LINEAR ALGEBRA: A BRIEF OVERVIEW
UNIVERSITY OF ARIZONA INTEGRATION WORKSHOP (2019)

YONG SUK MOON

Abstract. This is an expository note on linear algebra for University of Arizona Integration Workshop 2019. This note is largely based on the notes from previous years by Dan Madden, Bryden Cais, Nick Rogers, Dinesh Thakur, and Doug Ulmer .

1. Lecture 1

1.1. **Vector spaces.** The central objects we study in liner algebra are vector spaces.

Definition 1.1. A **field** K is a commutative ring such that $1 \neq 0$ and every non-zero element $x \in K$ has a multiplicative inverse x^{-1}.

Example 1.2. Typical examples of fields we can consider are $\mathbf{Q}, \mathbf{R}, \mathbf{C}$, and $\mathbf{Q}_p$. We can also consider finite fields $\mathbf{F}_{p^n}$.

Definition 1.3. A **vector space** over a field K is a set V equipped with maps $+ : V \times V \to V$ (addition) and $\cdot : K \times V \to V$ (scalar multiplication) which satisfy:

(1) The pair $(V, +)$ forms an abelian group:
  (a) $+$ is *commutative*: $\mathbf{v} + \mathbf{w} = \mathbf{w} + \mathbf{v}$ for all $\mathbf{v}, \mathbf{w} \in V$.
  (b) $+$ is *associative*: $\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}$ for all $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$.
  (c) There *exists an additive identity* $\mathbf{0} \in V$ such that $\mathbf{0} + \mathbf{v} = \mathbf{v}$ for all $\mathbf{v} \in V$.
  (d) There *exist additive inverses*: for all $\mathbf{v} \in V$, there exists $-\mathbf{v} \in V$ such that $\mathbf{v} + (-\mathbf{v}) = 0$.

(2) Scalar multiplication gives an action of the group $K^\times$ on V:
  (a) $a \cdot (b \cdot \mathbf{v}) = (ab) \cdot \mathbf{v}$ for all $a, b \in K$ and all $\mathbf{v} \in V$.
  (b) $1 \cdot \mathbf{v} = \mathbf{v}$ for all $\mathbf{v} \in V$.

(3) Scalar multiplication distributes over addition:
  (a) $a \cdot (\mathbf{v} + \mathbf{w}) = a \cdot \mathbf{v} + a \cdot \mathbf{w}$ for all $a \in K$ and all $\mathbf{v}, \mathbf{w} \in V$.
  (b) $(a + b) \cdot \mathbf{v} = a \cdot \mathbf{v} + b \cdot \mathbf{v}$ for all $a, b \in K$ and all $\mathbf{v} \in V$.

We often simply write $a\mathbf{v}$ for the scalar product $a \cdot \mathbf{v}$, and will not mention the underlying field K if clear from context. The elements of the set V are called **vectors**.

Definition 1.4. Let V be a vector space over a field 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.

Lemma 1.5. *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: $\mathbf{u} + \mathbf{w} \in W$ for all $u, w \in W$.
(3) W is closed under scalar multiplication: $a\mathbf{w} \in W$ for all $a \in K$ and all $\mathbf{w} \in W$.

Here are some fundamental examples and constructions of vector spaces.

Example 1.6.
(1) For any field K, we have the **zero vector space**, $\{\mathbf{0}\}$, consisting of the single vector $\mathbf{0}$.
(2) K is a vector space over itself equipped with the canonical addition and multiplication.
(3) More generally, $K^n := \{(a_1, \ldots, a_n) \ : \ a_i \in K\}$ is a vector space over K equipped with the coordinate-wise addition and scalar multiplication.
(4) For any vector space V, both V and $\{\mathbf{0}\}$ are subspaces of V.

Example 1.7. For a little more interesting examples:

(1) The polynomial ring $K[x]$ with the natural addition and scalar multiplication is a vector space over K.
(2) The set of continuous functions $f : [0,1] \to \mathbf{R}$ with the natural addition and scalar multiplication is a vector space over $\mathbf{R}$.

Definition 1.8. Let V be a vector space over a field K, and let I be an index set and $\{W_i\}_{i\in I}$ a collection of subspaces of V.

