Notes for MAT 310, Fall 2025

Joseph Helfer

Last updated: December 3, 2025

Contents

3B Null Spaces and Ranges

- Throughout these notes, we follow the Axler's notational convention: $\mathbf{F}$ is one of the fields $\mathbf{R}$ or $\mathbf{C}$, and V and W are vector spaces over $\mathbf{F}$.

Null Space and Injectivity

- Definition 3.11: the **null space**, or **kernel** of a linear map $T \in \mathcal{L}(V, W)$ is the subset of V defined by null $T := \{v \in V \mid Tv = 0\}$.
- Proposition 3.13: null T is a subspace of V.
- Examples 3.12:
  - If $T\colon V \to W$ is the zero map, then null $T = V$.
  - If $D \in \mathcal{L}(\mathcal{P}(\mathbf{R}))$ is the differentiation map, $Dp = p'$, then $D \subset \mathcal{P}(\mathbf{R})$ is the set of constant functions.
    - This is because (by the mean value theorem!) any differentiable function $f\colon \mathbf{R} \to \mathbf{R}$ with zero derivative is constant.
  - If $T \in \mathcal{L}(\mathcal{P}(\mathbf{R}))$ is the map defined by $(Tp)(x) = x^2 p(x)$, then null $T = \{0\}$.
    - This is because, if $x^2 p(x) = 0$ for all x, then $p(x) = 0$ for all $x \neq 0$, and hence also $p(0) = 0$ by continuity of p.
  - If $T \in \mathcal{L}(\mathbf{F}^\infty)$ is the map defined by $T(x_i)_{i=1}^\infty = (x_{i+1})_{i=1}^\infty$, then null $T = \{(a, 0, 0, \ldots) \mid a \in \mathbf{F}\}$.
- Definition 3.14: a function $f\colon X \to Y$ between sets X and Y is **injective** if $f(x_1) = f(x_2)$ implies $x_1 = x_2$ for all $x_1, x_2 \in X$.
- Proposition 3.15: $T \in \mathcal{L}(V, W)$ is injective if and only if null $T = \{0\}$.

Range and Surjectivity

- Definition 3.16: the **range**, or **image**, of a function $f\colon X \to Y$ between sets X and Y is the subset of Y defined by range $f := \{f(x) \mid x \in X\}$.
- Proposition 3.18: for any $T \in \mathcal{L}(V, W)$, range T is a subspace of W.
- Examples 3.17:
  - If $T\colon V \to W$ is the zero map, then range $T = \{0\}$
  - If $D \in \mathcal{L}(\mathcal{P}(\mathbf{R}))$ is the differentiation operator $Dp = p'$, then range $D = \mathcal{P}(\mathbf{R})$.
    - This is because for each $q \in \mathcal{P}(\mathbf{R})$, there is $p \in \mathcal{P}(R)$ with $p' = q$, namely $p(x) = \int_0^x q(t)\,\mathrm{d}t$.
- Definition 3.19: a function $f\colon X \to Y$ between sets X and Y is **surjective**, or **onto**, if range $f = Y$.
  - In other words, f is surjective if for each $y \in Y$, there is an $x \in X$ with $f(x) = y$.

- Example 3.20: the differentiation map $D \in \mathcal{L}(\mathcal{P}_5(\mathbf{R}))$ is not surjective but the differentiation map $D \in \mathcal{L}(\mathcal{P}_5(\mathbf{R}), \mathcal{P}_4(\mathbf{R}))$ is.
  - The lesson: whether a function f is surjective depends on what we are considering the codomain of f to be.

Fundamental Theorem of Linear Maps

- Theorem 3.21 ("fundamental theorem of linear maps" or "rank-nullity theorem" or "dim-sum theorem"): for any $T \in \mathcal{L}(V, W)$, if V is finite-dimensional, then $\operatorname{range} T$ is finite-dimensional, and

$$\dim V = \dim \operatorname{null} T + \dim \operatorname{range} T.$$

- Corollary 3.22: given $T \in \mathcal{L}(V, W)$, if $\dim V > \dim W$, then T is not injective.
- Example 3.23: define a linear map $T\colon \mathbf{F}^4 \to \mathbf{F}^3$ by some big crazy formula; then without doing any calculations, we can conclude that T is not injective.
- Corollary 3.24: given $T \in \mathcal{L}(V, W)$, if $\dim V < \dim W$, then T is not surjective.
- Reformulation of homogeneous systems of linear equations in terms of linear maps:
  - Consider a homogeneous system of m linear equations in n variables over $\mathbf{F}$ (*homogeneous* means that the constant terms are all zero)

$$\begin{cases} A_{1,1}x_1 & + & \cdots & + & A_{1,n}x_n & = & 0 \\ & & & & & \vdots & \\ A_{m,1}x_1 & + & \cdots & + & A_{m,n}x_n & = & 0. \end{cases}$$

  - We associate to this the linear map $T\colon \mathbf{F}^n \to \mathbf{F}^m$ given by

$$T(x_1, \ldots, x_n) = \Big(\sum_{k=1}^{n} A_{1,k}x_k, \ldots, \sum_{k=1}^{n} A_{m,k}x_k\Big)$$

.

  - We then have that $(x_1, \ldots, x_n)$ is a solution to the system if and only if $(x_1, \ldots, x_n) \in \operatorname{null} T$.

- Corollary 3.26: a homogeneous system of linear equations with more variables than equations has nonzero solutions.
- Reformulation of inhomogeneous systems of linear equations in terms of linear maps:
  - Now consider a possibly inhomogeneous system

$$\begin{cases} A_{1,1}x_1 & + & \cdots & + & A_{1,n}x_n & = & c_1 \\ & & & & & \vdots & \\ A_{m,1}x_1 & + & \cdots & + & A_{m,n}x_n & = & c_m. \end{cases}$$

  - Defining $T\colon \mathbf{F}^n \to \mathbf{F}^m$ as above, we see that $(x_1, \ldots, x_n)$ is a solution iff $T(x_1, \ldots, x_n) = (c_1, \ldots, c_m)$.
  - In particular, the system has a solution iff $(c_1, \ldots, c_m) \in \operatorname{range} T$.

- Corollary 3.28: if $m > n$ (i.e., if there are more equations than variables), then in the above system, no matter what the coefficients $A_{i,j}$ are, the constant terms c_k can be chosen so that the system has no solution.

3C Matrices

Representing a Linear Map by a Matrix

- Definition 3.29: an **m-by-n matrix over $\mathbf{F}$** is a family $A = (A_{j,k})_{1\leq j\leq m, 1\leq k\leq n}$ of elements of $\mathbf{F}$ indexed by pairs (j,k) with $j \in \{1,\dots,m\}$ and $k \in \{1,\dots,n\}$.
  - We typically represent a matrix as a rectangular array:
  $$A = \begin{pmatrix} A_{1,1} & \cdots & A_{1,n} \\ \vdots & & \vdots \\ A_{m,1} & \cdots & A_{m,n} \end{pmatrix}.$$
  - Note that in $A_{j,k}$, the *first* index j refers to the *row*, and the *second* entry k refers to the *column*.
  - Definition 3.39: we write $\mathbf{F}^{m,n}$ for the set m-by-n matrices.
- Definition 3.31: given $T \in \mathcal{L}(V,W)$ and bases $\mathcal{B}_1 = (v_1,\dots,v_n)$ of V and $\mathcal{B}_2 = (w_1,\dots,w_m)$, the **matrix of** T, denoted $\mathcal{M}(T)$, is the m-by-n matrix A with entries $A_{j,k}$ defined by
  $$Tv_k = A_{1,k}w_1 + \cdots + A_{m,k}w_m$$
  - Again, the matrix $\mathcal{M}(T)$ depends on the chosen bases $\mathcal{B}_1$ and $\mathcal{B}_2$; to make this dependence explicit in the notation, we may write $\mathcal{M}(T,\mathcal{B}_1,\mathcal{B}_2)$ in place of $\mathcal{M}(T)$.
  - Note that it is the k-th *column* of $\mathcal{M}(T)$ which is associated to the basis vector v_k (namely, its entries are the coefficients of the vector Tv_k with respect to the basis $\mathcal{B}_2$).
- Definition: we recall (from Example 2.27) that the **standard basis** of $\mathbf{F}^n$ is the basis $\mathcal{B}_{\text{std}} = (e_1,\dots,e_n)$ such that e_k has a 1 in the k-th entry and 0 elsewhere.
  - When we are considering a basis for $\mathbf{F}^n$, we will assume it is the standard basis unless stated otherwise.
- Example 3.32: if $T\colon \mathbf{F}^2 \to \mathbf{F}^3$ is given by $T(x,y) = (x+3y, 2x+5y, 7x+9y)$, then
  $$\mathcal{M}(T) = \begin{pmatrix} 1 & 3 \\ 2 & 5 \\ 7 & 9 \end{pmatrix}$$
  - More generally, recall from Example 3.3 that every $T \in \mathcal{L}(\mathbf{F}^n, \mathbf{F}^m)$ has the form
  $$T(x_1,\dots,x_n) = (A_{1,1}x_1 + \cdots + A_{1,n}x_n, \cdots, A_{m,1}x_1 + \cdots + A_{m,n}x_n)$$
  for some matrix A.
  - Then $\mathcal{M}(T) = A$.
    * Indeed, we see immediately that
    $$Te_k = (A_{1,k}, \cdots, A_{n,k}) = A_{1,k}e_1 + \cdots + A_{n,k}e_k.$$

- Example 3.33: the matrix of $D \in \mathcal{L}(\mathcal{P}_3(\mathbf{R}), \mathcal{P}_2(\mathbf{R}))$ with respect to the standard monomial bases is
$$\mathcal{M}(D) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}.$$

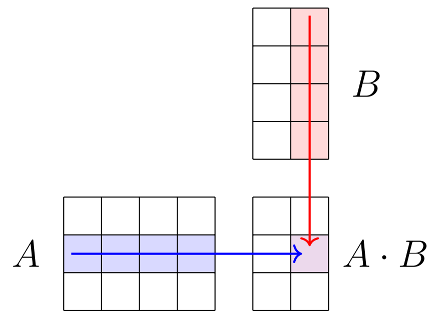
Addition and Scalar Multiplication of Matrices

- For the rest of chapter 3C, fix finite-dimensional vector spaces U, V, and W and bases $(u_1, \ldots, u_p)$, $(v_1, \ldots, v_n)$, and $(w_1, \ldots, w_m)$ for each of them.
- Definitions 3.34: the **sum** of two matrices $A, B \in \mathbf{F}^{m,n}$ is defined by $(A+B)_{j,k} = A_{j,k} + B_{j,k}$.
- Proposition 3.35: for any $S, T \in \mathcal{M}(V, W)$, we have $\mathcal{M}(S + T) = \mathcal{M}(S) + \mathcal{M}(T)$.
  - Exercise!
- Definitions 3.36: the **product of a scalar $\boldsymbol{\lambda}$ and a matrix $\boldsymbol{A}$** is defined by $(\lambda A)_{j,k} = \lambda \cdot A_{j,k}$.
- Proposition 3.35: for any $T \in \mathcal{M}(V, W)$, we have $\mathcal{M}(\lambda T) = \lambda \mathcal{M}(S)$.
  - Exercise!
- Proposition 3.40: $\mathbf{F}^{m,n}$, with addition and scalar multiplication as defined above, is a vector space of dimension $m \cdot n$.

Matrix multiplication

- We look for a formula for the matrix $\mathcal{M}(ST)$ of the composition ST of the linear maps S and T. Let $A = \mathcal{M}(S)$ and $B = \mathcal{M}(T)$. Then for each k:
$$(ST)u_k = S\Big(\sum_{r=1}^{n} B_{r,k} v_r\Big) = \sum_{r=1}^{n} B_{r,k} S v_r = \sum_{r=1}^{n} B_{r,k} \sum_{j=1}^{m} A_{j,r} w_j = \sum_{j=1}^{m} \Big(\sum_{r=1}^{n} A_{j,r} B_{r,k}\Big) w_j$$
  - Conclusion: $\mathcal{M}(ST)$ is the matrix C with entries $C_{k,j} = \sum_{r=1}^{n} A_{j,r} B_{r,k}$.
- Definition 3.41: the **matrix product** of $A \in \mathbf{F}^{m,n}$ and $B \in \mathbf{F}^{n,p}$ is the matrix $AB \in \mathbf{F}^{m,p}$ with entries
$$(AB)_{j,k} = \sum_{j=1}^{m} \Big(\sum_{r=1}^{n} A_{j,r} B_{r,k}\Big).$$
  - This is only defined when the number of columns of A equals the number of rows of B.

- By the calculation we did above, we immediately have (Proposition 3.43): $\mathcal{M}(ST) = \mathcal{M}(S)\mathcal{M}(T)$.
- Here is how I was taught to do matrix multiplication:



- That is, you line up A and B corner-to-corner; then, to compute a given entry in $A \cdot B$, you take the row in A to the left of it, the column in B above it, and you take their dot product.

