Math 115B Running Notes

Brendan Connelly

January to March 2025

Dual Spaces

Proposition. For any two vector spaces V and W, the set of linear functions from V to W, denoted $\mathcal{L}(V, W)$, is a vector space with the following operations:

- **Addition:** For $S, T \in \mathcal{L}(V, W)$, define $(S + T)(v) = S(v) + T(v)$ for all $v \in V$.
- **Scalar Multiplication:** For $\alpha \in k$ and $T \in \mathcal{L}(V, W)$, define $(\alpha T)(v) = \alpha(T(v))$ for all $v \in V$.

Definition (Dual Vector Space)**.** For any vector space V, the *dual vector space* V^* is the set of all linear functions from V to k, denoted:

$$V^* := \mathcal{L}(V, k).$$

Definition. Given a vector space V, the elements of the dual vector space V^* are known as *linear functionals*.

Proposition. For any basis $\beta = \{v_1, \dots, v_d\}$ of a finite-dimensional vector space V, there exists an isomorphism

$$[-]_\beta : \mathcal{L}(V, V) \to k^{d \times d}$$

defined by the formula:

$$[T]_\beta = \begin{pmatrix} [T(v_1)]_\beta & [T(v_2)]_\beta & \cdots & [T(v_d)]_\beta \end{pmatrix},$$

for any $T \in \mathcal{L}(V, V)$.

Theorem (2.20)**.** Let V and W be finite-dimensional vector spaces over $\mathbb{K}$, and let $\beta = \{v_1, \dots, v_m\}$ be a basis for V, and $\gamma = \{w_1, \dots, w_n\}$ be a basis for W. Then there exists a linear isomorphism:

$$[-]_{\gamma,\beta} : \mathcal{L}(V, W) \to k^{n \times m}.$$

Corollary. If V is a vector space of dimension m and W is a vector space of dimension n, then:

$$\dim(\mathcal{L}(V, W)) = mn.$$

Corollary. If V is a finite-dimensional vector space, then:

$$\dim(V^*) = \dim(V).$$

Definition (Dual Basis Vector)**.** Given a finite-dimensional vector space V and a basis $\beta = \{v_1, \dots, v_d\}$ of V, the i-th *dual basis vector* is the linear functional $v_i^* : V \to k$ defined by the formula:

$$v_i^*(\vec{v}) = \alpha_i,$$

where $\vec{v} = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_d v_d$ is the representation of $\vec{v} \in V$ in terms of the basis β.

Theorem (2.24)**.** If V is a finite-dimensional vector space and $\beta = \{v_1, \dots, v_d\}$ is a basis for V, then the set $\{v_1^*, v_2^*, \dots, v_d^*\}$ is a basis for V^*. Moreover, for any $f \in V^*$, we have:

$$f = f(v_1)v_1^* + f(v_2)v_2^* + \cdots + f(v_d)v_d^*.$$

Definition (Dual Basis)**.** If V is a finite-dimensional vector space with basis $\beta = \{v_1, \dots, v_d\}$, the basis

$$\beta^* = \{v_1^*, \dots, v_d^*\}$$

is called the *dual basis.*

Theorem (2.25)**.** Let V, W be finite-dimensional vector spaces over k. Let $\mathcal{B}$ be a basis for V and $\mathcal{C}$ be a basis for W. Let $T : V \to W$ be a linear transformation. Then, $T^* : W^* \to V^*$, given by

$$T^*(g) = g \circ T \quad \text{for any } g \in W^*,$$

is linear. Moreover,

$$[T^*]_{\mathcal{C}^*}^{\mathcal{B}^*} = \left([T]_{\mathcal{B}}^{\mathcal{C}}\right)^t.$$

Theorem (2.26)**.** Let V be a finite-dimensional vector space. Then, the map

$$\Psi : V \to (V^*)^*$$

given by the formula

$$\Psi(v)(f) = f(v),$$

for $v \in V$ and $f \in V^*$, is a linear isomorphism.