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

$$\sum_{i\in I} W_i := \{ \sum_{\substack{i\in J \\ J\subseteq I \text{ finite}}} \mathbf{w}_i \ : \ \mathbf{w}_i \in W_i \text{ for all } i\}.$$

It is a subspace of V, and is the smallest subspace of V (with respect to the partial ordering by inclusion) containing W_i for all i.

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

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

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

Definition 1.9. 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

$$\prod_{i\in I} V_i := \{(\mathbf{v}_i) \ : \ \mathbf{v}_i \in V_i\}$$

with coordinate-wise addition and scalar multiplication.

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

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

with coordinate-wise addition and scalar multiplication.

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

Definition 1.10. 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 $\{\mathbf{v} + W : \mathbf{v} \in V\}$ of (left) cosets of W in V, with addition and scalar multiplication given by

$$(\mathbf{u} + W) + (\mathbf{v} + W) := (\mathbf{u} + \mathbf{v}) + W,\ a \cdot (\mathbf{v} + W) := a\mathbf{v} + W.$$

We now go over some fundamental concepts and properties for studying the structure of vector spaces.

Definition 1.11. Let V be a vector space over K.

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

$$\operatorname{span}(S) := \sum_{\mathbf{v} \in S} K \cdot \mathbf{v}$$

.

Definition 1.12. Let V be a vector space over K.

(1) For vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n \in V$, we say that $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ are **linearly independent** if the equation

$$\sum a_i \mathbf{v}_i = \mathbf{0}$$

has the unique solution $a_1 = a_2 = \cdots = a_n = 0$ in K.
(2) Let $S \subseteq V$ be a 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 1.13. Let V be a vector space and S a subset of V. We say that S is a **basis** of V provided:

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

Theorem 1.14. *The following facts are fundamental:*

(1) Every vector space 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 1.15. (1) The **standard basis** of $K^n := \bigoplus_{i=1}^n K$ is the basis $\{\mathbf{e}_i\}_{i=1}^n$ with $\mathbf{e}_i := (0, \ldots, 0, 1, 0, \ldots, 0)$ the vector whose entries are all 0 except for a 1 in the i th position.
(2) $\{1, x, x^2, \ldots\}$ is a basis of $K[x]$.

Definition 1.16. 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 1.17. If V is a vector space over K and S is a basis of V, then every nonzero $\mathbf{v} \in V$ can be written as a *unique* linear combination of elements of S. That is, $\mathbf{v}$ 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 $\mathbf{v} = \sum_{s \in S} a_s s$. We say that $\{a_s\}_{s \in S}$ are the **coordinates of v with respect to the basis** S.

1.2. **Linear transformations.** We often encounter situations to study certain maps between vector spaces.

Definition 1.18. Let V and W be vector spaces over K.

(1) A **linear map** or a **linear transformation** is a map of sets $T : V \to W$ which respects addition and scalar multiplication:
   (a) $T(\mathbf{u}+\mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$ for all $\mathbf{u}, \mathbf{v} \in V$.
   (b) $T(a\mathbf{v}) = aT(\mathbf{v})$ for all $\mathbf{v} \in V$ and all $a \in K$.

(2) Let $T : V \to W$ be a linear map.
   (a) The **kernel of** T is
   $$\ker(T) := \{\mathbf{v} \in V \ : \ T(\mathbf{v}) = 0\}.$$
   It is a subspace of V.
   (b) The **image of** T is
   $$\mathrm{im}(T) := \{\mathbf{w} \in W \ : \ \mathbf{w} = T(\mathbf{v}) \text{ for some } \mathbf{v} \in V\}.$$
   It is a subspace of W.
   (c) We say T is an **isomorphism** if it is both injective and surjective, i.e., $\ker(T) = \{\mathbf{0}\}$ and $\mathrm{im}(T) = W$. When T is an isomorphism, we also say V and W are **isomorphic**.