- Definition 3.44: given $A \in \mathbf{F}^{m,n}$, we write $A_{j,\bullet} \in \mathbf{F}^{1,n}$ for the j-th row of A, and $A_{\bullet,k} \in \mathbf{F}^{m,1}$ for its k-th column.
  - Thus $(A_{j,\bullet})_{1,r} = A_{j,r}$ and $(A_{\bullet,k})_{s,1} = A_{s,k}$.
- Proposition 3.46: given $A \in \mathbf{F}^{m,n}$ and $B \in \mathbf{F}^{n,p}$, we have $(AB)_{j,k} = A_{j,\bullet}B_{\bullet,k}$.
  - This is another way of expressing of the above picture.
  - Note that there is a slight type-mismatch in the equation: the left-hand side is a scalar, while the right-hand side is a 1-by-1 matrix. However, by a common abuse of notation, we often identify a 1-by-1 matrix C with its single entry $C_{1,1}$.
  - The proof follows from the definitions:

$$(A_{j,\bullet}B_{\bullet,k})_{1,1} = \sum_{r=1}^{n} (A_{j,\bullet})_{1,r}(B_{\bullet,k})_{r,1} = \sum_{r=1}^{n} A_{j,r}B_{r,k} = (AB)_{j,k}.$$

- Proposition 3.48: given $A \in \mathbf{F}^{m,n}$ and $B \in \mathbf{F}^{n,p}$, we have $(AB)_{\bullet,k} = A \cdot B_{\bullet,k}$.
  - That is, the columns of AB are obtained by multiplying A by the columns of B.
  - The proof:

$$\left((AB)_{\bullet,k}\right)_{j,1} = (AB)_{j,k} = \sum_{r=1}^{n} A_{j,r}B_{r,k} = \sum_{r=1}^{n} A_{j,r}(B_{\bullet,k})_{r,1} = (A \cdot B_{\bullet,k})_{j,1}.$$

- Proposition 3.50: given $A \in \mathbf{F}^{m,n}$ and $b \in \mathbf{F}^{n,1}$ with entries $b_1, \ldots, b_n$, we have $Ab = \sum_{k=1}^{n} b_k A_{\bullet,k}$.
  - That is, the product is a linear combination of the columns of A, whose coefficients are the entries in b.
  - The proof:

$$(Ab)_{j,1} = \sum_{k=1}^{n} A_{j,k}b_{k,1} = \sum_{k=1}^{n} b_k(A_{\bullet,k})_{j,1}.$$

- Example 3.49: the product of $\begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}$ and $\begin{pmatrix} 5 \\ 1 \end{pmatrix}$.
- Note that the previous two propositions have analogues involving rows instead of columns.
- Proposition 3.51: given $C \in \mathbf{F}^{m,c}$ and $R \in \mathbf{F}^{c,n}$:
  (a) $(CR)_{\bullet,k} = \sum_{j=1}^{c} R_{j,k} C_{\bullet,j}$
  (b) $(CR)_{j,\bullet} = \sum_{k=1}^{c} C_{j,k} R_{k,\bullet}$
  - That is, the k-th column of CR is a linear combination of the columns of C, with coefficients given by the k-th column of R; and the j-th row of CR is a linear combination of the rows of R, with coefficients given by the j-th row of C.

Column-Row Factorization and Rank of a Matrix

- Definition 3.52: given $A \in \mathbf{F}^{m,n}$:
  The **column rank** of A is the dimension of $\operatorname{span}(A_{\bullet,1}, \ldots, A_{\bullet,n}) \subset \mathbf{F}^{m,1}$. The **row rank** of A is the dimension of $\operatorname{span}(A_{1,\bullet}, \ldots, A_{1,\bullet}) \subset \mathbf{F}^{1,n}$.
  - Note that the column rank is at most m and the row rank is at most n.
  - Soon, we will see that the column and row rank are always equal!
- Example 3.53: column and row rank of $A = \begin{pmatrix} 4 & 7 & 1 & 8 \\ 3 & 5 & 2 & 9 \end{pmatrix}$.
- Definition 3.54: the **transpose** of $A \in \mathbf{F}^{m,n}$ is the matrix $A^{\mathrm{t}} \in \mathbf{F}^{n,m}$ defined by $(A^{\mathrm{t}})_{k,j} = A_{j,k}$.
  - That is, the transpose is obtained by interchanging the columns and rows of A.
- Example 3.55: the transpose of $A = \begin{pmatrix} 5 & -7 \\ 3 & 8 \\ -4 & 2 \end{pmatrix}$.
- The transpose has nice algebraic properties:
  - $(A + B)^{\mathrm{t}} = A^{\mathrm{t}} + B^{\mathrm{t}}$
  - $(\lambda A)^{\mathrm{t}} = \lambda \cdot A^{\mathrm{t}}$
  - $(AC)^{\mathrm{t}} = C^{\mathrm{t}} A^{\mathrm{t}}$
- Proposition 3.56 ("column-row factorization"): given $A \in \mathbf{F}^{m,n}$ with column rank c, there exists $C \in \mathbf{F}^{m,c}$ and $R \in \mathbf{F}^{c,n}$ with $A = CR$.
- Theorem 3.57: for any $A \in \mathbf{F}^{m,n}$, the row rank of A equals its column rank.

- Definition 3.58: the **rank** of a matrix $A \in \mathbf{F}^{m,n}$ is the column of rank of A (or equivalently by (3.57) its row rank).

3D Invertibility and Isomorphisms

Invertible linear maps

- Definition: we recall (from Example 3.3) that for any vector space V, we denote by $I \in \mathcal{L}(V,V)$—or to be more explicit, by I_V—the **identity operator** on V, defined by $Iv = v$.
  - Of course, we can also talk about the identity function on any set X, but we usually only use the notation I when considering a vector space; in general, we may use the notation id or id_X for the identity function on X.
- Recollection on bijections and inverses:
  - Recall that given a function $f\colon X \to Y$ between sets X and Y, an **inverse** of f is a function $g\colon Y \to X$ with $g \circ f = \mathrm{id}_X$ and $f \circ g = \mathrm{id}_Y$—i.e., with $g(f(x)) = x$ for $x \in X$ and $f(g(y)) = y$ for $y \in Y$. If an inverse exists, we say that f is **invertible**. The basic facts about inverses are:
  - 1. If f is invertible, it has a unique inverse; we then denote the unique inverse of f by $f^{-1}\colon Y \to X$.
    - Indeed, if g_1 and g_2 are both inverses of f, then $g_1 = g_1 \circ \mathrm{id}_X = g_1 \circ f \circ g_2 = \mathrm{id}_Y \circ g_2 = g_2$.
  - 2. f is invertible iff it is a bijection (i.e., an injection and a surjection).
    - Indeed, if f is invertible, then $y = f(f^{-1}(y))$ for all $y \in Y$, hence f is surjective, and $f(x_1) = f(x_2)$ implies $x_1 = f^{-1}(f(x_1)) = f^{-1}(f(x_2)) = x_2$, hence f is injective.
    - And if f is bijective, then for each $y \in Y$, there exists by surjectivity some $x_y \in X$ with $f(x_y) = y$, hence we may define a function $g\colon Y \to X$ by $g(y) = x_y$; then g is the inverse of f, since $f(g(y)) = f(y_x) = y$ for $y \in Y$, and thus $f(g(f(x))) = f(x)$ for $x \in X$, hence $g(f(x)) = x$ by injectivity.
- Proposition 3.63: if a linear map $T\colon V \to W$ is invertible, then its inverse $T^{-1}\colon W \to V$ is also linear.
  - Indeed, given $v, w \in W$ and $\lambda \in \mathbf{F}$, we have

    $$T(\lambda T^{-1}v + T^{-1}w) = \lambda(T^{-1}Tv) + T^{-1}Tw = \lambda v + w$$

    by linearity of T and hence, applying T^{-1} to both sides,

    $$\lambda T^{-1}v + T^{-1}w = T^{-1}(\lambda v + w)$$

    as required.
- Proposition 3.65: if V and W are finite-dimensional and $\dim V = \dim W$ then T is invertible $\iff$ T is injective $\iff$ T is surjective.
- Example 3.64: when V is not finite-dimensional, then $T \in \mathcal{L}(V)$ can be injective or surjective, but fail to be invertible.
  - The multiplication by x^2 map in $\mathcal{L}(\mathcal{P}(\mathbf{R}))$ is injective but not surjective.
  - The backward shift map in $\mathcal{L}(\mathcal{P}(\mathbf{R}))$ is surjective but not injective.

- Example 3.67: for any $q \in \mathcal{P}(\mathbf{R})$, there is $p \in \mathcal{P}(R)$ with $\frac{\mathrm{d}^2}{\mathrm{d}x^2}\big((x^2+5x+7)p\big) = q$.
- Proposition 3.68: given $S \in \mathcal{L}(V, W)$ and $T \in \mathcal{L}(W, V)$, if $\dim V = \dim W$, then $ST = I \iff TS = I$.

Isomorphic Vector Spaces

- Definition 3.69: an **isomorphism** is an invertible linear map. V and W are **isomorphic**, denoted $V \cong W$, if there is an isomorphism $T\colon V \to W$.
- Proposition 3.70: if V and W are finite-dimensional, then $V \cong W$ iff $\dim V = \dim W$.
- Corollary: each finite-dimensional vector space V is isomorphic to $\mathbf{F}^n$ for some n (namely $n = \dim V$).
- Proposition 3.71: if $v_1, \ldots, v_n$ is a basis for V and $w_1, \ldots, w_m$ is a basis for W, then the map $\mathcal{L}(V, W) \to \mathbf{F}^{m,n}$ taking T to $\mathcal{M}(T)$ is an isomorphism.
- Corollary 3.72: it V and W are finite-dimensional, then $\dim \mathcal{L}(V, W) = \dim V \cdot \dim W$.

Linear Maps Thought of as Matrix Multiplication

- Definition 3.73: given a basis $\mathcal{B} = (v_1, \ldots, v_n)$ of V, the **matrix of $\boldsymbol{v}$**, or **coordinate column vector of $\boldsymbol{v}$** is the n-by-1 matrix

$$\mathcal{V}(v) = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix},$$

where $b_1, \ldots, b_n \in \mathbf{F}$ are the unique scalars such that

$$v = b_1 v_1 + \cdots + b_n v_n.$$

  - Such a unique list of scalars exists by (2.28) (criterion for basis).
  - Of course, the matrix of v depends on the chosen basis $\mathcal{B}$, though this is hidden from the notation. The intended basis will often be clear from context. To make it more explicit, we may write $\mathcal{V}(v, \mathcal{B})$.
  - The textbook uses the notation $\mathcal{M}(v)$ for the matrix of a vector, but it will later prove to be confusing to have the same notation for the matrix of a vector and the matrix of a linear map.
- Remark: $\mathcal{V}\colon V \to \mathbf{F}^{n,1}$ is an isomorphism.

- Example 3.74:
  - The matrix of the polynomial $2 - 7x + 5x^3 + x^4$ with respect to the monomial basis of $\mathcal{P}_4(\mathbf{R})$.
  - For $v \in \mathbf{F}^n$, the matrix $\mathcal{V}(v)$ with respect to the standard basis is simply v itself, written as a column; we may sometimes conflate elements $v \in \mathbf{F}^n$ with the corresponding elements $\mathcal{V}(v) \in \mathbf{F}^{n,1}$.
- Proposition 3.75: given bases for finite-dimensional vector spaces V and W, we have $\mathcal{M}(T)_{\bullet,k} = \mathcal{V}(Tv_k)$.
  - This is just a restatement of the definition of $\mathcal{M}(T)$.
- Proposition 3.76: given bases for finite-dimensional vector spaces V and W, we have $\mathcal{V}(Tv) = \mathcal{M}(T)\mathcal{V}(v)$ for all $v \in V$.
  - Note that this property determines $\mathcal{M}(T)$: is the unique matrix $A \in \mathbf{F}^{\dim V, \dim W}$ such that $A\mathcal{V}(v) = \mathcal{V}(Tv)$ for all $v \in V$.
    * This follows from the general fact that for $A, B \in \mathbf{F}^{m,n}$, if $A \cdot w = B \cdot w$ for all $w \in \mathbf{F}^{n,1}$, then $A = B$.
- Proposition 3.78: if V and W are finite-dimensional and $T \in \mathcal{L}(V, W)$, then $\dim \operatorname{range} T = \operatorname{rank} \mathcal{M}(T)$.