Remark (On Dual and Double Dual Vector Spaces)**.**

$$\dim(V) = \dim(V^*) = \dim((V^*)^*),$$

We also know that $V^* = \mathcal{L}(V, k)$, so

$$(V^*)^* = \mathcal{L}(V^*, k).$$

The main point is that for any finite-dimensional vector space V, there exists an isomorphism between V and its double dual $(V^*)^*$. This isomorphism does not depend on the choice of a basis.

Remark. *First equality easily shown in hw. Second equality is easy for one inclusion*

$$\ker(T^*) = (\operatorname{Im} T)^\circ.$$

$$\operatorname{Im}(T^*) = (\ker T)^\circ.$$

If W is a subspace of V where V may actually be infinite dimensional, then

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

1 Eigenvalues, Eigenvectors, & Diagonalizability

Definition (Eigenvector/Eigenvalue)**.** Assume $T : V \to V$ is linear, where V is a vector space. We say $v \in V$ is an eigenvector of T with eigenvalue $\lambda \in k$ if

$$T(v) = \lambda v \quad \text{and} \quad v \neq 0.$$

Definition (Diagonalizable)**.** A linear transformation $T : V \to V$, where V is a finite-dimensional vector space, is said to be *diagonalizable* if there exists a basis $\mathcal{B}$ of V such that the matrix $[T]_\mathcal{B}$ is a diagonal matrix.

Theorem (5.1)**.** Let V be a finite-dimensional vector space and $T : V \to V$ a linear transformation. Then, T is diagonalizable if and only if there exists a basis $\mathcal{B} = \{v_1, \ldots, v_d\}$ for V such that for any $i \in \{1, 2, \ldots, d\}$, v_i is an eigenvector of T with some eigenvalue $\lambda_i \in k$.

Theorem (5.2)**.** T has $\lambda \in k$ as an eigenvalue if and only if $\ker(T - \lambda I) \neq \{0\}$.

Corollary. λ is an eigenvalue of T if and only if

$$\det(T - \lambda I) = 0.$$

Definition (Determinant)**.** The determinant $\det(A) \in k$ is defined as given in the textbook on page 205.

Definition (Characteristic Polynomial)**.** The characteristic polynomial of $A \in k^{n\times n}$ is

$$\det(T - \lambda I) \in k[\lambda].$$

Definition (Determinant)**.** The determinant of a linear endomorphism $T : V \to V$ of a finite-dimensional vector space V is defined as

$$\det([T]_\mathcal{B}),$$

where $\mathcal{B}$ is a basis for V and $[T]_\mathcal{B}$ is the matrix representation of T with respect to $\mathcal{B}$.

Theorem (5.3)**.** The characteristic polynomial of T is a polynomial of degree n, where $n = \dim(V)$, and the coefficient on t^n is 1. More precisely, it is $(-1)^n$.

Corollary (Number of eigenvalues)**.** Because any polynomial $P_n(\lambda)$ can have at most n-roots (over any field), we conclude:

$$\text{If } \dim(V) = n, \text{ then } T \text{ has at most } n \text{ eigenvalues.}$$

Definition (Polynomial Splits)**.** A polynomial $p(t) \in k[t]$ splits over k if there exist $c, a_1, \ldots, a_d \in k$ such that

$$p(t) = c(t - a_1) \cdots (t - a_d).$$

Theorem (5.6)**.** [Diagonalizability and Splitting] If T is diagonalizable, then the characteristic polynomial of T splits over k.

Rmk: This is only a one-way implication. You can use the contrapositive to show that T is not diagonalizable.

Definition (Algebraic Multiplicity)**.** Given an eigenvalue λ of T, the *algebraic multiplicity* of λ is the largest positive integer j such that $(t - \lambda)^j$ divides the characteristic polynomial of T.

