MAT-395: Linear Algebra Notebook

Logan Goldberg

January 30, 2017

Contents

Chapter 0

Preliminaries

0.1 Preliminaries

Definition 0.1.1. *Let S be a nonempty set. A* ***multiset*** *$M \subseteq S \times \mathbb{Z}^+$ with* ***underlying set*** *S is a set of ordered pairs*

$$M = \{(s_i, n_i) \mid s_i \in S,\ n_i \in \mathbb{Z}^+,\ s_i \neq s_j \text{ whenever } i \neq j\}.$$

The number n_i is referred to as the ***multiplicity*** *of the elements s_i in M. If the underlying set of a multiset is finite, then we say that the multiset is* ***finite****. The* ***size*** *of the finite multiset is the sum of the multiplicies of all the elements. In other words, $size\colon M \to \mathbb{Z}^+$ is a function*

$$size(M) = \begin{cases} n_1 + size(M \backslash \{(s_1, n_1)\}) & \text{if } M \text{ is nonempty} \\ 0 & \text{otherwise} \end{cases}$$

which gives the size of a multiset M.

The symbol $\mathscr{M}_{m\times n}(F)$ is the set of all rectangular matrices of dimensions m by n with entries in the field F. $\mathscr{M}_{m\times n}$ is used when the field entries need not be mentioned. $\mathscr{M}_{n\times n}(F)$, $\mathscr{M}_n(F)$, or $\mathscr{M}_n$ is used to denote square matrices of dimension n with scalars in F. The identity matrix of size n is denoted by I_n and $A_{i,j}$ is used to denote the (i, j)th entry of matrix A.

Definition 0.1.2. *The* ***main diagonal*** *of an $m \times n$ matrix A is the sequence of entries*

$$A_{1,1}, A_{1,2}, \ldots, A_{k,k}$$

for $k = \min\{m, n\}$.

Definition 0.1.3. *The* ***trace*** *of a $m \times n$ matrix A is the sum over the main diagonal of A.*

Definition 0.1.4. *The* ***transpose*** *of $A \in \mathscr{M}_{m\times n}(F)$ is a matrix denoted by A^t and defined by*

$$(A^t)_{i,j} = A_{j,i}.$$

A matrix A is ***symmetric*** *if $A^t = A$ and* ***skew-symmetric*** *if $A^t = -A$.*

Theorem 0.1.5. *Let $A, B \in \mathscr{M}_{m,n}$. Then*

1) $(A^t)^t = A$

2) $(A+B)^t = A^t + B^t$

3) $(rA)^t = rA^t$ *for all* $r \in F$

4) $(AB)^t = B^t A^t$ *provided that the product* AB *is defined.*

5) $\det(A^t) = \det(A)$.

□

Definition 0.1.6. *Let* $M \in \mathcal{M}_{m\times n}$. *If* $B \subseteq \{1,2,\ldots,m\}$ *and* $C \subseteq \{1,2,\ldots,n\}$, *then the* ***submatrix*** $M[B,C]$ *is the matrix obtained by keeping only the rows with index in* B *and the columns with index in* C.

Proposition 0.1.7. *If* $M \in \mathcal{M}_{m\times n}$, $B \subseteq \{1,2,\ldots,m\}$, *and* $C \subseteq \{1,2,\ldots,n\}$, *then the submatrix* $M[B,C]$ *has size* $|B| \times |C|$.

Definition 0.1.8. *Let* A *be a set. A* ***partition*** *of a nonempty set* A *is a collection of subsets* $\mathcal{P}$ *of* A, *each called* ***blocks***, *such that*

1) $\bigcup P = A$.

2) *For any two distinct blocks* $a, b \in P$, $a \cap b = \emptyset$.

Theorem 0.1.9. *Let* $M \in \mathcal{M}_{m\times n}$ *and* $N \in \mathcal{M}_{n\times k}$. *Let*

1) $\mathcal{P} = \{B_1, B_2, \ldots, B_p\}$ *be a partition of* $\{1,2,\ldots,m\}$.

2) $\mathcal{Q} = \{C_1, C_2, \ldots, C_q\}$ *be a partition of* $\{1,2,\ldots,n\}$.

3) $\mathcal{R} = \{D_1, D_2, \ldots, B_r\}$ *be a partition of* $\{1,2,\ldots,k\}$.