Change of basis

- Definition 3.79: the **identity matrix** $I \in \mathbf{F}^{n,n}$ is the matrix with $I_{j,j} = 1$ for all j and $I_{j,k} = 0$ when $j \neq k$.
  - When we want to be explicit about the dimension n, we may write I_n instead of I.
  - Of course, this is the same symbol we are using for the identity operator, but they can usually be distinguished from the context; for example, we have the equation $\mathcal{M}(I) = I$ for any vector space V of dimension n and any basis, which really means $\mathcal{M}(I_V) = I_n$.
  - We have $AI = A$ and $IB = B$ for any $A \in \mathbf{F}^{m,n}$ and $B \in \mathbf{F}^{n,p}$.
- Definition 3.80: $A \in \mathbf{F}^{n,n}$ is **invertible** if there exists $B \in \mathbf{F}^{n,n}$ with $AB = BA = I$; we call B the **inverse** of A and denote it by A^{-1}.
  - As with functions, the inverse of a matrix is uniquely determined.
  - We have $(A^{-1})^{-1} = A$.
  - If $A, C \in \mathbf{F}^n$ are both invertible, then $(AC)^{-1} = C^{-1}A^{-1}$.
- Proposition 3.81: if $\mathcal{B}_1, \mathcal{B}_2, \mathcal{B}_3$ are bases for vector spaces U, V, W, then for any $T \in \mathcal{L}(U, V)$ and $S \in \mathcal{L}(V, W)$, we have $\mathcal{M}(S, \mathcal{B}_2, \mathcal{B}_3)\mathcal{M}(T, \mathcal{B}_1, \mathcal{B}_2) = \mathcal{M}(ST, \mathcal{B}_1, \mathcal{B}_3)$.
  - This is just a restatement of a calculation we already did when defining the matrix of a linear map.

- Proposition 3.82: if $\mathcal{B}_1$ and $\mathcal{B}_2$ are bases for V, then $\mathcal{M}(I_V, \mathcal{B}_1, \mathcal{B}_1)$ and $\mathcal{M}(I_V, \mathcal{B}_2, \mathcal{B}_1)$ are invertible, and are inverses of each other.
- Proposition 3.84: suppose $T \in \mathcal{L}(V)$ and $\mathcal{B}_1, \mathcal{B}_2$ are bases of V. If $A = \mathcal{M}(T, \mathcal{B}_1)$ and $B = \mathcal{M}(T, \mathcal{B}_2)$ and $C = \mathcal{M}(I, \mathcal{B}_1, \mathcal{B}_2)$, then $A = C^{-1}BC$.
- Proposition 3.86: given any basis $\mathcal{B}$ of V and any $T \in \mathcal{L}(V)$, we have $\mathcal{M}(T^{-1}) = \mathcal{M}(T)^{-1}$.
  - Exercise!

3E Products and Quotients of Vector Spaces

Products of vector spaces

- Definition 3.87: the **product** of vector spaces $V_1, \ldots, V_m$ over $\mathbf{F}$ is the vector space $V_1 \times \cdots \times V_m = \{(v_1, \ldots, v_m) \mid v_1 \in V_1, \cdots, v_m \in V_m\}$ with addition and scalar multiplication given entry wise.

- Example 3.90: $\mathbf{R}^2 \times \mathbf{R}^3$ is isomorphic (but not *equal*) to $\mathbf{R}^5$.

- Proposition 3.92: $\dim(V_1 \times \cdots \times V_m) = \dim V_1 + \cdots + \dim V_m$.

- Proposition 3.93: given subspaces $V_1, \ldots, V_m$ of V, if we define the linear map $\Gamma\colon V_1 \times \cdots \times V_m \to V_1 + \cdots + V_m$ by $\gamma(v_1, \ldots, v_m) = v_1 + \cdots + v_m$, then $V_1 + \cdots + V_m$ is a direct sum iff Γ is injective.

  - Since Γ is clearly surjective, we could equivalently say "iff Γ is an isomorphism".

- Corollary 3.94: $V_1 + \cdots + V_m$ is a direct sum iff $\dim(V_1 + \cdots + V_m) = \dim V_1 + \cdots + \dim V_m$.

Quotient spaces

- Definitions 3.95 and 3.97: given $v \in V$ and $U \subseteq V$, we write $v + U$ for the subset $v + U = \{v + u \mid u \in U\}$; we say that $v + U$ is a **translate** of v.

- Examples 3.96 and 3.98: translate of lines and planes

- Definitions 3.99 and 3.102: given a subspace $U \subseteq V$, we define the **quotient space** to be vector space V/U whose underlying set is the set of translates of U, thus $V/U = \{v + U \mid v \in V\}$, and whose addition and scalar multiplication are given by $(v + U) + (w + U) = (v + w) + U$ and $\lambda(v + U) = \lambda v + U$.

  - It is not immediately obvious that these operations are well-defined: we need to check that if $v + U = v' + U$ and $w + U = w' + U$, then $(v + w) + U = (v' + w') + U$, and likewise with scalar multiplication. We next prove that this is so.
  - Once we verify that the operations are well-defined, you should verify that V/U indeed forms a vector space.

- Lemma 3.101: any two translates of a subspace $U \subseteq V$ are either equal or disjoint, and $v + U = w + U$ iff $v - w \in U$.

  - In other words, V/U is the quotient of V by the equivalence relation $\sim$ given by $v \sim w \iff v - w \in U$.

- Proposition 3.103: the addition and scalar multiplication on V/U are well-defined.

- Definition 3.104: the **quotient map** $\pi\colon V \to V/U$ is the linear map defined by $\pi(v) = v + U$.
  - You should verify that it is indeed linear.
- Proposition 3.105: null $\pi = U$, and if V is finite-dimensional, then $\dim V/U = \dim V - \dim U$.
- Definition 3.106: given $T \in \mathcal{L}(V,W)$, we denote by $\widetilde{T}$ the linear map $\widetilde{T}\colon V/(\text{null } T) \to W$ defined by $\widetilde{T}(v + \text{null } T) = Tv$.
  - You should check this is well-defined, i.e., that $v + \text{null } T = v' + \text{null } T$ implies $Tv = Tv'$.
  - You should also verify that this is linear.
- Proposition 3.107: given $T \in \mathcal{L}(V,W)$, we have
  (a) $\widetilde{T} \circ \pi = T$
  (b) $\widetilde{T}$ is injective
  (c) $\operatorname{range} \widetilde{T} = \operatorname{range} T$
  (d) $V/\text{ null } T \cong \operatorname{range} T$
  - (d) is one version of the "first isomorphism theorem" from abstract algebra

3F Duality

Dual Space and Dual map

- Definitions 3.108 and 3.110: a **linear functional** on V (also known as a **linear form** or **covector**) is a linear map $V \to \mathbf{F}$. The **dual space** of V, denoted V', is the vector space $V' = \mathcal{L}(V, \mathbf{F})$ of linear functionals on V.
- Examples 3.109:
  - An example of a linear functional $\varphi\colon \mathbf{R}^3 \to \mathbf{R}$ on $\mathbf{R}^3$ is $\varphi(x, y, z) = 4x - 5y + 2z$.
  - More generally, a linear functional φ on $\mathbf{F}^n$ is a function of the form $\varphi(x_1, \ldots, x_n) = a_1x_1 + \cdots + a_nx_n$; it is expressions of this kind that were traditionally (i.e., prior to the introduction of abstract vector spaces) called "linear forms".
  - An example of a linear functional φ on $\mathcal{P}(\mathbf{R})$ is $\varphi(p) = \int_0^1 p(x)\,\mathrm{d}x$. It is examples of this kind that give rise to the name "linear functional": in general, the word "functional" is often used to refer to a function whose argument is itself a function.
- Proposition 3.111: $\dim V = \dim V'$.
- Definition 3.112: given a basis $v_1, \ldots, v_n$ of V, the **dual basis** is the basis $\varphi_1, \ldots, \varphi_n$ of V' defined by $\varphi_j(v_k) = \delta_{j,k}$.
  - Here, the notation $\delta_{j,k}$ is called the **Kronecker delta symbol**, and is defined to be 1 if $j = k$ and 0 if $j \neq k$.
  - The φ_j are well-defined by the linear map lemma (3.4).
  - We will soon see that this is indeed a basis of V'.
- Example 3.113: the dual basis $\varphi_1, \ldots, \varphi_n \in (\mathbf{F}^n)'$ of the standard basis $e_1, \ldots, e_n$ of $\mathbf{F}^n$ is given by $\varphi_j(x_1, \ldots, x_n) = x_j$.
- Proposition 3.114: given a basis $v_1, \ldots, v_n$ of V, if $\varphi_1, \ldots, \varphi_n$ is the dual basis, then $v = \sum_{j=1}^n \varphi_j(v) \cdot v$ for all $v \in V$.
  - This is a direct generalization of Example 3.113: it says that j-th element of the dual basis of V' is the function giving the j-th coordinate of a vector with respect to the given basis of V.
- Lemma for 3.116: given a basis $v_1, \ldots, v_n$ of V, if $\varphi_1, \ldots, \varphi_n$ is the dual basis, $\varphi = \sum_{k=1}^n \varphi(v_k) \cdot \varphi_k$ for each $\varphi \in V'$.
  - Proof: Each side of the equation is a linear function $V \to \mathbf{R}$; to see that they agree, it suffices to check that they agree on the basis vectors v_j.
  - But $(\sum_{k=1}^n \varphi(v_k) \cdot \varphi_k)(v_j) = \sum_{k=1}^n \varphi(v_k) \cdot \varphi_k(v_j) = \sum_{k=1}^n \varphi(v_k)\delta_{j,k} = \varphi(v_j)$, as required.
- Proposition 3.116: The dual basis is indeed a basis.
  - Proof: The above lemma implies that the dual basis is a spanning set, and it has $\dim V'$ elements by (3.111), therefore its a basis by (2.42).

- Definition 3.118: given a linear map $T \in \mathcal{L}(V, W)$, the **dual map** $T' : W' \to V'$ is the linear map defined by $T'(\varphi) = \varphi \circ T \in V'$ for $\varphi \in W'$.
  - Here, we are using that the composition of linear maps is linear: we have $T \in \mathcal{L}(V, W)$ and $\varphi \in \mathcal{L}(W, \mathbf{F})$, and hence $\varphi \circ T \in \mathcal{L}(V, \mathbf{F})$.
  - That T' is *itself* a linear map from W' to V' follows from the distributivity of composition of linear maps (3.8), and from the following general scalar multiplication property for composition of linear maps: if $S_1 \in \mathcal{L}(U, V)$ and $S_2 \in \mathcal{L}(V, W)$ and $a \in \mathbf{F}$, then $(aS_2)S_1 = S_2(aS_1) = a(S_2 S_1)$.

- Proposition 3.120: the function taking $T \in \mathcal{L}(V, W)$ to $T' \in \mathcal{L}(W, V)$ is a linear map $\mathcal{L}(V, W) \to \mathcal{L}(W', V')$. Moreover, given $S \in \mathcal{L}(W, U)$ and $T \in \mathcal{L}(V, W)$, we have $(ST)' = T'S' \in \mathcal{L}(V, U)$.

- Remark: in Exercise 3.F.32, you will prove that there is a canonical linear map $V \to V''$ from any vector space V to its double dual, and that if V is finite-dimensional, this map is an isomorphism. For this reason, a finite-dimensional vector space is often considered to be "the same as" its double dual.

Null Space and Range of Dual of Linear map

- Definition 3.121: given any subset $U \subset V$, the **annihilator** of U, denoted U^0, is the subset of V' defined by $U^0 = \{\varphi \in V' \mid \varphi(u) = 0 \text{ for all } u \in U\}$.

- Example 3.123: the annihilator of $U = \operatorname{span}(e_1, e_2) \subset \mathbf{R}^5$.

- Proposition 3.124: for any subset $U \subseteq V$, U^0 is a subspace of V'.

- Example: if $V \subseteq \mathbf{R}^n$ is a subspace, then under the isomorphism $\mathbf{R}^n \xrightarrow{\sim} (\mathbf{R}^n)'$ taking the standard basis to its dual basis, the annihilator $V^0 \subseteq (\mathbf{R}^n)'$ corresponds to the *orthogonal complement* $V^\perp \subseteq \mathbf{R}^n$.
  - This illustrates one of the main utilities of the concept of dual space: it provides a substitute for such concepts as "orthogonality" even in the absence of an inner product. (We will discuss inner products more later on).

- Proposition 3.125: if V is finite-dimensional and $U \subseteq V$ is a subspace, then $\dim U^0 = \dim V - \dim U$.

- Proposition 3.127: given a subspace $U \subseteq V$, we have $U^0 = \{0\} \iff U = V$ and $U^0 = V' \iff U = \{0\}$.

- Proposition 3.128: given $T \in \mathcal{L}(V, W)$, we have
  (a) $\operatorname{null} T' = (\operatorname{range} T)^0$
  (b) if $\dim V, \dim W < \infty$, then $\dim \operatorname{null} T' = \dim \operatorname{null} T + \dim W - \dim V$.

- Corollary 3.129: $T \in \mathcal{L}(V, W)$ is surjective iff T' injective
- Proposition 3.130: if $T \in \mathcal{L}(V, W)$ and $\dim V = \dim W < \infty$, then
  (a) $\dim \operatorname{range} T' = \dim \operatorname{range} T$
  (b) $\operatorname{range} T' = (\text{null } T)^0$.
- Corollary 3.131: if $\dim V, \dim W < \infty$, then $T \in \mathcal{L}(V, W)$ is injective iff T' is surjective.
  - Unlike in Corollary 3.129, we need to assume V and W are finite-dimensional.