Definition 1.19. Let V and W be vector spaces over K. We define
$$\mathrm{Hom}(V, W) := \{\text{linear maps } T : V \to W\}.$$
This set is naturally a vector space over K with addition $(S+T)(\mathbf{v}) := S(\mathbf{v}) + T(\mathbf{v})$ and scalar multiplication $(aT)(\mathbf{v}) := aT(\mathbf{v})$. In the special case that $W = K$, we obtain the **dual** of V given by $V^* := \mathrm{Hom}(V, K)$. If S is a basis of V, and $s \in S$, the **dual functional** $s^* : V \to K$ 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^*; when it is a basis (e.g. when V has finite dimension) it is called the **dual basis**.

Theorem 1.20. *Let V be a vector space over K. Then there is a canonical linear map*
$$V \to V^{**} \qquad \textit{given by} \qquad \mathbf{v} \mapsto (f \mapsto f(\mathbf{v})) \tag{1.1}$$
which is injective. When V has finite dimension n, then V^* also has dimension n, and the map (1.1) *is a canonical isomorphism of V with V^{**}.*

Definition 1.21. Let $T : V \to W$ be a linear map of vector spaces over K.

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

Theorem 1.22. *Let V and W be vector spaces over K, and $T : V \to W$ a linear map. There is a canonical isomorphism of K-vector spaces*
$$V/\ker(T) \simeq \mathrm{im}(T).$$
In particular, if V and W are finite dimensional then
$$\mathrm{rk}(T) + \mathrm{null}(T) = \dim(V).$$

Remark 1.23 (Matrix associated with a linear map). Let V and W be vector spaces over K of finite dimension and $T : V \to W$ a linear transformation. Let $\mathcal{B} := \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ and $\mathcal{B}' := \{\mathbf{w}_1, \dots, \mathbf{w}_m\}$ be bases of V and of W, respectively. There are unique $a_{ij} \in K$ such that

$$T(\mathbf{v}_j) = \sum_{i=1}^{m} a_{ij}\mathbf{w}_i \qquad \text{for } 1 \le j \le n.$$

The **matrix of T with respect to the bases $\mathcal{B}$, $\mathcal{B}'$** is the $m \times n$ matrix ${}_{\mathcal{B}'}[T]_{\mathcal{B}} := (a_{ij})$. If $\mathbf{v} \in V$ has coordinates $(b_1, \dots, b_n)$ with respect to the basis $\mathcal{B}$, then the coordinates of $T(\mathbf{v})$ with respect to $\mathcal{B}'$ are given by the matrix product

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} \tag{1.2}$$

Conversely, if M is an arbitrary $m \times n$ matrix, then M may be viewed as the linear transformation $M : K^n \to K^m$ 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

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

We can utilize this perspective and treat $m \times n$ matrices as linear transformations $K^n \to K^m$. In fact, the set $\operatorname{Mat}_{m\times n}(K)$ is a K-vector space via addition and scalar multiplication of matrices, and the bijection $\operatorname{Hom}(V, W) \simeq \operatorname{Mat}_{m\times n}(K)$ is an isomorphism of vector space. In particular, $\operatorname{Hom}(V, W)$ has dimension $\dim(V) \cdot \dim(W)$.

Remark 1.24 (Change of basis). As a special case of Remark 1.23, we have the following. Let V be a finite-dimensional vector space over K with fixed basis $\mathcal{B} = \{\mathbf{e}_i\}_{i=1}^{n}$. Let $\mathcal{B}' = \{\mathbf{v}_i\}_{i=1}^{n}$ be another basis of V. The **change of basis matrix from $\mathcal{B}$ to $\mathcal{B}'$** is the $n \times n$ matrix $A = (a_{ij}) := {}_{\mathcal{B}'}[\mathrm{Id}_V]_{\mathcal{B}}$.

(1) It follows easily from definitions that if $B := {}_{\mathcal{B}}[\mathrm{id}_V]_{\mathcal{B}'}$ is the change of basis matrix from $\mathcal{B}'$ to $\mathcal{B}$ then

$$AB = BA = \mathrm{Id}_n,$$

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

(2) If $\mathbf{v} \in V$ has coordinates $(b_1, \dots, b_n)$ with respect to $\mathcal{B}$, then the coordinates of $\mathbf{v}$ with respect to $\mathcal{B}'$ are given by the matrix product as in (1.2).