Definition (Eigenspace)**.** Given an eigenvalue λ of T, the *eigenspace* for λ is the span of its eigenvectors with eigenvalue λ. Denote this eigenspace by V_λ.

Example: $V_\lambda = \text{span}\{\text{eigenvectors of } T \text{ with eigenvalue } \lambda\}$.

Definition (Geometric Multiplicity)**.** Given an eigenvalue λ of T, its *geometric multiplicity* is $\dim(V_\lambda)$.

Theorem (5.7)**.** If λ is an eigenvalue for T and has algebraic multiplicity m, then

$$\dim(V_\lambda) \leq m.$$

Equivalently,

$$\text{geo}(\lambda) \leq \text{alg}(\lambda).$$

Theorem (5.8)**.** T is diagonalizable if and only if for every eigenvalue λ_i of T, the geometric multiplicity of λ_i equals its algebraic multiplicity:

$$\text{geo}(\lambda_i) = \text{alg}(\lambda_i).$$

2 Cayley-Hamilton

Theorem (Cayley-Hamilton)**.** If $A \in k^{n\times n}$ and the characteristic polynomial of A is

$$(-1)^d t^d + a_{d-1}t^{d-1} + \cdots + a_1 t + a_0,$$

where $a_{d-1}, \ldots, a_0 \in k$, then

$$(-1)^d A^d + a_{d-1}A^{d-1} + \cdots + a_1 A + a_0 I = 0,$$

where 0 is the zero matrix.

Definition (Nilpotent Maps/Matrices)**.** T is nilpotent if $T^k = 0$ for some $k \in \mathbb{N}$.

Proposition (Eigenvalues of Nilpotent Matrices)**.** Let T be a nilpotent linear map (or matrix). Then, the only eigenvalue of T is 0.

Proof. Suppose T is nilpotent, so there exists some positive integer k such that $T^k = 0$.

Let λ be an eigenvalue of T with corresponding eigenvector $v \neq 0$, i.e.,

$$T(v) = \lambda v.$$

Applying T^k to v, we get:

$$T^k(v) = T^{k-1}(T(v)) = T^{k-1}(\lambda v) = \lambda T^{k-1}(v).$$

Repeating this process iteratively, we find:

$$T^k(v) = \lambda^k v.$$

However, since $T^k = 0$, it follows that:

$$T^k(v) = 0 = \lambda^k v.$$

Because $v \neq 0$, we must have $\lambda^k = 0$. The only solution in the field of scalars (typically $\mathbb{C}$ or $\mathbb{R}$) is $\lambda = 0$. Therefore, the only eigenvalue of a nilpotent matrix T is 0. □

Corollary (Cayley-Hamilton for Linear Transformations)**.** Let $T : V \to V$ be a linear transformation for V a finite-dimensional vector space over a field k, and let

$$p(t) = (-1)^{\dim(V)} t^d + a_{d-1}t^{d-1} + \cdots + a_1 t + a_0$$

be the characteristic polynomial for T.

Then, in $\mathcal{L}(V, V)$,

$$p(T) = (-1)^{\dim(V)} T^d + a_{d-1}T^{d-1} + \cdots + a_1 T + a_0 I = 0.$$

Definition (T-invariant Subspace)**.** A subspace W of V is called T-*invariant* if $T(W) \subseteq W$, i.e.,

$$\{T(w) \mid w \in W\} \subseteq W.$$

Proposition. If v_1, v_2 are eigenvectors for T with possibly different eigenvalues, then $\operatorname{span}\{v_1, v_2\}$ is T-invariant.

More generally, if $v_1, \ldots, v_k$ are eigenvectors for T, then $\operatorname{span}\{v_1, \ldots, v_k\}$ is T-invariant.