Matrix of Dual of Linear map

- Lemma for 3.132: given a basis $\mathcal{B}$ for V and hence a dual basis $\mathcal{B}'$ for V', we have $\mathcal{V}(\varphi, \mathcal{B}') = \mathcal{M}(\varphi, \mathcal{B}, (1))^{\mathrm{t}}$ for all $\varphi \in V'$, where (1) is the standard basis for $\mathbf{F}$.
  - Let $v_1, \ldots, v_n$ be the elements of $\mathcal{B}$ and $\varphi_1, \ldots, \varphi_n$ be the elements of $\mathcal{B}'$.
  - By the lemma for 3.116, we have $\varphi = \sum_{j=1}^{n} \varphi(v_i)\varphi_i$.
  - Hence by definition $\mathcal{V}(\varphi, \mathcal{B}')$ is the vector with entries $\varphi(v_1), \ldots, \varphi(v_n)$.
  - But by definition (more precisely, by (3.75)), the entries of $\mathcal{M}(\varphi, \mathcal{B}, (1)) \in \mathbf{F}^{1,n}$ are precisely $\varphi(v_1), \ldots, \varphi(v_n)$.
- Proposition 3.132: given bases $\mathcal{B}_V$ and $\mathcal{B}_W$ for V and W, and hence dual bases $\mathcal{B}'_V$ and $\mathcal{B}'_W$ for V' and W', and given $T \in \mathcal{L}(V, W)$, we have $\mathcal{M}(T', \mathcal{B}'_W, \mathcal{B}'_V) = (\mathcal{M}(T, \mathcal{B}_V, \mathcal{B}_W))^{\mathrm{t}}$.
  - By (3.76), it suffices to show $\mathcal{M}(T)^{\mathrm{t}}\mathcal{V}(\varphi) = \mathcal{V}(T'\varphi)$ for all $\varphi \in V'$.
  - But using the above lemma, we have

    $$\mathcal{V}(T'\varphi) = \mathcal{M}(T'\varphi)^{\mathrm{t}} = \mathcal{M}(\varphi \circ T')^{\mathrm{t}} = (\mathcal{M}(\varphi)\mathcal{M}(T'))^{\mathrm{t}} = \mathcal{M}(T')^{\mathrm{t}}\mathcal{M}(\varphi)^{\mathrm{t}} = \mathcal{M}(T')^{\mathrm{t}}\mathcal{V}(\varphi),$$

    as desired.
- Corollary 3.133: this gives another proof of (3.57) (row rank = column rank) using (3.78) (column rank is dimension of range) and (3.130(a)) ($\dim \operatorname{range} T = \dim \operatorname{range} T'$).

4 Polynomials

More review of complex numbers

- Definition 4.1: the **real part** and **imaginary part** of a complex number $z = a + bi$ are defined as $\operatorname{Re} z = a$ and $\operatorname{Im} z = b$.
- Definition 4.2: the **conjugate** of a complex number $z = a + bi$ is $\bar{z} = a - bi$, and the **absolute value** of z is $|z| = \sqrt{a^2 + b^2}$.
- Proposition 4.4: some nice algebraic properties of complex numbers:
  - $z + \bar{z} = 2 \operatorname{Re} z$
  - $z - \bar{z} = 2(\operatorname{Im} z)i$
  - $z\bar{z} = |z|^2$
  - $\overline{w + z} = \bar{w} + \bar{z}$ and $\overline{wz} = \bar{w}\bar{z}$
  - $\bar{\bar{z}} = z$
  - $|\operatorname{Re} z|, |\operatorname{Im} z| \leq z$
  - $|\bar{z}| = |z|$
  - $|zw| = |z||w|$
  - $|w + z| \leq |w| + |z|$

Zeros of Polynomials

- Recall that a **polynomial of degree m over $\mathbf{F}$** is a function $p \colon \mathbf{F} \to \mathbf{F}$ for which there exist $a_0, \ldots, a_m \in \mathbf{F}$ with $a_m \neq 0$ such that $p(z) = a_0 + a_1 z + \cdots + a_m z^m$ for all $z \in \mathbf{F}$, and that we write $\mathcal{P}(\mathbf{F})$ for the vector space of polynomials over $\mathbf{F}$, and $\mathcal{P}_m(F)$ for the subspace consisting of polynomials of degree $\leq m$.
  - It is not immediate from the definition, but we will soon show that the coefficients $a_0, \ldots, a_m$ are uniquely determined by the function p, so that there is a bijection between the set of degree m polynomials, and the set of lists $a_0, \ldots, a_m \in \mathbf{F}$ with $a_m \neq 0$.
    * (Warning: this is true when $\mathbf{F}$ is $\mathbf{R}$ or $\mathbf{C}$, as is always the case for us, but it is not always true when $\mathbf{F}$ is an arbitrary *field*.)
- Definition 4.5: a **zero** (or **root**) of the polynomial $p \in \mathcal{P}(\mathbf{F})$ is an element $\lambda \in \mathbf{F}$ with $p(\lambda) = 0$.
- Proposition 4.6: if $p \in \mathcal{P}(\mathbf{F})$ has degree $m > 0$, then $\lambda \in \mathbf{F}$ is a root of p if and only if there exists a polynomial $q \in \mathcal{P}(\mathbf{F})$ of degree $m - 1$ such that for all $z \in \mathbf{F}$: $p(z) = (z - \lambda)q(z)$.
- Corollary 4.8: given a non-zero polynomial $p \in \mathcal{P}(\mathbf{F})$, if p has degree $m \geq 0$ with $p \neq 0$, then p has at most m zeros in $\mathbf{F}$.

- Corollary of 4.8: the coefficients of a polynomial are uniquely determined; that is, given $a_0, \ldots, a_m$ and $b_0, \ldots, b_n$ if $\sum_{j=0}^m a_j x^j = \sum_{k=0}^n b_k x^n$ for all $x \in \mathbf{F}$, then $a_j = b_k = 0$ for $j > m$ or $k > n$, and $a_j = b_j$ for all $j = 0, \ldots, \min(m, n)$.
  - In particular, the *degree* of a non-zero polynomial $p \in \mathcal{P}(\mathbf{F})$ is uniquely determined (it is the unique m for which there exists $a_0, \ldots, a_m$ with $a_m \neq 0$ and $p(x) = \sum_{j=0}^m a_j x^j$ for all x), and we denote it by $\deg p$.
    - A common and useful convention is that the zero polynomial has degree $-\infty$.

Division Algorithm for Polynomials

- Proposition 4.9: for any $p, s \in \mathcal{P}(\mathbf{F})$ with $s \neq 0$, there are unique polynomials q ("quotient") and r ("remainder") with $p = sq + r$ and $\deg r < \deg s$.

Factorization of Polynomials over C

- Theorem 4.12 (fundamental theorem of algebra, first version): every nonconstant polynomial $p \in \mathcal{P}(\mathbf{C})$ has a zero.

- Corollary 4.14 (fundamental theorem of algebra, second version): every nonconstant polynomial $p \in \mathcal{P}_m(\mathbf{C})$ can be factored as $p(z) = c(z-\lambda_1) \cdots (z-\lambda_m)$; the number $c \in \mathbf{C}$ is uniquely determined, and the list $\lambda_1, \ldots, \lambda_m \in \mathbf{C}$ is uniquely determined, up to reordering.

Factorization of Polynomials over R

- Proposition 4.14: suppose $p \in \mathcal{P}(\mathbf{C})$ has real coefficients. Then $p(\lambda) = 0$ implies $p(\bar{\lambda}) = 0$ for all $\lambda \in \mathbf{C}$.

- Proposition 4.15: for any $b, c \in \mathbf{R}$, there exists a factorization $x^2 + bx + c = (x - \lambda_1)(x - \lambda_2)$ with $\lambda_1, \lambda_2 \in \mathbf{R}$ iff $b^2 - 4c \geq 0$.

- Theorem 4.16: if $p \in \mathcal{P}(\mathbf{R})$ is nonconstant, then p has a factorization $p(x) = c \prod_{i=1}^m (x - \lambda_i) \prod_{j=1}^M (x^2 + b_j x + c_j)$, for a unique $c \in \mathbf{R}$ and a unique (up to reordering) list $\lambda_1, \ldots, \lambda_m, b_1, \ldots, b_M, c_1, \ldots, c_M \in \mathbf{R}$ with $b_j^2 - 4c_j < 0$ for $j = 1, \ldots, M$.
  - This implies a strengthening of Proposition 4.14: not only is $\bar{\lambda}$ a zero of $p \in \mathcal{P}(\mathbf{R})$ whenever λ is, but moreover the zeroes λ and $\bar{\lambda}$ have the same *multiplicity*
    - Here, the **multiplicity** of a root λ of p is the number of times the linear factor $(z - \lambda)$ appears in the decomposition of p into linear factors over $\mathbf{C}$.

5A Invariant subspaces

Eigenvalues

- Definition 5.1: an **operator** on V is a linear map $T \in \mathcal{L}(V)$
- Definition 5.2: a subspace $U \subseteq V$ is **invariant** under $T \in \mathcal{L}(V)$ if $Tu \in U$ for all $u \in U$.
  - In this case, we obtain an operator by restriction: $T|_U \in \mathcal{L}(U)$
- Example 5.3: $\mathcal{P}_m(\mathbf{R}) \subset \mathcal{P}(\mathbf{R})$ is invariant under the differentiation operator.
- Example 5.4: $\{0\}, V, \text{null } T, \text{range } T \subset V$ are all invariant under $T \in \mathcal{L}(V)$.
- Definitions 5.5 and 5.8: an **eigenvalue** of $T \in \mathcal{L}(V)$ is a number $\lambda \in \mathbf{F}$ for which there exists $v \in V$ with $v \neq 0$ and $Tv = \lambda v$; such a v is called an **eigenvector** of T.
- Remark: a 1-dimensional subspace $U \subseteq V$ is invariant iff $U = \text{span}(v)$ for some eigenvector v.
- Proposition 5.7: given $T \in \mathcal{L}(V)$ and $\lambda \in \mathbf{F}$, if V is finite-dimensional, then the following are equivalent:
  (a) λ is an eigenvalue of T
  (b) $T - \lambda I$ is not injective
  (c) $T - \lambda I$ is not surjective
  (d) $T - \lambda I$ is not bijective
- Example 5.9: the 90-degree rotation operator $T(w, z) = (-z, w)$ has no eigenvalues over $\mathbf{R}$ but has eigenvalues $\pm i$ over $\mathbf{C}$.
- Proposition 5.11: if $v_1, \ldots, v_n$ are eigenvectors of $T \in \mathcal{L}(V)$ with distinct eigenvalues, then they are linearly independent.
- Corollary 5.12: if V is finite-dimensional, then any $T \in \mathcal{L}(V)$ has at most $\dim V$ eigenvalues

Polynomials Applied to Operators

- Definition 5.13: given $T \in \mathcal{L}(V)$, we define T^m for $m \geq 0$ inductively by $T^0 = I$ and $T^{m+1} = T^m T$. If T is invertible we also define $T^m = (T^{-1})^{-m}$ for $m < 0$.
- Remark: you should check that the usual exponent laws $T^{m+n} = T^m T^n$ and $(T^m)^n = T^{mn}$ for all $m, n \in \mathbb{Z}_{\geq 0}$, and for all $m, n \in \mathbb{Z}$ when T is invertible.

- Definition 5.14: if $T \in \mathcal{L}(V)$ and $p \in \mathcal{P}(\mathbf{F})$ is given by $p(z) = \sum_{k=0}^{m} a_k z^k$, we define $p(T) \in \mathcal{L}(V)$ by $p(T) = \sum_{k=0}^{m} a_k T^k$.
- Remark: you should check that for a fixed $T \in \mathcal{L}(V)$, the function $\mathcal{P}(\mathbf{F}) \to \mathcal{L}(V)$ taking p to $p(T)$ is linear.
- Definition 5.16: the **product** $pq \in \mathcal{P}(\mathbf{F})$ of polynomials $p, q \in \mathcal{P}(\mathbf{F})$ is the polynomial given by $(pq)(z) = p(z)q(z)$.
- Proposition 5.17: given $p, q \in \mathcal{P}(\mathbf{F})$ and $T \in \mathcal{L}(V)$, we have $(pq)(T) = p(T)q(T)$ and $p(T)q(T) = q(T)p(T)$.
- Corollary of 5.17: $Tp(T) = p(T)T$ for any $p \in \mathcal{P}(\mathbf{F})$.
- Proposition 5.18: given $T \in \mathcal{L}(V)$ and $p \in \mathcal{P}(\mathbf{F})$, the subspaces null $p(T)$, range $p(T) \subseteq V$ are invariant under T.
  - More generally, given any $A \in \mathcal{L}(V)$ with $AT = TA$, we have that null A, range $A \subseteq V$ are invariant under T.

