which the matrix of f is diagonal, with exactly r nonzero (diagonal) entries.*
- (2) *If $k = \mathbf{R}$ and f is a symmetric bilinear form on V of rank r then there exists a basis \mathbf{e} of V with respect to which the matrix of f is diagonal, with diagonal entries*

$$f(e_i, e_i) = \pm 1 \quad \text{for } 1 \leq i \leq r \quad \text{and } f(e_i, e_i) = 0 \text{ for } r < i \leq n.$$

Furthermore, the number of -1 's on the diagonal depends only on f , and not on any choice of basis.

- (3) *If $k = \mathbf{C}$ and f is an Hermitian form on V then there exists a basis \mathbf{e} of V with respect to which the matrix of f is diagonal, with all diagonal entries equal to 1 or 0.*

Remark 3.2.4. The result of Theorem 3.2.3 (2) allows us to define the **signature** of a symmetric bilinear form on a vector space V over the real numbers as the triple (r^-, r^+, s) where r^- (respectively r^+) is the number of -1 's (respectively $+1$'s) on the diagonal and s is the number of zeroes. Clearly $r = r^- + r^+$ and $r + s = \dim(V)$. The signature of a bilinear form depends only on the form.

The method of proof of Theorem 3.2.3 is interesting and important. Rather than prove the theorem, however, we discuss the Gram-Schmidt algorithm, which gives a constructive proof in a special case. It is straightforward to adapt these ideas to the general case.

Definition 3.2.5. Let V be a vector space over \mathbf{R} or \mathbf{C} . An **inner product** on V is a map $\langle, \rangle : V \times V \rightarrow \mathbf{C}$ satisfying

- (1) $\langle v, v \rangle$ lies in \mathbf{R} and is nonnegative.
- (2) $\langle v, v \rangle = 0$ if and only if $v = 0$.
- (3) $\langle au + v, w \rangle = a\langle u, w \rangle + \langle v, w \rangle$.
- (4) $\langle u, v \rangle = \overline{\langle v, u \rangle}$.

for all $u, v, w \in V$ and any $a \in k$. The first two conditions are called **positivity** and **definiteness**, respectively. Note that when the base field is \mathbf{R} , an inner product is the same as a positive definite symmetric bilinear form (automatically nondegenerate), and when the base field is \mathbf{C} , an inner product is a positive definite Hermitian form. We define the **norm** of a vector $v \in V$ to be $\|v\| := \sqrt{\langle v, v \rangle}$; the square root makes sense due to positivity. An **inner product space** is a vector space V equipped with an inner product.

Technique 3.2.6 (Gram-Schmidt). Let V be an n -dimensional inner product space (over \mathbf{R} or \mathbf{C}) and let $\{v_1, \dots, v_n\}$ be a basis of V . We will construct an **orthogonal basis** of V , i.e. a basis $\{u_1, \dots, u_n\}$ with $\langle u_i, u_j \rangle$ nonzero if and only if $i = j$. It is clear that in this basis, the matrix of \langle, \rangle is diagonal.

We inductively define

$$\begin{aligned} u_1 &:= v_1 \\ u_2 &:= v_2 - \frac{\langle v_2, u_1 \rangle}{\langle u_1, u_1 \rangle} u_1 \\ &\vdots \\ u_n &:= v_n - \frac{\langle v_n, u_1 \rangle}{\langle u_1, u_1 \rangle} u_1 - \frac{\langle v_n, u_2 \rangle}{\langle u_2, u_2 \rangle} u_2 - \dots - \frac{\langle v_n, u_{n-1} \rangle}{\langle u_{n-1}, u_{n-1} \rangle} u_{n-1} \end{aligned}$$

As the v_i may be recovered from the u_i , we see that $\{u_i\}$ is a basis of V , and one easily checks that it is an orthogonal basis. Given an orthogonal basis $\{u_i\}$ of V , we can construct an **orthonormal basis** of V , i.e. a basis $\{w_i\}$ satisfying $\langle w_i, w_j \rangle = \delta_{ij}$ by setting $w_i := \frac{1}{\|u_i\|} u_i$.