(3) Let W be another finite dimensional vector space, and let $\mathcal{C}$ and $\mathcal{C}'$ be bases of W. Let $T : V \to W$ be a linear transformation. If $A := {}_{\mathcal{B}'}[\mathrm{Id}_V]_{\mathcal{B}}$ and $B := {}_{\mathcal{C}'}[\mathrm{Id}_W]_{\mathcal{C}}$ are the change of basis matrices, then the equality of matrices

$${}_{\mathcal{C}'}[T]_{\mathcal{B}'} = B({}_{\mathcal{C}}[T]_{\mathcal{B}})A^{-1}$$

holds.

We now discuss a system of linear equations and its connection with linear maps. A **system of m linear equations in n unknowns** over K is any collection of equations

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ \vdots \qquad \cdots \qquad & \quad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned} \tag{1.3}$$

in which the a_{ij} and b_i are given scalars in K and the x_j are unknowns. A **solution** of (1.3) is any vector $(x_1, \ldots, x_n) \in K^n$ satisfying the system. Let $A = (a_{ij}) \in \mathrm{Mat}_{m\times n}(K)$, and let $\mathbf{b} := (b_1, \ldots, b_n) \in K^n$ and $\mathbf{x} := (x_1, \ldots, x_n)$ viewed as column vectors. Then the system (1.3) is equivalent to the matrix equation $A\mathbf{x} = \mathbf{b}$. 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 when $\mathbf{b} = \mathbf{0}$, if we identify A with a linear map $K^n \to K^m$, then the set of solutions can be identified with $\ker(A)$, and the above procedure allows us to determine a basis of $\ker(A)$. In general, the system has a solution if and only if $\mathbf{b}$ lies in $\mathrm{im}(A)$. If it has a solution, then the set of solutions can be identified with a translation of $\ker(A)$ by any choice of solution.

1.3. **Bilinear forms.**

Definition 1.25. Let V be a vector space over K. A **bilinear form on** V is a map $f : V \times V \to K$ satisfying

1. $f(a\mathbf{u} + \mathbf{v}, \mathbf{w}) = af(\mathbf{u}, \mathbf{w}) + f(\mathbf{v}, \mathbf{w})$
2. $f(\mathbf{u}, a\mathbf{v} + \mathbf{w}) = af(\mathbf{u}, \mathbf{v}) + f(\mathbf{u}, \mathbf{w})$

for all $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ 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 $\langle \mathbf{u}, \mathbf{v} \rangle$ for $f(\mathbf{u}, \mathbf{v})$ if f is clear from context.

We denote by $\mathrm{Bilin}(V, V)$ the set of all bilinear forms on V. Note that any scalar multiple of a bilinear form or any sum of two bilinear forms is again a bilinear form. This gives $\mathrm{Bilin}(V, V)$ the structure of a vector space over K. Moreover, any bilinear form f on V gives rise to *two* canonical linear maps

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

Definition 1.26. Let $f : V \times V \to K$ be a bilinear form. We say that f is:

1. **Left (respectively right) non-degenerate** if L_f (respectively R_f) is injective.
2. **Non-degenerate** if it is both left and right non-degenerate.

(3) **Symmetric** if $f(\mathbf{u},\mathbf{v}) = f(\mathbf{v},\mathbf{u})$ for all $\mathbf{u},\mathbf{v} \in V$ (equivalently $L_f = R_f$).
(4) **Alternating** if $f(\mathbf{v},\mathbf{v}) = 0$ for all $\mathbf{v} \in V$.

Definition 1.27. Let f be a symmetric bilinear form on a vector space V.

(1) We say that $\mathbf{u},\mathbf{v} \in V$ are **orthogonal** with respect to f if $f(\mathbf{u},\mathbf{v}) = 0$.
(2) If $W \subseteq V$ is a subspace of V, we define the **orthogonal complement of** W **in** V to be

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

We first consider the special case when $K = \mathbf{R}$ and the bilinear form f is symmetric and **positive definite**, i.e., $f(\mathbf{v},\mathbf{v}) \geq 0$ for all $\mathbf{v} \in V$ and the equality implies $\mathbf{v} = \mathbf{0}$.