5B The minimal polynomial

Existence of Eigenvalues on Complex Vector Spaces

- Proposition 5.19: every operator on a finite-dimensional non-zero complex vector space has an eigenvalue
  - The proof uses the fundamental theorem of algebra. Exercise 5B.16 (in combination with (5.27) below) shows, conversely, how to deduce the fundamental theorem of algebra from (5.19).
- Example 5.20: (5.19) can fail in infinite dimensions. For example, the operator $T \in \mathcal{P}(\mathbf{C})$ given by $(Tp)(z) = zp(z)$ has no eigenvalue.

Eigenvalues and the Minimal Polynomial

- Definition 5.21: a polynomial is **monic** if its highest-degree coefficient equals 1.
  - Example: $z^7 + 9z^2 + 2$ is monic of degree 7.
- Definition: a **polynomial ideal** over $\mathbf{F}$ is a subspace $J \subseteq \mathcal{P}(\mathbf{F})$ with the property that $p \in J$ implies $pq \in J$ for all $p, q \in \mathcal{P}(\mathbf{F})$.
- Proposition 5.29: any non-zero polynomial ideal J contains a unique polynomial $p \neq 0$ of minimal degree. Moreover, $J = \{p \cdot q \mid q \in \mathcal{P}(\mathbf{F})\}$. ($p$ is called the **generator** of J)
  - Proof:
  - Let $p \in J$ be any non-zero element of minimal degree; any scalar multiple of p must still be in J, so we can assume p is monic.
  - Since J is an ideal, we automatically have $J \supseteq \{p \cdot q \mid q \in \mathcal{P}(\mathbf{F})\}$.
  - Conversely, given $u \in J$, by polynomial division, there exists $q, s \in \mathcal{P}(\mathbf{F})$ with $\deg s < \deg p$ and $p \cdot q + s = u$. Since J is an ideal, it follows that $s = u - p \cdot q \in J$, hence $s = 0$ by the minimality of p. Hence $u = p \cdot q$. This proves $J \subseteq \{p \cdot q \mid q \in \mathcal{P}(\mathbf{F})\}$.
  - Finally, to see that p is unique, note that any other monic polynomial u of least degree must be of the form $u = p \cdot q$, hence q must be of degree 0 (a constant), and hence $q = 1$ since u and p are both monic.
- Proposition: if V is finite-dimensional and $T \in \mathcal{L}(V)$, the set $J_T \subseteq \mathcal{P}(\mathbf{F})$ of polynomials p with $p(T) = 0$ is a non-zero ideal.
  - The proof that J is an ideal is a straightforward exercise.
  - That J is non-zero follows from the fact that the $\dim \mathcal{L}(V)$ is finite-dimensional, hence there must be some linear combination $c_0 I + c_1 T + c_2 T^2 + \cdots + c_{\dim \mathcal{L}(V)} T^{\dim \mathcal{L}(V)}$ which is equal to 0.

- Definition 5.24: for V finite-dimensional and $T \in \mathcal{L}(V)$, the generator m_T of the ideal J_T is called **minimal polynomial** of T, i.e., m_T is the unique monic polynomial of smallest degree with $m_T(T) = 0$.
  - By (5.29), it follows that m_T is also "minimal" in the sense that any other polynomial with q with $q(T) = 0$ is a multiple of m_T.
- Proposition 5.31: if V is finite-dimensional $T \in \mathcal{L}(V)$, and $U \subseteq V$ is a T-invariant subspace, then the minimal polynomial of T is a multiple of the minimal polynomial of $T|_U$.
- Proposition 5.22: if V is finite-dimensional and $T \in \mathcal{L}(V)$, then $\deg m_T \leq \dim V$.
- Proposition 5.27: if $\dim V < \infty$ and $T \in \mathcal{L}(V)$, then the zeros of the minimal polynomial of T are the eigenvalues of T.
  - Note that this gives an alternate proof that any operator on V has at most $\dim V$ distinct eigenvalues.
- Proposition 5.32: If $\dim V < \infty$ and $T \in \mathcal{L}(V)$, then T is not invertible iff the constant term of its minimal polynomial is 0.

Eigenvalues on Odd-Dimensional Real Vector Spaces

- Lemma 5.33: if $\dim V < \infty$ and $\mathbf{F} = \mathbf{R}$, and $T \in \mathcal{L}(V)$, then $\dim \operatorname{null}(T^2 + bT + cI)$ is even for any $b, c \in \mathbf{R}$ with $b^2 - 4c < 0$.
- Proposition 5.34: every operator on an odd-dimensional real vector space has an eigenvalue.

5C Upper-Triangular Matrices

- Definition 5.38: a square matrix $A \in \mathbf{F}^{n,n}$ is **upper-triangular** if the entries below the diagonal are zero, i.e., if $A_{i,j} = 0$ when $j < i$; an operator $T \in \mathcal{L}(V)$ is **upper-triangulable** if there is some basis of V with respect to which $\mathcal{M}(T)$ is upper-triangular.

- Proposition 5.39: given $T \in \mathcal{L}(V)$ and a basis $v_1, \dots, v_n$ of V, the following are equivalent:
  (a) $\mathcal{M}(T)$ is upper triangular
  (b) $\text{span}(v_1, \dots, v_k)$ is T-invariant for each $k = 1, \dots, n$
  (c) $Tv_k \in \text{span}(v_1, \dots, v_k)$ for each $k = 1, \dots, n$

- Proposition 5.40: given $T \in \mathcal{L}(V)$, if there is a basis with respect to which $\mathcal{M}(T)$ is upper-triangular with diagonal entries $\lambda_1, \dots, \lambda_n$, then the polynomial $p(z) = \prod_{i=1}^{n}(z - \lambda_i)$ satisfies $p(T) = 0$.

- Proposition 5.41: given $T \in \mathcal{L}(V)$, if there is a basis with respect to which $\mathcal{M}(T)$ is upper-triangular, then the eigenvalues of T are precisely the diagonal entries of $\mathcal{M}(T)$.

- Warning 5.43: it can happen that a linear operator $T \in \mathcal{L}(\mathbf{R}^n)$ is not triangulable, but that the operator on $\mathcal{L}(\mathbf{C}^n)$ given by the same formula is triangulable.
  - An example is the rotation by 90 degrees operator $T(x, y) = (-y, x)$, which as we saw has two distinct eigenvalues over $\mathbf{C}$ (hence is triangulable), but has no eigenvalues over $\mathbf{R}$ (hence cannot be triangulable).

- Proposition 5.44: if V is finite-dimensional and $T \in \mathcal{L}(V)$, then T has an upper-triangular matrix with respect to some basis iff the minimal polynomial of T splits into linear factors.

- Corollary 5.48: if $F = \mathbf{C}$, then every $T \in \mathcal{L}(V)$ has an upper-triangular matrix with respect to some basis.

5D Diagonalizable Operators

Diagonal matrices

- Definition 5.48: a square matrix $A \in \mathbf{F}^{n,n}$ is **diagonal** if $A_{i,j} = 0$ when $i \neq j$.
- Definition 5.50: an operator $T \in \mathcal{L}(V)$ on a finite-dimensional space V is **diagonalizable** if there is some basis with respect to which $\mathcal{M}(T)$ is diagonal.
- Example: the operator $T \in \mathcal{L}(\mathbf{R}^2)$ given by $T(x, y) = (y, x)$ is diagonalizable.
  - The reason is that its basis with respect to the matrix $(1, 1), (1, -1)$ is diagonal.
  - Note, however, that its matrix with respect to the *standard* basis is $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, which is not diagonal.
- Definition 5.52: the **$\boldsymbol{\lambda}$-eigenspace** of an operator $T \in \mathcal{L}(V)$ is the subspace $E(\lambda, T) = \operatorname{null}(T - \lambda I) \subseteq V$ consisting of all eigenvectors with eigenvalue λ, along with the zero vector.
  - Remark: λ is an eigenvalue of T iff $E(\lambda, T) \neq \{0\}$.
  - Remark: $E(\lambda, T)$ is an invariant subspace, and $T|_{E(\lambda,T)} = \lambda I$.
- Proposition 5.54: if $\lambda_1, \dots, \lambda_m$ are distinct eigenvalues of T, then $E(\lambda_1, T) + \cdots + E(\lambda_m, T)$ is a direct sum.
- Corollary: if $\dim V < \infty$, then $\dim E(\lambda_1, T) + \cdots + \dim E(\lambda_m, T) \leq \dim V$.

Conditions for Diagonalizability

- Proposition 5.55: suppose $\dim V < \infty$ and $T \in \mathcal{L}(V)$, and $\lambda_1, \dots, \lambda_m$ are distinct eigenvalues of T. Then the following are equivalent:
  (a) T is diagonalizable
  (b) V has a basis consisting of eigenvectors of T
  (c) $V = E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_m, T)$
  (d) $\dim V = \dim E(\lambda_1, T) + \cdots + \dim E(\lambda_m, T)$
- Corollary 5.58: if $\dim V = n$ and $T \in \mathcal{L}(V)$ has n distinct eigenvalues, then T is diagonalizable.
- Example: $T(x, y, z) = (2x, 3y + z, 3z)$ is not diagonalizable (whether $\mathbf{F} = \mathbf{R}$ or $\mathbf{F} = \mathbf{C}$).
- Example 5.59: by diagonalizing $T \in \mathcal{L}(\mathbf{F}^3)$ given by $T(x, y, z) = (2x + y, 5y + 3z, 8z)$, we can easily compute T^{100}.

- Proposition 5.62: if $\dim V < \infty$, then $T \in \mathcal{L}(V)$ is diagonalizable iff its minimal polynomial is a product of *distinct* linear factors.
- Corollary 5.65: the restriction of a diagonalizable operator $T \in \mathcal{L}(V)$ to an invariant subspace $U \subseteq V$ is again diagonalizable.

5E Commuting Operators

- Definition 5.71: we say that two operators $S, T \in \mathcal{L}(V)$ **commute** if $ST = TS$; similarly, two matrices $A, B \in \mathcal{F}^{n,n}$ commute if $AB = BA$.
  - Example: We have seen that $p(T)$ and $q(T)$ commute for any $T \in \mathcal{L}(V)$ and any $p, q \in \mathcal{P}(\mathbf{F})$.
- Example 5.7.2: on the vector space $\mathcal{P}(\mathbf{R}^2)$ of polynomials in two variables, the partial differentiation operators $D_x, D_y \in \mathcal{L}(\mathcal{P}(\mathbf{R}^2))$ commute.
- Proposition 5.74: given $S, T \in \mathcal{L}(V)$ and a basis for V, the operators S and T commute if and only if the matrices $\mathcal{M}(S)$ and $\mathcal{M}(T)$ commute.
- Proposition 5.75: if $S, T \in \mathcal{L}(V)$ commute, then $E(\lambda, S)$ is T-invariant for any $\lambda \in \mathcal{F}$.
- Definition: we say two operators $S, T \in \mathcal{L}(V)$ are **simultaneously diagonalizable** (or **simultaneously upper-triangulable**) if there is a basis $\mathcal{B}$ of V for which $\mathcal{M}(S, \mathcal{B})$ and $\mathcal{M}(T, \mathcal{B})$ are both diagonal (or upper-triangular).
- Proposition 5.76: if $S, T \in \mathcal{L}(V)$ are both diagonalizable, then they are simultaneously diagonalizable if and only if they commute.
- Proposition 5.78: if $\mathbf{F} = \mathbf{C}$ and $\dim V < \infty$ and $S, T \in \mathcal{L}(V)$ commute, then there is $v \in V$ which is an eigenvector for both S and T.
- Proposition 5.80: if $F = \mathbf{C}$ and $\dim V < \infty$, then any commuting $S, T \in \mathcal{L}(V)$ are simultaneously triangulable.
- Proposition 5.81: if $F = \mathbf{C}$ and $\dim V < \infty$ and $S, T \in \mathcal{L}(V)$ commute, then each eigenvalue of $S + T$ or ST is of the form $\lambda + \mu$ or $\lambda \cdot \mu$, respectively, where λ is an eigenvalue of S and μ is an eigenvalue of T.