3.3. Operators on inner product spaces. Let V be a finite dimensional inner product space. The inner product on V gives a map

$$(3.3.1) \quad V \rightarrow V^* \quad \text{via} \quad v \mapsto \langle \cdot, v \rangle,$$

which is injective (see Corollary 3.1.13) as any inner product is nondegenerate. It follows that (3.3.1) is an isomorphism, though be warned that when the base field is \mathbf{C} , the complex structure on the target must be twisted by complex conjugation (so the isomorphism in this case is “semilinear over complex conjugation”).

Definition 3.3.1. Let V and W be inner product spaces and $T : V \rightarrow W$ a linear map. The **adjoint of T** is the linear map $T^* : W \rightarrow V$ defined by the composition

$$W \xrightarrow[\simeq]{(3.3.1)} W^* \xrightarrow{T^\vee} V^* \xrightarrow[\simeq]{(3.3.1)^{-1}} V$$

By definition, T^* is the unique mapping $W \rightarrow V$ satisfying

$$\langle Tv, w \rangle_W = \langle v, T^*w \rangle_V$$

for all $v \in V$ and $w \in W$. One checks that the mapping

$$\Theta : \text{Hom}(V, W) \xrightarrow{T \mapsto T^*} \text{Hom}(W, V)$$

is semi-linear over complex conjugation (i.e. it is additive and satisfies $\Theta(af) = \bar{a}\Theta(f)$). Moreover, one has $T^{**} = T$ for any $T \in \text{Hom}(V, W)$, and if U is a third inner product space then $(ST)^* = T^*S^*$ for any $S \in \text{Hom}(W, U)$ and $T \in \text{Hom}(V, W)$.

Remark 3.3.2. Let \mathbf{e} be a basis of V and \mathbf{f} a basis of W . For $T \in \text{Hom}(V, W)$ one checks that

$$\mathbf{f}[T]_{\mathbf{e}} = \mathbf{e}[\overline{T^*}]_{\mathbf{f}}^t.$$

That is, the matrix of the adjoint of T is the conjugate transpose of the matrix of T .

Definition 3.3.3. Let $T : V \rightarrow V$ be a linear map on an inner product space V . We say that T is:

- (1) **Self adjoint** if $T^* = T$. If we wish to emphasize that the ground field is \mathbf{R} , we will say that T is **symmetric**, while in the case that the ground field is \mathbf{C} , we will say that T is **Hermitian**.
- (2) **Normal** if T commutes with its adjoint; i.e. if $TT^* = T^*T$. Observe that every self adjoint operator is normal. There are, however, normal operators which are not self adjoint.
- (3) An **Isometry** if $\|Tv\| = \|v\|$ for all $v \in V$. If we wish to emphasize that the ground field is \mathbf{R} , we will say in this case that T is **orthogonal**; in the complex case we will say **unitary**.

Remark 3.3.4. It is easy to see that any eigenvalue of an Hermitian operator must be real, while any eigenvalue of a unitary operator must have absolute value 1.

Theorem 3.3.5 (Spectral Theorem). *Let V be an inner product space and $T : V \rightarrow V$.*

- (1) *If the base field is \mathbf{C} , then V has an orthonormal basis consisting of eigenvectors for T if and only if T is normal.*
- (2) *If the base field is \mathbf{R} , then V has an orthonormal basis consisting of eigenvectors for T if and only if T is self adjoint.*

From part (1) of the Spectral Theorem we deduce:

Corollary 3.3.6. *Let $T : V \rightarrow V$ be a normal linear transformation on a complex inner product space. Then T is Hermitian if and only if all eigenvalues are real and unitary if and only if all eigenvalues have absolute value 1.*