Definition (T-cyclic subspace)**.** The T-cyclic subspace at a vector $v \in V$ is defined as

$$\operatorname{span}\{v, T(v), T^2(v), \ldots\} = \operatorname{span}\{T^j(v) : j \in \mathbb{Z}_{\geq 0}\}.$$

Remark (Infinite Span)**.** Recall that if S is a possibly infinite set of vectors in a vector space W, then

$$\operatorname{span}(S) = \left\{ \sum_{j \in S} \alpha_j s_j : \alpha_j \in \mathbb{R}, \text{ and only finitely many } \alpha_j \neq 0 \right\}.$$

This allows us to pick or combine finitely many vectors from S in linear combinations.

Theorem (5.21)**.** Let T be a linear operator on a finite-dimensional vector space V, and let W denote the T-cyclic subspace of V generated by a nonzero vector $v \in V$. Let $k = \dim(W)$. Then:

(a) $\{v, T(v), T^2(v), \ldots, T^{k-1}(v)\}$ is a basis for W.

(b) If $a_0 v + a_1 T(v) + \cdots + a_{k-1} T^{k-1}(v) + T^k(v) = 0$, then the characteristic polynomial of $T|_W$ is

$$f(t) = (-1)^k \left(a_0 + a_1 t + \cdots + a_{k-1} t^{k-1} + t^k\right).$$

Theorem (5.20)**.** Let T be a linear operator on a finite-dimensional vector space V, and let W be a T-invariant subspace of V. Then the characteristic polynomial of $T|_W$ divides the characteristic polynomial of T.

Theorem (Characteristic Polynomial of a Cyclic Subspace)**.** Let $T : V \to V$ be a linear operator on a finite-dimensional vector space V, and let $W \subseteq V$ be the T-cyclic subspace generated by a vector $v \in V$. If $\{v, T(v), T^2(v), \ldots, T^{n-1}(v)\}$ is a basis for W, then:

1. The matrix representation of $T|_W$ with respect to this basis is

$$[T]_{\mathcal{B}} = \begin{pmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & \cdots & 0 & -a_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -a_{n-1} \end{pmatrix},$$

   where $T^n(v) = -a_0 v - a_1 T(v) - \cdots - a_{n-1} T^{n-1}(v)$.

2. The characteristic polynomial of $T|_W$ is

$$f_{T|W}(t) = (-1)^n (a_0 + a_1 t + \cdots + a_{n-1} t^{n-1} + t^n).$$

Proposition (Characteristic Polynomial Decomposition)**.** Let $W \subseteq V$ be a T-invariant subspace of a vector space V. Then the characteristic polynomial of T satisfies the relation

$$f_T(t) = p(t) \cdot q(t),$$

where $p(t)$ is the characteristic polynomial of $T|_W$, and $q(t)$ is the characteristic polynomial of $T|_{V/W}$.

Remark. To see this explicitly, take any basis $\mathcal{B}_1$ for W and extend it to a basis $\mathcal{B}_2 = \mathcal{B}_1 \cup \mathcal{Q}$ for the entire vector space V. Then, the matrix representation of T with respect to $\mathcal{B}_2$ is block-upper triangular:

$$[T]_{\mathcal{B}_2} = \begin{pmatrix} [T|_W]_{\mathcal{B}_1} & A_1 \\ 0 & A_2 \end{pmatrix},$$

where $A_1 = 0$ if and only if $\operatorname{span}(\mathcal{Q})$ is T-invariant. The determinant of $tI_V - [T]_{\mathcal{B}_2}$ decomposes as

$$\det(tI_V - [T]_{\mathcal{B}_2}) = \det(tI_W - [T|_W]_{\mathcal{B}_1}) \cdot \det(tI_{V/W} - A_2),$$

which corresponds to the factorization $f_T(t) = p(t) \cdot q(t)$.

Proposition. If V is T cyclic, then S commutes with T if and only if $S = g(T)$ for polynomial g.