8A Generalized Eigenvectors and Nilpotent Operators

- The following lectures will be building up to the following theorem, which is a strengthening of the theorem that every operator on a complex finite-dimensional vector space is triangulable:
- Theorem 8.46: if $F = \mathbf{C}$ and $\dim V < \infty$, then for any $T \in \mathcal{L}(V)$, there is a basis of V such that $\mathcal{M}(T)$ has the following form:

$$\left(\begin{array}{ccc|c|ccc} \lambda_1 & 1 & & & & & \\ & \ddots & 1 & & & & \\ & & \lambda_1 & & & & \\ \hline & & & \ddots & & & \\ \hline & & & & \lambda_k & 1 & \\ & & & & & \ddots & 1 \\ & & & & & & \lambda_k \end{array}\right)$$

Null Spaces of Powers of an Operator

- Proposition 8.1: for any $T \in \mathcal{L}(V)$ and any $k \geq 0$, we have null $T^k \subseteq$ null T^{k+1}.
- Definition: we define the **stable nullspace** of $T \in \mathcal{L}(V)$, denoted null T^∞, to be the union $\bigcup_{k=0}^\infty$ null T^k.
- Proposition: null T^∞ is a subspace.
  - Indeed, if $v, w \in$ null T^∞, then there are $k, l \geq 0$ with $v \in$ null T^k and $v \in$ null T^l, hence $v, w \in$ null $T^{\max(k,l)}$, hence $av, v + w \in$ null $T^{\max(k,l)} \subseteq$ null T^∞ for any $a \in \mathbf{F}$.
- Proposition 8.2: given $T \in \mathcal{L}(V)$, if null $T^m =$ null T^{m+1} for some $m \geq 0$, then null $T^k =$ null T^{k+1} for all $k \geq m$ (in other words, null $T^k =$ null T^∞ for any $k \geq m$).
- Proposition 8.3: if $\dim V = n$ and $T \in \mathcal{L}(V)$, then null $T^n =$ null T^{n+1} (in other words, null $T^k =$ null T^∞ for any $k \geq n$).
- Proposition 8.4: if $\dim V = n$ and $T \in \mathcal{L}(V)$, then $V =$ null $T^n \oplus \operatorname{range} T^n$.

Generalized Eigenvectors

- Definition 8.8: a **generalized eigenvector** of T with eigenvalue λ is a non-zero vector $v \in V$ such that $(T - \lambda I)^k v = 0$ for some $k > 0$ (i.e., $v \in$ null $(T - \lambda I)^\infty$).

- Remark: λ is then an eigenvalue of T, since taking the minimal k with $(T-\lambda I)^k v = 0$, we have $(T-\lambda I)(T-\lambda I)^{k-1}v = 0$ and $(T-\lambda I)^{k-1}v \neq 0$.
- Remark: if $\dim V = n$, then this condition is equivalent to $(T-\lambda I)^n = 0$ by (8.3).

- Definition 8.19: the **generalized eigenspace** of T with eigenvalue λ, denoted $G(\lambda, T)$, is the subspace $\operatorname{null}(T-\lambda I)^\infty$.
  - Note that each $\operatorname{null}(T-\lambda I)^k$ is T-invariant since $(T-\lambda I)^k$, and it follows that $G(\lambda, T)$ is also T-invariant. (See (8.22)).

- Example 8.21: the generalized eigenspaces of $T \in \mathcal{L}(\mathbf{C}^3)$ defined by $T(x, y, z) = (4x, 0, 5z)$ are $G(0, T) = \operatorname{span}(e_1, e_2)$ and $G(5, T) = \operatorname{span}(e_3)$.

- Proposition 8.9: if $F = \mathbf{C}$ and $\dim V < \infty$, then for any $T \in \mathcal{L}(V)$, there is a basis for V consisting of generalized eigenvectors for V.
  - Equivalently: V is the direct sum of the generalized eigenspaces $G(\lambda, T)$. (See also (8.22).)

- Example 8.10: $\begin{bmatrix} 0 & 4 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 5 \end{bmatrix}$ has generalized eigenvalues e_1, e_2 with eigenvalue 0 and e_3 with eigenvalue 5.

- Lemma 8.11: given $T \in \mathcal{L}(V)$ and any two eigenvalues, λ_1, λ_2 of T, we have $G(\lambda_1, T) \cap G(\lambda_2, T) = \{0\}$.
  - Note that if $F = \mathbf{C}$ and $\dim V < \infty$, this follows from (8.9).

- Proposition 8.12: if $\lambda_1, \ldots, \lambda_k$ are distinct eigenvalues of $T \in \mathcal{L}(V)$, then any list $v_1, \ldots, v_k$ such that v_i is a generalized eigenvector with eigenvalue λ_i is independent.
  - In other words, the sum $G(\lambda_1, T) + \cdots + G(\lambda_k, T)$ is direct.
  - Note that if $F = \mathbf{C}$ and $\dim V < \infty$, this follows from (8.9).

Nilpotent Operators

- Definition 8.14: $T \in \mathcal{L}(V)$ is **nilpotent** if $T^k = 0$ for some $k \geq 0$.
  - If $\dim V = n < \infty$, this is equivalent to $\operatorname{null} T^n = \operatorname{null} T^\infty = G(0, T) = V$.

- Example 8.15: the differentiation operator on $\mathcal{P}_m(\mathbf{R})$ is nilpotent, as is any operator represented by a triangular matrix with zeros on the diagonal.

- Example 8.22: if λ is an eigenvalue of T and $\dim G(\lambda, T) < \infty$, then the restriction $(T - \lambda_k I)|_{G(\lambda, T)}$ is nilpotent.

- Proposition 8.17:

  (a) if T is nilpotent and $V \neq 0$, then 0 is an eigenvalue of T, and T has no other eigenvalues

  (b) if $F = \mathbf{C}$ and $\dim V < \infty$ and 0 is the only eigenvalue of T, then T is nilpotent.

- Proposition 8.18: given $T \in \mathcal{L}(V)$ with $\dim V < \infty$, the following are equivalent:

  (a) T is nilpotent

  (b) $m_T(z) = z^m$ for some m

  (c) there is a basis of V such that $\mathcal{M}(T)$ is triangular with zeros on the diagonal.

8B Generalized Eigenspace Decomposition

Multiplicity of an Eigenvalue

- Definition 8.23: given $T \in \mathcal{L}(V)$, the **multiplicity** of an eigenvalue λ of T is the dimension of $G(\lambda, T)$.
- Proposition 8.25: if $\mathbf{F} = \mathbf{C}$ and $\dim V < \infty$, then the sum of the multiplicities of the eigenvalues of any $T \in \mathcal{L}(V)$ is equal to $\dim V$.
- Definition 8.26: given $T \in \mathcal{L}(V)$, and assuming $\dim V < \infty$ and $\mathbf{F} = \mathbf{C}$, the **characteristic polynomial** of T is the polynomial $p_T(z) = (z - \lambda_1)^{d_1} \cdots (z - \lambda_m)^{d_m}$, where $\lambda_1, \dots, \lambda_m$ are the distinct eigenvalues of T and $d_1, \dots, d_m$ are their multiplicities.
  - This agrees with the usual definition $p_T(z) = \det(zI - T)$ of the characteristic polynomial (which isn't available to us since we haven't introduced determinants); indeed, in (8.31), we will show that $p_T(z) = (z - c_1) \cdots (z - c_n)$, where $c_1, \dots, c_n$ are the entries appearing in any triangulation of T, which agrees with the present definition, since determinant of a triangular matrix is the product of the diagonal entries.
- Proposition 8.28: if $F = \mathbf{C}$ and $\dim V < \infty$, then for any $T \in \mathcal{L}(V)$, we have
  (a) $\deg p_T = \dim V$
  (b) the zeros of p_T are the eigenvalues of T.
- Theorem 8.29: if $F = \mathbf{C}$ and $\dim V < \infty$, then for any $T \in \mathcal{L}(V)$, we have $p_T(T) = 0$.
- Corollary 8.30: p_T is a multiple of m_T.
- Proposition 8.31: if $F = \mathbf{C}$ and $\dim V < \infty$, then for any $T \in \mathcal{L}(V)$, we have $p_T(z) = (z - c_1) \cdots (z - c_n)$, where $c_1, \dots, c_n$ are the diagonal entries appearing in any triangulation of T.
  - In other words, the multiplicity of each eigenvalue λ is the number of occurrences of λ among the diagonal entries $c_1, \dots, c_n$.

Block diagonal matrices

- Definition: a matrix A is **block diagonal** if it consists of a sequence of square matrices $A_1, \dots, A_m$ arranged along the diagonal of A, and has zeros everywhere else.
  - Formally, for $A \in \mathbf{F}^{n,n}$ to be block diagonal means that there is a decomposition of the list $(1, \dots, n)$ into a concatenation of sub-lists $S_1 \dots S_m$ such that for each $1 \leq i, j \leq n$, we have $A_{i,j} = 0$ unless i and j both appear in the same sub-list S_k.
  - If $S_k = (r + 1, \dots, r + l)$, we then define the ***k*-th block** of A to be the matrix $A_k \in F^{l,l}$ defined by $(A_k)_{i,j} = A_{r+i,r+j}$.

- Proposition: given $T \in \mathcal{L}(V)$ and a basis $\mathcal{B}$ for V, the matrix $\mathcal{M}(T, \mathcal{B})$ is block diagonal if and only if the basis $\mathcal{B}$ is a concatenation of sub-lists $\mathcal{B}_1 \cdots \mathcal{B}_m$ such that the subspace $W_k = \text{span}(\mathcal{B}_k) \subseteq V$ is T-invariant for each $1 \leq k \leq m$.
  - Moreover, the k-th block of $\mathcal{M}(T, \mathcal{B})$ is then $\mathcal{M}(T|_{W_k}, \mathcal{B}_k)$.
  - Also, note that in this case, we have a direct sum decomposition $V = W_1 \oplus \cdots \oplus W_k$.
- Proposition 8.37: if $\mathbf{F} = \mathbf{C}$ and $\dim V < \infty$ and $T \in \mathcal{L}(V)$, and if $\lambda_1, \ldots, \lambda_m$ are the distinct eigenvalues of T with multiplicities $d_1, \ldots, d_m$, then there is a basis of V such that $\mathcal{M}(T)$ is block diagonal with blocks $A_1, \ldots, A_m$ such that A_k is an upper-triangular d_k-by-d_k matrix with λ_k along the diagonal.
- Example 8.38 works this out explicitly for the operator $T(x, y, z) = (6x + 3y + 4z, 6y + 2z, 7z)$.

8C Jordan Form

- Definition 8.44: Jordan blocks and Jordan bases
  - Let $N_n \in \mathbf{F}^{n,n}$ (or just N for short) be the matrix with ones just above the diagonal, and zeros everywhere else (so $N_{j,k} = \delta_{j+1,k}$).
  - A **Jordan block** is a square matrix of the form $\lambda I + N$.
  - Given $T \in \mathcal{L}(V)$, a **Jordan block basis** for T is a basis $(v_1, \ldots, v_n)$ such that $\mathcal{M}(T)$ is a Jordan block, meaning that $Tv_1 = \lambda v_1$ and $Tv_i = \lambda v_i + v_{i-1}$ for $i > 1$.
  - A matrix is in **Jordan form** if it is block diagonal where each block is a Jordan block.
  - Given $T \in \mathcal{L}(V)$, a **Jordan basis** for T is a basis $\mathcal{B}$ such that $\mathcal{M}(T)$ is in Jordan form, meaning that there is a T-invariant direct sum decomposition $V = V_1 \oplus \cdots \oplus V_p$ and Jordan block bases $\mathcal{B}_k$ of V_k such that $\mathcal{B} = \mathcal{B}_1 \cdots \mathcal{B}_p$.
- Example 8.42: the matrix of $T(w, x, y, z) = (0, x, y, z)$ with respect to the basis $(T^3 e_1, T^2 e_1, Te_1, e_1)$ is $N = 0I + N$; thus, this is a Jordan block basis for T.
- Example 8.43: the matrix of $T(a, b, c, d, e, f) = (0, a, b, 0, d, 0)$ with respect to the basis $(T^2 e_1, Te_1, e_1, Te_4, e_4, e_6)$ is in Jordan form with blocks N_3, N_2, and N_1; thus, this is a Jordan basis for T.
- Proposition 8.45: if $\dim V < \infty$ and $T \in \mathcal{L}(V)$ is nilpotent, then there exists a Jordan basis for T.
- Theorem 8.46: if $F = \mathbf{C}$, then for any $T \in \mathcal{L}(V)$, there is a Jordan basis for T.

6A Inner products

Inner products

- Definition 6.2: an **inner product** on V is a function $\beta\colon V \times V \to \mathbf{F}$ satisfying the following properties:
  - ("positive definite") $\beta(v,v) \geq 0$ for all non-zero $v \in V$
  - ("linear in the first slot") The function $\beta(-,v)\colon V \to \mathbf{F}$ is linear for each $v \in V$.
  - ("conjugate-symmetric") $\beta(u,v) = \overline{\beta(v,u)}$ for all $u, v \in V$.
- Remarks:
  - Usually, when we are considering a fixed inner product β on V, we write $\langle u, v\rangle$ in place of $\beta(u,v)$.
  - The linearity in the first slot together with conjugate symmetry implies **anti-linearity** in the second slot, meaning that $\langle u, av+w\rangle = \bar{a}\langle u,v\rangle + \langle u,w\rangle$ for $u,v,w \in V$ and $a \in \mathbf{F}$.
  - When $\mathbf{F} = \mathbf{R}$, the conjugate symmetry simply becomes symmetry ($\langle u,v\rangle = \langle v,u\rangle$), and the anti-linearity in the second slot becomes linearity (thus in this case β is simply **bilinear**).
- Examples 6.3:
  - (a) The **Euclidean inner product** on $\mathbf{F}^n$ is defined by $\langle u,v\rangle = \sum_{i=1}^n u_i \bar{v}_i$ for $u, v \in \mathbf{F}^n$.
  - (b) For any $c_1, \ldots, c_n > 0$ we can define an inner product on $\mathbf{F}^n$ by $\langle u,v\rangle = \sum_{i=1}^n c_i u_i \bar{v}_i$.
  - (c) On the vectors space $\mathcal{C}([0,1])$ of continuous real-valued functions $[0,1] \to \mathbf{R}$, there is an inner product given by $\langle f, g\rangle = \int_{-1}^1 fg$.
- Definition 6.4: an **inner product space** is a pair (V, β) with V a vector space over $\mathbf{F}$ and β an inner product on V.