Matrix multiplication at the block level is a well-defined operation, that is

$$[MN][B_i, D_j] = \sum_{C_h \in \mathcal{Q}} M[B_i, C_h] N[C_h, D_j].$$

Definition 0.1.10. *If* $B_{i,j}$ *are matrices of appropriate sizes, then by the* ***block matrix***

$$M = \begin{bmatrix} B_{1,1} & B_{1,2} & \cdots B_{1,n} & \\ \vdots & \vdots & \ddots & \cdots \\ B_{m,1} & B_{m,2} & \cdots & B_{m,n} \end{bmatrix}_{block}$$

we mean the matrix whose upper left submatrix is $B_{1,1}$, *and so on.*

Definition 0.1.11. *An* ***elementary matrix*** *of type* k *is an elementary row or column operation applied to* I_n.

1) *If* $k = 1$, *then a row or column is scaled.*

2) If $k = 2$, then a row or column is interchanged.

3) If $k = 3$, then a scalar multipled of a row or column is added to another.

An elementary matrix E applied from the left onto A, e.g EA, is a row operation on A An elementary matrix E applied from the right onto A, e.g AE, is a column operation on A

Definition 0.1.12. *A matrix R is said to be in **row reduced echelon form** if*

1. *A rows consisting of 0's appear at the bottom of the matrix.*
2. *In any nonzero row, the first nonzero entry is a 1. This entry is called a **leading entry**.*
3. *For any two consecutive rows, the leading entry of the lower row is to the right of the leading entry of the upper row.*
4. *Any column that contains a leading entry has 0's in all other positions.*

Theorem 0.1.13. *Matrices $A, B \in \mathcal{M}_{m\times n}$ are **row equivalent**, denoted by $A \sim B$ if either one can be obtained from the other by a series of elementary row operations.*

1. *Row equivalence is an equivalence relation.*
2. *A marix A is row equivalent to a unique matrix R that is in row reduced echelon form. R is called the **row reduced echelon form** of A.*
3. *A is invertible if and only if its row reduced echelon form is the identity matrix. Hence a matrix is invertible if and only if it is the product of elementary matrices.*

Definition 0.1.14. *A square matrix is **upper triangular** if all of its entries below the main diagonal are 0. Similarly, a square matrix is lower triangular if all its entries above the main diagonal are 0. A square matrix is diagonal if it is both upper triangular and lower triangular.*

Theorem 0.1.15. *Let $F[x]$ be a polynomial ring over the field F.*

1. *A nonconstant polynomial $f(x) \in F[x]$ is irreducible if and only if it has the property that whenever $f(x) \mid p(x)q(x)$, then $f(x) \mid p(x)$ or $f(x) \mid q(x)$. In other words, $f(x)$ is prime.*
2. *If F is a field, $F[x]$ is a Euclidean domain, a principal ideal domain, and a unique factorization domain.*
3. *For any two polynomials $f(x), g(x) \in F[x]$, the **greatest common divisor** of $f(x)$ and $g(x)$ is the unique monic polynomial $p(x)$ over F up to associates for which*
   (a) $p(x) \mid f(x)$ and $p(x) \mid g(x)$
   (b) if $r(x) \mid f(x)$ and $r(x) \mid g(x)$, then $r(x) \mid p(x)$.
4. *For any two polynomials $f(x)$ and $g(x)$, there exist polynomials $a(x)$ and $b(x)$ in $F[x]$ for which*
$$\gcd(f(x), g(x)) = a(x)f(x) + b(x)g(x).$$

Definition 0.1.16. *Let $f : S \to T$. Assuming $0 \in T$, the **support** of f is*

$$supp(f) = \{s \in S \mid f(s) \neq 0\}.$$

Definition 0.1.17. *Let $\sim$ be an equivalence relation on S. For $a \in S$,*

$$[a] = \{b \in S \mid b \sim a\}$$

*is called the **equivalence class** of a.*

Theorem 0.1.18. *The set of all equivalence relations on a set S is in bijective correspondence to the set of all partitions of S. Furthermore,*

1. *The set of distinct equivalence classes corresponds to a unique partition.*

2. *If $\mathscr{P}$ is a partition of S, then the binary relation $\sim$ defined by, $a \sim b$ if and only if a and b belong to the same block in $\mathscr{P}$, is an equivalence relation on S whose equivalence classes are the blocks of $\mathscr{P}$.*