Proposition 1.28. *Let V be a non-zero finite dimensional vector space over $\mathbf{R}$, equipped with a symmetric positive definite bilinear form. Then V has an **orthogonal basis**, i.e. a basis $\{\mathbf{u}_1,\ldots,\mathbf{u}_n\}$ with $\langle \mathbf{u}_i,\mathbf{u}_j\rangle$ nonzero if and only if $i = j$.*

If W is a subspace of V, then $V = W \oplus W^{\perp}$. In particular, $\dim(W) + \dim(W^{\perp}) = \dim(V)$.

Proof. We apply *Gram-Schmidt process* to any given basis $\{\mathbf{v}_1,\ldots,\mathbf{v}_n\}$ of V. We inductively define

$$\begin{aligned}
\mathbf{u}_1 &:= \mathbf{v}_1 \\
\mathbf{u}_2 &:= \mathbf{v}_2 - \frac{\langle \mathbf{v}_2,\mathbf{u}_1\rangle}{\langle \mathbf{u}_1,\mathbf{u}_1\rangle}\mathbf{u}_1 \\
&\ \vdots \\
\mathbf{u}_n &:= \mathbf{v}_n - \frac{\langle \mathbf{v}_n,\mathbf{u}_1\rangle}{\langle \mathbf{u}_1,\mathbf{u}_1\rangle}\mathbf{u}_1 - \frac{\langle \mathbf{v}_n,\mathbf{u}_2\rangle}{\langle \mathbf{u}_2,\mathbf{u}_2\rangle}\mathbf{u}_2 - \cdots - \frac{\langle \mathbf{v}_n,\mathbf{u}_{n-1}\rangle}{\langle \mathbf{u}_{n-1},\mathbf{u}_{n-1}\rangle}\mathbf{u}_{n-1}.
\end{aligned}$$

As $\mathbf{v}_i$ may be recovered from the $\mathbf{u}_i$, we see that $\{\mathbf{u}_i\}$ is a basis of V, and one easily checks that it is an orthogonal basis.

The second statement is an immediate consequence. □

In general situations, we have following theorems which you will verify in the problem set.

Theorem 1.29. *Let V be a non-zero finite dimensional vector over a field K with $\mathrm{char}(K) \neq 2$. Suppose V is equipped with a symmetric bilinear form. Then V has an orthogonal basis.*

Theorem 1.30. *Let V be a finite dimensional vector space with a non-degenerate symmetric bilinear form. Then for any subspace $W \subset V$, we have*

$$\dim(W) + \dim(W^{\perp}) = \dim(V).$$

2. Lecture 2

2.1. **Determinants and characteristic polynomial.**

We study some important properties of square matrices and associated linear transformations.

Definition 2.1. Let n be a positive integer. A **determinant function** is a map $D : \mathrm{Mat}_{n\times n}(K) \to K$ 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 column when the other $n-1$ columns are fixed.

(2) $D(A) = 0$ whenever A has two equal columns.
(3) $D(\mathrm{Id}_n) = 1$.

Theorem 2.2. *For any positive integer n, a determinant function exists and is unique. It can be given by*

$$\det((a_{ij})) := \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) \prod_{i=1}^{n} a_{i\sigma(i)}. \tag{2.1}$$

Here, the sum is over all permutations σ of the set $\{1, 2, \ldots, n\}$ and $\mathrm{sgn}(\sigma)$ *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 $A, B \in \mathrm{Mat}_{n\times n}(k)$ with A invertible, then $\det(ABA^{-1}) = \det(B)$.

Remark 2.3. 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

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

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

Definition 2.4. For an $n \times n$ matrix A, its **adjoint** $\mathrm{adj}(A)$ is the $n \times n$ matrix whose (i, j)-entry is equal to the (j, i)-cofactor of A.

From Laplace's formula, we deduce

Proposition 2.5. $A \cdot \mathrm{adj}(A) = \det(A) \cdot \mathrm{Id}_n = \mathrm{adj}(A) \cdot A$.