Norms

- For the rest of Chapters 6 and 7, V and W denote inner product spaces over $\mathbf{F}$.
- Definition 6.7: the **norm** of $v \in V$ is defined by $\|v\| = \sqrt{\langle v,v\rangle}$.
- Example 6.8: the norm of $z \in \mathbf{F}^n$ with the standard inner product is $\|v\| = \left(\sum_{i=1}^n |v_i|^2\right)^{1/2}$, and the norm of a continuous function $f\colon [0,1] \to \mathbf{R}$ with the inner product from (6.3) is $\|f\| = \left(\int_{-1}^1 f^2\right)^{1/2}$.
- Proposition 6.9: given $v \in V$,
  (a) $\|v\| = 0$ iff $v = 0$
  (b) $\|\lambda v\| = |\lambda| \|v\|$ for all $\lambda \in \mathbf{F}$.

- Definition 6.10: $u, v \in V$ are **orthogonal**, written $u \perp v$, if $\langle u, v\rangle = 0$.
- Proposition 6.11: (a) $0 \perp v$ for all $v \in V$ and (b) 0 is the only $v \in V$ with $v \perp v$.
- Proposition 6.12 (Pythagorean theorem): given $u, v \in V$, if $u \perp v$, then $\|u+v\|^2 = \|u\|^2 + \|v\|^2$.
- Proposition 6.13: given $u, v \in V$ with $v \neq 0$, if we define $c = \frac{\langle u,v\rangle}{\langle v,v\rangle}$ and $w = u - cv$, then $u = cv + w$ and $w \perp v$.
  - The vector $cv = \frac{\langle u,v\rangle}{\langle v,v\rangle}v$ is called the **projection** of u onto v, and denoted $P_u v$; we will return to it later.
- Theorem 6.14 (Cauchy-Schwarz inequality): for any $u, v \in V$, we have $|\langle u, v\rangle| \leq \|u\|\|v\|$, where equality holds if and only if one of u and v is a scalar multiple of the other.
- Remark: when $\mathbf{F} = \mathbf{R}$ an equivalent statement of the Cauchy-Schwarz inequality is $-1 \leq \frac{\langle u,v\rangle}{\|u\|\|v\|} \leq 1$ for all $u, v \neq 0$. It follows that there is a unique $\theta \in [-\pi/2, \pi/2]$ satisfying $\cos\theta = \frac{\langle u,v\rangle}{\|u\|\|v\|}$. In the case $V = \mathbf{R}^3$, this θ is precisely the angle between the vectors u and v. In a general real inner product space V, we can thus **define** the "angle" between two vectors u and v to be this number $\theta = \arccos\frac{\langle u,v\rangle}{\|u\|\|v\|} \in [-\pi/2, \pi/2]$.
  - Be warned that it is much less obvious how one might define the angle between two vectors in a **complex** inner product space.
- Example 6.16: it follows that $\left(\int_{-1}^{1} fg\right)^2 \leq \left(\int_{-1}^{1} f^2\right)\left(\int_{-1}^{1} g^2\right)$ for continuous function f and g.
- Corollary 6.17 (triangle inequality): given $u, v \in V$, we have $\|u+v\| \leq \|u\| + \|v\|$, where equality holds if and only if one of u and v is a nonnegative real multiple of the other.
- Proposition 6.21 (parallelogram law): given $u, v \in V$, we have $\|u+v\|^2 + \|u-v\|^2 = 2(\|u\|^2 + \|v\|^2)$.

6B Orthonormal Bases

- V and W continue to denote inner product spaces over $\mathbf{F}$.

Orthonormal Lists and the Gram-Schmidt Procedure

- Definition 6.22: a list of vectors $e_1, \ldots, e_m$ is **orthonormal** if $\langle e_i, e_j \rangle = \delta_{ij}$ for all $1 \leq i, j \leq m$.
- Example 6.23: in the space $\mathcal{C}([0,1])$ of continuous functions $[0,1] \to \mathbf{R}$ with the inner product $\langle f, g \rangle = 2\int_0^1 fg$, the functions $\frac{1}{\sqrt{2}}, \cos(2\pi nx), \sin(2\pi nx)$ for $0 \neq n \in \mathbb{Z}$, form an infinite orthonormal list.
  - This orthonormal list serves as the basis of the theory of *Fourier analysis*.
- Proposition 6.24: if $e_1, \ldots, e_m$ is an orthonormal list, then $\|\sum_{i=1}^m a_i e_i\|^2 = \sum_{i=1}^m |a_i|^2$ for any $a_1, \ldots, a_m \in \mathbf{F}$.
- Corollary 6.25: any orthonormal list is linearly independent.
- Proposition 6.26 (Bessel's inequality): if $e_1, \ldots, e_m$ is an orthonormal list, then for any $v \in V$, we have $\sum_{i=1}^m |\langle v, e_i \rangle|^2 \leq \|v^2\|$.
  - What is usually called "Bessel's inequality" is a generalization of this which allows the orthonormal list to be infinite (for example, the list of trigonometric functions listed above, which was the original context in which Bessel proved the inequality).
- Definition 6.27: an **orthonormal basis** of V is an orthonormal list of vectors that is also a basis.
- Proposition 6.28: any orthonormal list of length $\dim V$ is an orthonormal basis.
- Proposition 6.30: if $e_1, \ldots, e_n$ is an orthonormal basis of V and $u, v \in V$, then:
  (a) $v = \sum_{i=1}^n \langle v, e_i \rangle e_i$
  (b) $\|v\|^2 = \sum_{i=1}^n |\langle v, e_i \rangle|^2$
  (c) $\langle u, v \rangle = \sum_{i=1}^n \langle u, e_i \rangle \overline{\langle v, e_i \rangle}$.
- Proposition 6.32 (Gram-Schmidt procedure): if $v_1, \ldots, v_m$ are linearly independent, then defining f_k inductively by $f_1 = v_1$ and $f_k = v_k - \sum_{i=1}^{k-1} P_{f_i} v$ for $k > 1$, and defining $e_k = f_k / \|f_k\|$, the list $e_1, \ldots, e_m$ is orthonormal and satisfies $\operatorname{span}(v_1, \ldots, v_m) = \operatorname{span}(e_1, \ldots, e_m)$.
- Corollary 6.35: any finite-dimensional inner product space has an orthonormal basis.

- A consequence of this is that every finite-dimensional inner-product space V over $\mathbf{F}$ is isomorphic, as an inner-product space, to $\mathbf{F}^n$ with the standard inner product, where $n = \dim V$; that is, there is an isomorphism $T\colon V \to \mathbf{F}^n$ such that $\langle Tu, Tv \rangle = \langle u, v \rangle$ for all $u, v \in V$.

- Corollary 6.36: any orthonormal list of vectors in a finite-dimensional inner product space can be extended to an orthonormal basis.
- Corollary 6.37: if $\dim V < \infty$ and $T \in \mathcal{L}(V)$ is triangulable, then there is an orthonormal basis such that $\mathcal{M}(T)$ is triangular.
- Corollary 6.38: any operator on a finite-dimensional inner product space has an upper-triangular matrix with respect to an orthonormal basis.

Linear Functionals on Inner Product Spaces

- Proposition 6.42 (finite-dimensional Riesz representation theorem): if $\dim V < \infty$, then for any $\varphi \in V'$, there is a unique $v \in V$ with $\varphi(u) = \langle u, v \rangle$ for all $u \in V$.
  - What is usually called the Riesz representation theorem is a generalization of this that holds for certain infinite-dimensional inner-product spaces.
  - Proof (of the finite-dimensional version):
  - In the case $\mathbf{F} = \mathbf{R}$, we define a map $F\colon V \to V'$ taking v to $\langle -, v \rangle$. Then F is a linear map by the bilinearity of the inner product.
    - F is injective by positive-definiteness of the inner product: if $F(v) = 0$, then by definition $\langle u, v \rangle = F(v)(u) = 0$ for all $u \in V$, hence $\langle v, v \rangle = 0$, hence $v = 0$.
    - Since F is injective and V and V' have the same dimension, it follows that F is surjective, which is what we wanted to show.
  - In the case $\mathbf{F} = \mathbf{C}$, define a new vector space $\overline{V}$ which has the same underlying set as V and the same addition operation, but has scalar multiplication $\star\colon \mathbf{C} \times \overline{V} \to \overline{V}$ defined by $a \star v = \bar{a} \cdot v$ for $a \in \mathbf{C}$ and $v \in \overline{V} = V$, where "$\cdot$" denotes the scalar multiplication in V.
    - It is straightforward to verify that $\overline{V}$ still satisfies the axioms of a complex vector space.
    - It is also easy to check that any basis for V is also a basis for $\overline{V}$, hence $\dim \overline{V} = \dim V$.
    - We now again define a map $F\colon \overline{V} \to V'$ taking $v \to \langle -, v \rangle$ which is again linear because of the **anti-linearity** of the inner product in the second slot: $F(a \star v)(u) = \langle u, a \star v \rangle = \langle u, \bar{a} \cdot v \rangle = a \cdot \langle u, v \rangle = (a \cdot F(v))(u)$.
    - Injectivity and surjectivity of F then from the same arguments above.

6C Orthogonal Complements and Minimization Problems

- V and W continue to denote inner product spaces over $\mathbf{F}$.

Orthogonal Complements

- Definition 6.46: The **orthogonal complement** of any subset $U \subseteq V$ is the subset $U^\perp = \{v \in V \mid u \perp v \text{ for all } u \in U\}$.
  - We are mainly interested in the case in which U is a subspace.
- Examples 6.47:
  - The orthogonal complement of a line/plane through the origin in $\mathbf{R}^3$ is a plane/line through the origin.
  - In $\mathbf{F}^n$, the orthogonal complement of $\operatorname{span}(e_1, \ldots, e_k)$ is $\operatorname{span}(e_{k+1}, \ldots, e_n)$.
  - This holds more generally whenever $e_1, \ldots, e_n$ is an orthonormal basis of V.
- Proposition 6.48:
  (a) $U^\perp$ is a subspace
  (b) $\{0\}^\perp = V$
  (c) $V^\perp = \{0\}$
  (d) $U \cap U^\perp \subseteq \{0\}$
  (e) if $U_1 \subseteq U_2$, then $U_2^\perp \subseteq U_1^\perp$.
- Proposition 6.49: if $U \subseteq V$ is a finite-dimensional subspace, then $V = U \oplus U^\perp$.
- Definition 6.55: if $U \subseteq V$ is a finite-dimensional subspace, then the **orthogonal projection** onto U is the operator $P_U \in \mathcal{L}(V)$ such that $P_U(u + w) = u$ for $u \in U$ and $w \in U^\perp$.
  - Explicitly, if $u_1, \ldots, u_k$ is an orthogonal basis of U, then $P_U(w) = \sum_{i=1}^k \frac{\langle w, u_i \rangle}{\langle u_i, u_i \rangle} u_i$.
- Corollary 6.51: if V is finite-dimensional and $U \subseteq V$ is a subspace, then $\dim U^\perp = \dim V - \dim U$.
- Corollary 6.52: if $U \subseteq V$ is a finite-dimensional subspace, then $(U^\perp)^\perp = U$.
- Corollary 6.54: if $U \subseteq V$ is a finite-dimensional subspace, then $U^\perp = \{0\} \iff U = V$.
- Proposition 6.57: if U is a finite-dimensional subspace of V, then (i) P_U is a projection operator onto U (see below), and (ii) $\|P_U v\| \leq \|v\|$ for all $v \in V$.

Aside on projection operators

- This is something that we should have covered when we discussed direct sums; during this aside, let V be an arbitrary vector space (rather than an inner product space).

- Definition: given a subspace $U \subseteq V$ of a vector space V, a **projection operator** onto U is an operator $P \in \mathcal{L}(V)$ with $\mathrm{range}(P) = U$ and $P|_U = I$.