Definition 0.1.19. *Let $\sim$ be an equivalence relation on S. A function $f : S \to T$, where T is any set, is called an **invariant** of $\sim$ if*

$$a \sim b \Rightarrow f(a) = f(b).$$

*and is a **complete invariant** if*

$$a \sim b \Leftrightarrow f(a) = f(b).$$

*A **complete system of invariants** is a finite set of invariants of $\sim$ on S.*

Definition 0.1.20. *Let $\sim$ be an equivalence relation on S. A subset $C \subseteq S$ is said to be a set of **canonical forms** for $\sim$ if for every $s \in S$, there is exactly one $c \in C$ such that $c \sim s$. Put another way, each equivalence class intersects C at exactly one element.*

Simply put, what subset of S provides a unique representation for each equivalence class of $\sim$?

Definition 0.1.21.

Definition 0.1.22.

Definition 0.1.23.

Definition 0.1.24.

Definition 0.1.25.

0.2 Algebraic Structures

Chapter 1

Vector Spaces

1.1 Vector Spaces

1.2 Subspaces

Below is a proof of Theorem 1.1.

Theorem 1.2.1. *A nonempty subset S of a vector space V is a subspace of V if and only if S is closed under addition and scalar multiplication or, equivalently, for all $a, b \in F$ and $u, v \in S$, $au + bv \in S$.*

Proof. Suppose S is a subspace of V. Then by definition S is a vector space over F with vector addition $S \times S \to V$ and with scalar multiplication $F \times S \to V$. Let $u, v \in S$ and $a, b \in F$. By definition of scalar multiplication, $au \in S$ and $bv \in S$. Also from the defintion of vector addition on S, $au + bv \in S$. Hence S is closed under linear combination.

On the other hand, let $S \subseteq V$ be an arbitrary nonempty subset that is closed under linear combination. Since V is a vector space, vector addition associates and commutes on the elements of a nonempty subset S. S is nonempty, so fix $v \in S$. Because S is closed under linear combination and $0u = 0$ for any $u \in V$ by Exercise 1 from this chapter, we conclude that $0v = 0 \in S$. Now since V is a vector space, fix $w \in V$ for which $v + w = 0$. Since $0 \in S$,

$$v + w = 0 \Leftrightarrow w = (-1)v \in S.$$

Again since V is a vector space, scalar multiplication associates and distributes on the elements of a nonempty subset S. Since F is a field and V is a vector space, $1 \neq 0$ and $1 \in F$ with the property that $1v = v$ for all $v \in V$. Since $S \subseteq V$, $1v = v$ for all $v \in S$. Hence S is a subspace of V. □

The proof of Theorem 1.2 can use a little more elaboration. Following the conclusion of the proof, it should state the following.

Proof. Since F is an infinite field, the infinite set A cannot be the union of finitely many proper subspaces since each subspaces of V only contains at most one element of A. Hence $V \neq S_1 \cup S_2 \cup \cdots \cup S_n$. □

1.3 Direct Sums

Theorem 1.3.1. *The set $S(V)$ of all subspaces of a vector space V is a complete lattice under set inclusion, with samllest element $\{0\}$, and largest element V, meet*

$$\sup\{S_i \mid i \in K\} = \bigcap_{i\in K} S_i$$

and join

$$\operatorname{lub}\{S_i \mid i \in K\} = \sum_{i\in K} S_i = \{s_1 + \ldots + s_n \mid s_j \in \bigcup_{i\in K} S_i\}$$

Proof. Let V be a vector space over a field F. We begin by showing that $S(V)$ is a lattice under set inclusion. We will first establish that every collection of sets has a greatest lower bound and a least upper bound under set inclusion.

We claim that the greatest lower bound of a collection of subspaces is a subspace formed by the intersection over those subspaces. Let $P = \{S_i \mid i \in K\}$, where K is an indexing set, be a family of subspaces of V. Since each S_i if a subspace, each contains 0, so $\bigcap P$ is nonempty. Therefore, fix $a, b \in F$ and $u, v \in \bigcap P$. From set theory, $u, v \in S_i$ for each $i \in K$. Since each S_i is a subspace, $au + bv \in S_i$ for each $i \in K$ by Theorem 1.2.1. Thus, $au + bv \in \bigcap P$ and so $\bigcap P$ is a subspace by Theorem 1.2.1.