Theorem/Definition 2.6. *Let V be a finite dimensional vector space over K and $T : V \to V$ a linear map. For each basis $\mathcal{B}$ of V, the number* $\det({}_{\mathcal{B}}[T]_{\mathcal{B}})$ *is independent of $\mathcal{B}$; we call this number the* ***determinant*** *of T and denote it* $\det(T)$.

In the following, we are given an n-dimensional vector space V over K and a linear map $T : V \to V$.

Definition 2.7. The **characteristic polynomial** of T is

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

Using the definition (2.1) of the determinant and its basic properties, it is easy to see that the following hold:

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

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

Theorem 2.8 (Cayley–Hamilton)**.** *Every linear map satisfies its characteristic polynomial:* $\mathrm{char}_T(T) = 0$.

Proof. Let A be the matrix associated with T (for some choice of basis of V), and write $\mathrm{char}_A(x) = \det(x \cdot \mathrm{Id} - A) = x^n + a_{n-1}x^{n-1} + \cdots + a_0$. We can write

$$\mathrm{adj}(x \cdot \mathrm{Id} - A) = B_{n-1}x^{n-1} + \cdots + B_1 x + B_0$$

for some $n \times n$ matrices $B_0, \ldots, B_{n-1}$. Since

$$(x \cdot \mathrm{Id} - A) \cdot \mathrm{adj}(x \cdot \mathrm{Id} - A) = \det(x \cdot \mathrm{Id} - A) \cdot \mathrm{Id}$$

by Proposition 2.5, we have a system of equations

$$B_{n-1} = \mathrm{Id},\ -AB_{n-1} + B_{n-2} = a_{n-1}\mathrm{Id},\ -AB_{n-2} + B_{n-3} = a_{n-2}\mathrm{Id}, \ldots, -AB_0 = a_0\mathrm{Id}.$$

From this recursive system of equations, we obtain

$$A^n + a_{n-1}A^{n-1} + \cdots + a_0\mathrm{Id} = 0.$$

□

Since $K[x]$ is a principal ideal domain, we then have:

Definition/Proposition 2.9. *The* ***minimal polynomial*** *of T, written $\min_T(x)$, is the monic polynomial of least degree satisfied by T. If $f(x) \in K[x]$ is a polynomial such that $f(T) = 0$, then $\min_T(x)$ divides $f(x)$ in $K[x]$. In particular, $\min_T(x)$ divides $\mathrm{char}_T(x)$.*

Conversely, by considering an algebraic closure of K and eigenvalues of T, we deduce

Proposition 2.10. $\mathrm{char}_T(x)$ *divides some power of* $\min_T(x)$*, i.e., the roots of* $\mathrm{char}_T(x)$ *are precisely the roots of* $\min_T(x)$ *(but the multiplicities may be different).*

In the case K is algebraically closed, $\min_T(x)$ can be characterized via Jordan canonical form, which will be discussed in Section 3.2.

2.2. **Eigenvalues and eigenvectors.** For this section, we assume K is algebraically closed.

Definition 2.11. A scalar $\lambda \in K$ is said to be an **eigenvalue** of T if there exists a *nonzero* vector $\mathbf{v} \in V$ such that

$$T\mathbf{v} = \lambda\mathbf{v} \qquad \text{or equivalently} \qquad (T - \lambda \cdot \mathrm{Id})\mathbf{v} = \mathbf{0}. \tag{2.2}$$

Any nonzero vector $\mathbf{v}$ satisfying (2.2) is called an **eigenvector** of T corresponding to λ. For an eigenvector λ, the associated **eigenspace** is the subspace V_λ of V consisting of all vectors satisfying (2.2). A **generalized eigenvector corresponding to** λ is any $\mathbf{v} \in V$ satisfying

$$(T - \lambda \cdot \mathrm{Id})^j\mathbf{v} = 0 \qquad \text{for some } j > 0. \tag{2.3}$$

For an eigenvalue λ of T, the associated **generalized eigenspace** is the set U_λ of all $\mathbf{v} \in V$ which satisfy (2.3). It is easy to see that U_λ is a subspace of V. The **geometric multiplicity** $m_{\mathrm{geom}}(\lambda)$ of an eigenvalue λ of T is by definition $\dim(V_\lambda)$. The **algebraic multiplicity** $m_{\mathrm{alg}}(\lambda)$ of an eigenvalue λ is its multiplicity as a root of $\mathrm{char}_T(x)$. 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.