Proof. Assume that V is a cyclic T-module, generated by a vector v, so that

$$V = \operatorname{span}\{v, Tv, T^2v, \dots\}.$$

Let $m(x)$ be the minimal polynomial of T with respect to v, i.e., the monic polynomial of smallest degree such that

$$m(T)v = 0.$$

Then every vector in V can be expressed as a polynomial in T of degree less than $\deg m$ applied to v.

($\implies$) Suppose that S commutes with T, i.e., $ST = TS$.

Since V is generated by v, the action of S is determined by its action on v. Let us express Sv as

$$Sv = p(T)v,$$

for some polynomial $p(x)$.

We need to show that $S = p(T)$. For any non-negative integer k,

$$ST^kv = T^kSv = T^kp(T)v = p(T)T^kv.$$

On the other hand,

$$ST^kv = p(T)T^kv.$$

This equality holds for all k, and since $\{v, Tv, T^2v, \dots\}$ spans V, it follows that

$$S = p(T).$$

Therefore, S is a polynomial in T.

($\impliedby$) Conversely, suppose that $S = g(T)$ for some polynomial $g(x)$.

Since polynomials in T commute with T, we have

$$ST = g(T)T = Tg(T) = TS.$$

Thus, S commutes with T.

Combining both directions, we conclude that S commutes with T if and only if $S = g(T)$ for some polynomial g. $\square$

3 Inner Product Spaces and Adjoints

Definition (Standard Inner Product (Real))**.** Let $\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \mathbf{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in \mathbb{R}^n$. The standard inner product (dot product) is defined as:

$$\langle \mathbf{x}, \mathbf{y} \rangle = x_1 y_1 + \cdots + x_n y_n \in \mathbb{R}.$$

Remark. For $\mathbf{x} \in \mathbb{R}^n$, $\langle \mathbf{x}, \mathbf{x} \rangle = x_1^2 + \cdots + x_n^2 = \|\mathbf{x}\|^2$.

Definition (Standard Inner Product (Complex))**.** Let $\mathbf{z} = \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix}, \mathbf{w} = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} \in \mathbb{C}^n$. The standard inner product of vectors is defined as:

$$\langle \mathbf{z}, \mathbf{w} \rangle = \overline{z_1} w_1 + \cdots + \overline{z_n} w_n.$$

Remark. For any $\mathbf{w} \in \mathbb{C}^n$, $\langle \mathbf{w}, \mathbf{w} \rangle \in \mathbb{R}_{\geq 0}$, and it is equal to zero if and only if $\mathbf{w} = \mathbf{0}$.

Remark. In $\mathbb{R}^2$, the cosine of the angle θ between two vectors $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ and $\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$ is given by:

$$\cos\theta = \frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\|\mathbf{x}\| \|\mathbf{y}\|}.$$

Definition (Inner Product)**.** An *inner product* on an F-vector space V is the data of a scalar $\langle v, w \rangle \in F$ for every $v, w \in V$, such that the following properties hold:

1. **Linearity in the First Variable:** For all $v_1, v_2, w \in V$ and $\alpha_1, \alpha_2 \in F$,

$$\langle \alpha_1 v_1 + \alpha_2 v_2, w \rangle = \alpha_1 \langle v_1, w \rangle + \alpha_2 \langle v_2, w \rangle.$$

2. **Conjugate Symmetry:** For all $v, w \in V$,

$$\langle v, w \rangle = \overline{\langle w, v \rangle}.$$

3. **Positive Definiteness:** If $v \in V$ is a nonzero vector, then

$$\langle v, v \rangle > 0,$$

where the result is a positive real number (even if $F = \mathbb{C}$).

The inner product is a map $\langle \cdot, \cdot \rangle : V \times V \to F$.

Definition (Inner Product Space)**.** An *inner product space* is the data of a vector space V over F and an inner product on V.