Now that $\bigcap P \in S(V)$, we can show that it is the greatest lower bound of P. Clearly, $\bigcap P \subseteq S_i$ for each $i \in K$ by definition of intersection. Let T be an arbitrary lower bound of P. Since T is a lower bound of P, T is contained in every $S \in P$. As such if $v \in T$, then $v \in S_i$ for each i. Therefore $v \in \bigcap P$ and so $T \subseteq \bigcap P$.

Hence $\operatorname{glb}\{S_i \mid i \in K\} = \bigcap_{i\in K} S_i$ is the greatest lower bound of any collection of subspaces $\{S_i \mid i \in K\}$. Now we will show that every collection of subspaces has a least upper bound under set inclusion. Let $P = \{S_i \mid i \in K\}$. We claim that the least upper bound over a collection of sets is the set

$$\operatorname{lub} P = \sum P$$

for which $\sum P$ is the set of finite sums of elements in $\bigcup P$.

We begin by showing that $\sum P$ is a subspace of V. Notice first that by definition of $\sum P$ that $\bigcup P \subseteq \sum P$. Therefore if $\sum P$ is a subspace of V, $\sum P$ is an upper bound of P. Recall that each $S \in P$ is a subspace of V, so $0 \in S$. Thus $\sum P$ is nonempty. Therefore fix $a, b \in F$ and $u, v \in \sum P$. Because $\sum P$ contains the union of each $S \in P$, we encounter two different cases.

Case 1: Suppose there is a $j \in K$ with $u, v \in S_j$. Since S_j is a subspace, $au + bv \in S_j$ and so $au + bv \in \sum P$.

Case 2: Suppose u and v hail from distinct subspaces, that is $u \in S_i \setminus S_j$ and $v \in S_j \setminus S_i$ for some $i, j \in K$. Since S_i and S_j are closed under scalar multiplication, $au \in S_i \setminus S_j$ and $bv \in S_j \setminus S_i$. Now because $\sum P$ is closed under finite sums of vectors, and $au, bv \in \bigcup P$, $au + bv \in \sum P$. Hence $\sum P$ is a subspace of V by Theorem 1.2.1.

Last we show that $\sum P$ is the smallest upper bound. Let $T \in S(V)$ be an upper bound of P. Then $S \subseteq T$ for each $S \in P$. Because $S \subseteq T$ for all $S \in P$, if $\{s_j\}_{j=1}^n \subseteq \bigcup P$ is a finite set of elements, then $\sum_{j=1}^n \{s_j\} \in T$ since T is closed under linear combination. Given that $\{s_j\}_{j=1}^n$ was an arbitrary finite subset of elements of $\bigcup P$, we conclude that all finite sums of elements in $\bigcup P$ are in T. Hence $\sum P \subseteq T$. Therefore $\sum P$ is the least upper bound of P and thus every collection of subspaces of V has a least upper bound under set inclusion.

Therefore $S(V)$ is a lattice under set inclusion. Finally we need to show that $S(V)$ is a complete lattice. To do this, we need to show that $\{0\}$ is the smallest element of $S(V)$ and V is the largest element of $S(V)$.

Notice that $\{0\} \subseteq S$ for all $S \in S(V)$ by definition of subspace. Therefore if $T \in S(V)$ is smaller than $\{0\}$, T cannot contain 0 and is thus not in $S(V)$. Therefore $\{0\}$ is the smallest element of $S(V)$.

Because every $S \in S(V)$ is a subspace of V, every S is a subset of V. Now if $T \in S(V)$ is such that $V \subseteq T$, then $V = T$ since $T \in S(V)$. Hence V is the largest element of $S(V)$.

Thus $S(V)$ is a complete lattice as desired. □

Below, I've rewritten the proof of Theorem 1.5 for clarity.

Theorem 1.3.2. *Let $\mathscr{F} = \{S_i \mid i \in I\}$ be a family of distinct subspaces of V. The following are equivalent:*

1) ***(Independence of the family)*** *For each $i \in I$,*

$$S_i \cap \left(\sum_{j \neq i} S_j \right) = \{0\}$$

2) ***(Uniqueness of expression for 0)*** *The zero vector $\vec{0}$ cannot be written as a sum of nonzero vectors from distinct subspaces of $\mathscr{F}$.*

3) ***(Uniqueness of expression)*** *Every nonzero vector $v \in V$ has a unique expression as a sum*

$$v = s_{\sigma(1)} + \cdots + s_{\sigma(n)}$$

of nonzero vectors from distinct subspaces of $\mathscr{F}$ for any permutation $\sigma : \{1,2,\ldots,n\} \to \{1,2,\ldots,n\}$.