The following structural result can be checked easily.

Theorem 2.12. *Let V and T be as above.*

1. *The eigenvalues of T are precisely the roots of* $\mathrm{char}_T(x)$.
2. $m_{\mathrm{geom}}(\lambda) = \dim(V_\lambda) \le \dim_k(U_\lambda) = m_{\mathrm{alg}}(\lambda)$ *with equality for all λ if and only if T is diagonalizable.*
3. *Each U_λ is stable under T; i.e. the restriction $T\big|_{U_\lambda}$ 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 $\sum_{i=1}^r U_{\lambda_i} \to V$ induces an isomorphism*

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

We end this section with the following useful result in determining the growth rate of A^k as $k \to \infty$.

Theorem 2.13 (Perron-Frobenius)**.** *Suppose $K = \mathbf{C}$ and A is an $n \times n$ matrix whose entries are positive real numbers. Then A has a unique eigenvector with positive entries (up to a multiplication by a positive scalar), and the corresponding eigenvalue has algebraic multiplicity one and is strictly greater than the absolute value of any other eigenvalue.*

You will verify this theorem in the problem set. As a preliminary, we recall the following fact from real analysis.

Definition 2.14. A **metric space** is a set X with a function $\delta : X \times X \to [0, \infty)$ satisfying

1. $\delta(x, y) = 0$ if and only if $x = y$;
2. $\delta(x, y) = \delta(y, x)$ for all $x, y \in X$;
3. $\delta(x, y) + \delta(y, z) \ge \delta(x, z)$ for all $x, y, z \in X$.

A metric space is **complete** if every Cauchy sequence converges to a limit in X. A function on a metric space $f : X \to X$ is **contractive** if for any $x \neq y$, we have $\delta(f(x), f(y)) \le c\delta(x, y)$ for some fixed constant c with $0 < c < 1$.

Lemma 2.15. *A contractive function on a complete metric space has a unique fixed point.*

3. Lecture 3

3.1. **Spectral theorem.** We study an important structural result on an inner product space. Let V be a finite dimensional vector space over $K = \mathbf{R}$ or $K = \mathbf{C}$.

Definition 3.1. An **inner product** on V is a map $\langle \cdot, \cdot \rangle : V \times V \to \mathbf{C}$ satisfying

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

for all $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ 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. An **inner product space** is a vector space V equipped with an inner product.

Suppose V is an inner product space. The inner product on V gives a map

$$V \to V^* \qquad \text{via} \qquad \mathbf{v} \mapsto \langle \cdot, \mathbf{v} \rangle, \tag{3.1}$$

which is injective as any inner product is non-degenerate. It follows that (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 "semi-linear over complex conjugation").

Definition 3.2. Let $T : V \to V$ a linear map. The **adjoint** of T is the linear map $T^* : V \to V$ defined by the composition

$$V \xrightarrow[\simeq]{(3.1)} V^* \xrightarrow{T^\vee} V^* \xrightarrow[\simeq]{(3.1)^{-1}} V$$

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

$$\langle T\mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{v}, T^*\mathbf{w} \rangle$$

for all $\mathbf{v}, \mathbf{w} \in V$. We can check that the mapping

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

is semi-linear over complex conjugation (i.e. it is additive and satisfies $\Theta(af) = \overline{a}\Theta(f)$). Moreover, we have $T^{**} = T$ for any $T \in \mathrm{Hom}(V, V)$, and $(ST)^* = T^*S^*$ for any $T, S \in \mathrm{Hom}(V, V)$.

Remark 3.3. Let $\mathcal{B}$ be a basis of V. For $T : \mathrm{Hom}(V, V)$, we have

$$_{\mathcal{B}}[T]_{\mathcal{B}} = {}_{\mathcal{B}}\overline{[T^*]}^{t}_{\mathcal{B}}.$$

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

Definition 3.4. Let $T : V \to V$ be a linear map on an inner product space V. We say that T is:

(1) **Self adjoint** if $T^* = T$. When the base field is $\mathbf{R}$, we also say that T is **symmetric**, while when the base field is $\mathbf{C}$, we also 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 $\langle T\mathbf{v}, T\mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{v} \rangle$ for all $\mathbf{v} \in V$. When the base field is $\mathbf{R}$, we also say T is **orthogonal**; in the complex case we also say **unitary**.