- Proposition: $P \in \mathcal{L}(V)$ is a projection operator if and only if $P^2 = P$ (an operator with this property is also called **idempotent**), in which case it is a projection operator onto $U = \mathrm{range}(P)$.
  - Proof: if P is a projection onto $U \subseteq V$, then $P(v) \in U$ for any $v \in V$, hence $P(P(v)) = P(v)$.
  - Conversely, if $P^2 = P$, we may set $U = \mathrm{range}\, P$; then any $u \in U$ is of the form $u = Pv$, and hence $Pu = PPv = Pv = u$.

- Definition: given a subspace $U \subseteq V$, a **complement** of U is a subspace $W \subseteq V$ with $U \oplus W = V$.

- Theorem: let $U \subseteq V$ be a subspace. Then
  (a) if P is a projection onto U, then null P is a complement of U;
  (b) conversely, if W is any complement of U, there is a unique projection P onto U with null $P = W$.
  - Proof of (a): if $u \in U \cap \text{null } P$, then $u = Pu = 0$, where the first equality follows from $P|_U = I$; and given any $v \in V$, we have $P(v - Pv) = Pv - PPv = Pv - Pv = 0$, hence $v - Pv \in \text{null } P$, and so $v = Pv + (v - Pv) \in U + \text{null } P$.
  - Proof of (b): we define a projection onto U as follows: given $v \in V$, we can write v uniquely as $v = u + w$ with $u \in U$ and $w \in W$, and we set $Pv = u$; it's easy to check that this is a projection on to U with nullspace W. To prove uniqueness, given any projection P onto U with nullspace W, if $v = u + w$ as above, then we must have $Pv = Pu + Pw = u + 0 = u$.

- Proposition: if P is a projection onto U, then $I - P$ is a projection onto the complement null P of U.
  - Proof: we have $(I - P)^2 = I^2 - 2P + P^2 = I - 2P + P = I - P$, hence $I - P$ is a projection operator; it remains to see that $\mathrm{range}(I - P) = \text{null } P$.
  - For any $v \in V$, we have $P((I - P)v) = P(v - Pv) = Pv - PPv = 0$, so $\mathrm{range}(I - P) \subseteq \text{null } P$.
  - And for any $v \in \text{null } P$, we have $v = v - Pv = (I - P)v \in \mathrm{range}(I - P)$, so null $P \subseteq \mathrm{range}(I - P)$.

Orthogonal complements and annihilators

- We return to our convention that V and W denote inner product spaces over $\mathbf{F}$.
- Proposition: if V is finite-dimensional, then for any subspace $U \subseteq V$, the isomorphism $\overline{V} \to V'$ defined in the proof of (6.42) satisfies $F(U^{\perp}) = U^0$.
  - This follows directly from the definition: $v \in U^{\perp}$ if and only if $\langle u, v \rangle = 0$ for all $u \in U$ if and only if $F(v)|_U = 0$ if and only if $F(v) \in U^0$.

Minimization Problems

- Proposition 6.61: if U is a finite-dimensional subspace of V, then for any $v \in V$, $P_U v$ is the closest point to v in U, in the sense that $\|v - P_U v\| < \|v - u\|$ for any other $u \in U$.
- Example 6.63: we can use this to approximate the function $\sin\colon [-\pi, \pi]$ by a polynomial by letting $V = \mathcal{C}^0([-\pi, \pi])$ with inner product $\langle f, g \rangle = \int_{-\pi}^{\pi} fg$, and letting $U = \operatorname{span}(1, x, \ldots, x^n)$.

7A Self-Adjoint and Normal Operators

- V and W continue to denote inner product spaces over $\mathbf{F}$.
- We will now introduce self-adjoint operators—which are an analogue for operators of symmetric matrices—and normal operators, which are a generalization of self-adjoint operators.
- The most important fact about these, which we will prove in the next section, is the (finite-dimensional) **spectral theorem** (Theorems 7.29 and 7.31), which says that

  (a) every self-adjoint operator T on a finite-dimensional vector space V is diagonalizable with respect to an orthonormal basis,

  (b) if $\mathbf{F} = \mathbf{C}$, then more generally, every **normal** operator T is diagonalizable with respect to an orthonormal basis,

  (c) conversely, if an operator is diagonalizable with respect to an orthonormal basis then, if $\mathbf{F} = \mathbf{R}$, it is self-adjoint, and if $\mathbf{F} = \mathbf{C}$, it is normal.

Adjoints

- Proposition/Definition 7.1 and 7.4: for any $T \in \mathcal{L}(V, W)$, there exists at most one linear map $T^*\colon W \to V$ satisfying $\langle Tv, w\rangle = \langle v, T^*w\rangle$ for all $v \in V$ and $w \in W$; if it exists, T^* is called the **adjoint** of T. Moreover, if $\dim V, \dim W < \infty$, then the adjoint T^* always exists.
- Remark: supposing $\dim V, \dim W < \infty$, let $\varphi_V\colon \overline{V} \xrightarrow{\sim} V'$ and $\varphi_W\colon \overline{W} \xrightarrow{\sim} W'$ be the isomorphisms given by $\varphi_V(v) = \langle -, v\rangle$ and $\varphi_W(w) = \langle -, w\rangle$. Then for any $T\colon \mathcal{L}(V, W)$ and $w \in \overline{W} = W$, we have $\varphi_V(T^*w) = T'(\varphi_W w)$.

$$\begin{array}{ccc} \overline{W} & \xrightarrow{T^*} & \overline{V} \\ {\scriptstyle \varphi_W}\big\downarrow{\scriptstyle \wr} & & {\scriptstyle \wr}\big\downarrow{\scriptstyle \varphi_V} \\ W' & \xrightarrow{T'} & V' \end{array}$$

  - Thus, roughly speaking, the adjoint of a linear map between inner product spaces is "isomorphic" to the dual map; in particular, all the properties of the dual map carry over to give properties of the adjoint map.
- Example: for a linear map $T\colon \mathbf{R}^{m,1} \to \mathbf{R}^{n,1}$ (with the Euclidean inner product) given by multiplying by a matrix A, the adjoint $T^*\colon \mathbf{R}^{n,1} \to \mathbf{R}^{m,1}$ is given by multiplying by A^{t}.
  - Indeed, one way of writing the Euclidean inner product on $\mathbf{R}^{n,1}$ is as the matrix product $\langle u, v\rangle = u^{\mathrm{t}} \cdot v \in \mathbf{R}^{1,1}$.

- Hence $\langle A \cdot u, v\rangle = (A \cdot u)^{t} \cdot v = u^{t} \cdot A^{t} \cdot v = u^{T} \cdot (A^{t} \cdot v) = \langle u, A^{t} \cdot v\rangle$.

- Proposition 7.5: Suppose V and W are finite-dimensional. Then
  (a,b) the map $T \mapsto T^*$ is an anti-linear map $\mathcal{L}(V, W) \to \mathcal{L}(W, V)$
  (c) $(T^*)^* = T$ for any $T \in \mathcal{L}(V, W)$
  (d) $(ST)^* = T^*S^*$ for $T \in \mathcal{L}(V, W)$ and $S \in \mathcal{L}(W, U)$ (where U is a finite-dimensional inner-product space over $\mathbf{F}$)
  (e) $I_V^* = I_V$
  (f) if $T \in \mathcal{L}(V, W)$ is invertible, then T^* is invertible and $(T^*)^{-1} = (T^{-1})^*$.

- Proposition 7.6: if V and W are finite-dimensional, then for any $T \in \mathcal{L}(V, W)$:
  (a) null $T^* = (\operatorname{range} T)^{\perp}$
  (b) $\operatorname{range} T^* = (\text{null } T)^{\perp}$
  (c) null $T = (\text{null } T^*)^{\perp}$
  (d) $\operatorname{range} T = (\text{null } T^*)^{\perp}$

- Definition 7.7: the **conjugate transpose** of a matrix $A \in \mathbf{F}^{m,n}$ is the matrix A^* defined by $(A^*)_{j,k} = \overline{A_{k,j}}$.

- Proposition 7.9: given $T \in \mathcal{L}(V, W)$ and orthonormal bases for V and W, we have $\mathcal{M}(T^*) = \overline{\mathcal{M}(T)}$.

Self-Adjoint Operators

- Definition 7.10: $T \in \mathcal{L}(V)$ is **self-adjoint** if $T = T^*$, i.e., if $\langle Tu, v\rangle = \langle u, Tv\rangle$ for all $u, v \in V$.
  - Remark: given an orthonormal basis V, T is self-adjoint if and only if $\mathcal{M}(T)$ is **conjugate-symmetric** (also known as being **Hermitian**), meaning that $\mathcal{M}(T)^* = \mathcal{M}(T)$.

- Proposition 7.12: Every eigenvalue of a self-adjoint operator is real.
  - Of course, this is only an interesting result when $\mathbf{F} = \mathbf{C}$.

- Proposition 7.13: if $\mathbf{F} = \mathbf{C}$, then for any $T \in \mathcal{L}(V)$, we have $Tv \perp v$ for all $v \in V$ if and only if $T = 0$.
  - This is not true when $\mathbf{F} = \mathbf{R}$, for instance when $T(x, y) = (-y, x)$, a rotation by 90°.

- Proposition 7.14: if $F = \mathbf{C}$, then $T \in \mathcal{L}(V)$ is self-adjoint if and only if $\langle Tv, v\rangle \in \mathbf{R}$ for all $v \in V$.
  - This is obviously false when $\mathbf{F} = \mathbf{R}$, since there are real operators that are not self adjoint.

- Proposition 7.16: if $T \in \mathcal{L}(V)$ is self-adjoint, then $Tv \perp v$ for all $v \in V$ if and only if $T = 0$.
  - When $\mathbf{F} = \mathbf{C}$, this is implied by the stronger result (7.13), which does not have the self-adjoint hypothesis.

Normal operators

- Definition 7.18: $T \in \mathcal{L}(V)$ is **normal** if T and T^* commute.
  - Note that every self-adjoint operator is normal.
  - The main significance of normal operators is the **spectral theorem** in section 7B: if $F = \mathbf{C}$ and $\dim V < \infty$, then T is normal if and only if there is an orthonormal basis $\mathcal{B}$ such that $\mathcal{M}(T, \mathcal{B})$ is diagonal.
- Example 7.19: the operator on $\mathbf{F}^2$ with matrix $\begin{bmatrix} 2 & -3 \\ 3 & 2 \end{bmatrix}$ is normal but not self-adjoint.
- Proposition 7.20: $T \in \mathcal{L}(V)$ is normal iff $\|Tv\| = \|T^*v\|$ for all $v \in V$.
- Proposition 7.21: if $T \in \mathcal{L}(V)$ is normal, then
  (a) $\text{null } T = \text{null } T^*$
  (b) $\operatorname{range} T = \operatorname{range} T^*$ assuming $\dim V < \infty$
  (c) $V = \text{null } T \oplus \operatorname{range} T$ assuming $\dim V < \infty$
  (d) $T - \lambda I$ is normal for all $\lambda \in \mathbf{F}$
  (e) $Tv = \lambda v$ iff $T^*v = \bar{\lambda} v$ for $v \in V$ and $\lambda \in F$.
- Proposition 7.22: if $T \in \mathcal{L}(V)$ is normal, then eigenvectors of T with distinct eigenvalues are orthogonal.
  - This has a simpler proof in the special case that T is self-adjoint.
- Proposition 7.23: when $F = \mathbf{C}$, then $T \in \mathcal{L}(V)$ is normal if and only if there are self-adjoint commuting $A, B \in \mathcal{L}(V)$ with $T = A + iB$.
  - A version of the statement that makes sense over $\mathbf{R}$ is: T is normal if and only if there are commuting $A, B \in \mathcal{L}(V)$ with $T = A + B$ and $A = A^*$ and $B = -B^*$.

7B Spectral Theorem

- V and W continue to denote inner product spaces over $\mathbf{F}$.

Real Spectral Theorem

- Lemma 7.26: if $\dim V < \infty$ and $T \in \mathcal{L}(V)$ is self-adjoint and $b, c \in \mathbf{R}$ satisfy $b^2 - 4c < 0$, then $T^2 + bT + cI \in \mathcal{L}(V)$ is invertible.
- Lemma 7.27: if $\dim V < \infty$ and $T \in \mathcal{L}(V)$ is self-adjoint, then the minimal polynomial m_T is a product of real linear factors.
- Theorem 7.29 (real finite-dimensional spectral theorem): if $\mathbf{F} = \mathbf{R}$ and $\dim V < \infty$, then any $T \in \mathcal{L}(V)$ is self-adjoint if and only if it is diagonalizable with respect to some orthonormal basis.
  - Remark: an equivalent condition to being diagonalizable with respect to an orthonormal basis is that the eigenspaces of T are mutually orthogonal and their sum is V.

Complex Spectral Theorem

- Theorem 7.31 (complex finite-dimensional spectral theorem): if $\mathbf{F} = \mathbf{C}$ and $\dim V < \infty$, then any $T \in \mathcal{L}(V)$ is normal if and only if it is diagonalizable with respect to some orthonormal basis.