Corollary (Orthogonal Basis Expansion)**.** Assume V is an inner product space (IPS), and let $\{v_1, \ldots, v_d\}$ be an orthonormal basis (ONB) for V. Then, for any $v \in V$, we have:

$$v = \langle v, v_1 \rangle v_1 + \langle v, v_2 \rangle v_2 + \cdots + \langle v, v_d \rangle v_d.$$

Theorem (Gram-Schmidt Process)**.** Let $S = \{v_1, \ldots, v_m\}$ be a set of a finite number of vectors in an inner product space (IPS) V. Then, there exists an orthonormal set of vectors $\{v_{r+1}, \ldots, v_d\} \subset V$ such that $\{v_1, \ldots, v_r, v_{r+1}, \ldots, v_d\}$ forms an orthonormal basis (ONB) for V.

Definition (Orthogonal Complement)**.** Given a subspace W of an inner product space V, its *orthogonal complement* is defined as:

$$W^\perp = \{v \in V \mid \langle v, w \rangle = 0 \text{ for all } w \in W\}.$$

Theorem. If W is any subspace of a finite-dimensional inner product space (IPS) V, then:

$$V = W \oplus W^\perp,$$

where $W^\perp$ is the orthogonal complement of W.

Proof (Sketch). Use the Gram-Schmidt process to construct an orthonormal basis $\{w_1, w_2, \ldots, w_r\}$ for W. Then, extend this basis to an orthonormal basis for V by adding vectors from $W^\perp$. The resulting basis $\{w_1, \ldots, w_r, w_{r+1}, \ldots, w_d\}$ satisfies the decomposition $V = W \oplus W^\perp$. □

Theorem. Fix an inner product space V. The function $P : V \to V^*$, defined by:

$$P(v)(w) = \langle w, v \rangle \quad \text{for } v, w \in V,$$

is a bijection. However, P is not linear over $\mathbb{C}$ if V is a complex vector space.

Theorem. Let $T : V \to V$ be a linear endomorphism of a finite-dimensional inner product space (IPS) V. Then, there exists a unique linear map $T^* : V \to V$ such that:

$$\langle T(v), w \rangle = \langle v, T^*(w) \rangle \quad \text{for all } v, w \in V.$$

This function T^* is linear.

Definition (Adjoint or Conjugate Transpose)**.** The linear operator $T^* : V \to V$ is called the *conjugate transpose* or *adjoint* of T.

Theorem. If we choose a basis $\mathcal{B} = \{v_1, \ldots, v_d\}$ for a finite-dimensional inner product space (IPS) V, then the conjugate transpose of T satisfies:

$$[T^*]_\mathcal{B} = ([T]_\mathcal{B})^\dagger,$$

where $([T]_\mathcal{B})^\dagger = ([T]_\mathcal{B})^T$ is the transpose (or conjugate transpose in the complex case) of the matrix representation of T in the basis $\mathcal{B}$.

Theorem. Let $T, U : V \to V$ be linear operators on a finite-dimensional inner product space (IPS) V. Then, the following properties hold:

1. $(U + T)^* = U^* + T^*$,
2. If $\alpha \in F$, then $(\alpha T)^* = \overline{\alpha} T^*$,
3. $(U \circ T)^* = T^* \circ U^*$,
4. $(T^*)^* = T$,
5. $I^* = I$, where I is the identity operator.

Remark. These properties hold because for any composition of operators, $(AB)^* = B^* A^*$, which can be verified using the definition of the adjoint:

$$\langle (AB)v, w \rangle = \langle v, (AB)^* w \rangle = \langle v, B^* A^* w \rangle.$$

Definition (Normal Operator)**.** A linear operator $T : V \to V$ on a finite-dimensional inner product space (IPS) V is called *normal* if:

$$TT^* = T^*T.$$

Theorem (Properties of Normal Operators)**.** Let T be a normal operator on a finite-dimensional inner product space V. Then:

1. $\|T(v)\| = \|T^*(v)\|$ for any $v \in V$,
2. $T - \alpha I$ is normal for any $\alpha \in F$.