Remark 3.5. It is easy to check that any eigenvalue of an Hermitian operator must be real, and any eigenvalue of a unitary operator must have absolute value 1.

Theorem 3.6 (Spectral Theorem)**.** *Let V be an inner product space and $T : V \to V$ be a linear map.*

(1) If the base field is $\mathbf{C}$, then V has an orthogonal 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 orthogonal basis consisting of eigenvectors for T if and only if T is self adjoint.

Corollary 3.7. *Let $T : V \to V$ 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.

For proving spectral theorem, we discuss here the case when T is self adjoint, and let you verify the full statement in the problem set. Suppose first that $K = \mathbf{R}$ and T is symmetric. We can deduce that every eigenvalue of T is real from T being symmetric. So there exists a real eigenvector $\mathbf{v}$ of T, and let $W = \operatorname{span}\{\mathbf{v}\}$. Then W is stable under T, and we can also check that $W^{\perp}$ is stable under T. By inducting on the dimension, we deduce that V has an orthogonal basis consisting of eigenvectors for T. The Hermitian case is similar.

3.2. **Jordan canonical form.** In general, we cannot expect a given square matrix to be diagonalizable even when the base field K is algebraically closed. However, we still have a nice way to decompose a matrix into certain blocks which holds quite generally. Suppose first that K is algebraically closed.

Definition 3.8. 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.

Let V be a finite dimensional vector space over K, and let $T : V \to V$ a linear map. For an eigenvalue λ of T and $\mathbf{v}$ in the generalized eigenspace U_λ, if s is the smallest positive integer such that $(T - \lambda \cdot \mathrm{Id})^s \mathbf{v} = \mathbf{0}$, we can check $\mathbf{v}, (T - \lambda \cdot \mathrm{Id})\mathbf{v}, \ldots, (T - \lambda \cdot \mathrm{Id})^{s-1}\mathbf{v}$ are linearly independent, and form a **cyclic** subspace of V which is T-invariant. We can show further that V can be decomposed as a direct sum of cyclic subspaces, and deduce:

Theorem 3.9 (Jordan Canonical Form)**.** *Let $T : V \to V$ 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****)*

$$\begin{pmatrix} J_{\lambda_1,s_1} & & & \\ & J_{\lambda_2,s_2} & & \\ & & \ddots & \\ & & & J_{\lambda_t,s_t} \end{pmatrix}. \tag{3.2}$$

Here, $\lambda_1, \ldots, \lambda_t$ 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.

Remark 3.10. $\min_T(x)$ can be determined by considering the maximal size of the Jordan blocks for each eigenvalue appearing in the Jordan canonical form.

Now, we consider the case of a general base field K. We cannot have a "direct" generalization of Theorem 3.9 since an eigenvalue may not lie in K. So we have to introduce some abstract notions generalizing the situation when K is algebraically closed.

Definition 3.11. For a linear map $T : V \to V$, we say

(1) T is **nilpotent** if $T^s = 0$ for some positive integer s.
(2) T is **semisimple** if any T-stable subspace $W \subseteq V$ admits a T-stable complement.

Definition 3.12. A field K is **perfect** if every algebraic extension of K is separable.

Remark 3.13. K is perfect if and only if either $\mathrm{char}(K) = 0$ or $\mathrm{char}(K) = p$ and $K^p = K$.

Theorem 3.14 (Jordan-Chevalley Decomposition)**.** *Suppose K is perfect and $T : V \to V$ is a linear map on a finite dimensional vector space. Then T can be written uniquely as a sum $T = T_{ss} + T_n$ such that $T_{ss} : V \to V$ is semisimple, $T_n : V \to V$ is nilpotent, and $T_{ss}T_n = T_nT_{ss}$.*

You will study this theorem in detail in the problem set, and discuss what happens in the case when K is not perfect.