Hence, a sum

$$V = \sum_{i \in I} S_i$$

is direct if and only if any one of 1)-3) holds.

Proof. Suppose that 2) fails, that is if $\{j_i\}_{i=1}^n \subseteq I$ is a finite indexing set, then

$$0 = s_{j_1} + \cdots + s_{j_n}$$

where the nonzero vectors s_{j_i} are from distinct subspaces S_{j_i}. Then $n > 1$ and so

$$-s_{j_1} = s_{j_2} + \cdots s_{j_n}$$

which violates 1) because s_{j_1} is now the combination of vectors from distinct subspaces. In other words, $s_{j_1} \in S_i \cap \left(\sum_{i \in I} S_i\right)$. Hence 1) implies 2). Suppose 2) holds and

$$v = s_1 + \cdots + s_n \text{ and } v = t_1 + \cdots t_m$$

where the terms are nonzero and the s_i's belong to distinct subspaces in $\mathscr{F}$ and similarly for the t_i's, then

$$0 = s_1 + \cdots + s_n + t_1 + \cdots + t_m.$$

By collecting terms from the same subspaces, we may write

$$0 = (s_{i_1} - t_{i_1}) + \cdots (s_{i_k} - t_{i_k}) + s_{i_{k+1}} + \cdots + s_{i_n} + t_{i_{k+1}} + \cdots + t_{i_m}$$

with index set $\{i_k \mid 1 \leq k \leq \max\{n, m\}\}$. 2) implies that $n = m = k$ and $s_{i_u} = t_{i_u}$ for all $u \in \{1, \ldots, k\}$. Hence 2) implies 3).
Finally, suppose that there is a nonzero $v \in V$ with

$$v \in S_i \cap \left(\sum_{j \neq i} S_j\right)$$

Then $v = s_i$ for some $s_i \in S_i$ and

$$s_i = s_{j_1} + \cdots s_{j_n}$$

where $s_{j_k} \in S_{j_k}$ are nonzero. But this violates 3) because $s_i \in S$ now has two representations as sums of vectors s_{j_i} from distinct subspaces.

□

1.4 Spanning Sets

Theorem 1.4.1. *A finite set $S = \{v_1, v_2, \ldots, v_n\}$ of vectors in V is a basis of V if and only if*

$$V = \langle v_1 \rangle \oplus \cdots \oplus \langle v_n \rangle.$$

Assume $V = \langle v_1 \rangle \oplus \cdots \oplus \langle v_n \rangle$. Let $u \in V$ be arbitrary. Then v is a linear combination of vectors from the spans of each v_i with nonzero coefficients

$$u = r_1 v_1 + r_2 v_2 + \cdots + r_n v_n$$

Because the family of spans is a direct sum, u has a unique expression of vectors from the family v_i by Theorem 1.5. Therefore v is an essentially unique combination of vectors from S and hence is a basis.
If $V \neq \langle v_1 \rangle \oplus \cdots \oplus \langle v_n \rangle$, then S ceases to be a basis. One of two conditions fails:

Case 1: If $\sum_{i=1}^{n} \langle v_i \rangle \subset V$, then S does not span V.

Case 2: Alternatively, $\langle v_i \rangle \cap \sum_{j \neq i} \langle v_j \rangle \neq \{0\}$. In this case, an arbitrary nonzero $v \in \langle v_i \rangle \cap \sum_{j \neq i} \langle v_j \rangle$ does not have a unique expression as the sum of vectors from each span $\langle v_i \rangle$. Thus v is not an essentially unique combination of vectors in S, so S is not linearly independent.

1.5 The Dimension of a Vector Space

1.6 Ordered Bases and Coordinate Matrices

1.7 The Row and Column Spaces of a Matrix

1.8 The Complexification of a Real Vector Space

Chapter 2

Linear Transformations

2.1 Linear Transformations

2.2 The Kernel and Image of a Linear Transformation

2.3 Isomorphisms

2.4 The Rank Plus Nullity Theorem

2.5 Linear Transformations from F^n to F^m

2.6 Change of Basis Matrix

2.7 The Matrix of a Linear Transformation

2.8 Change of Bases for Linear Transformations

2.9 Equivalence of Matrices

2.10 Similarity of Matrices

2.11 Similarity of Operators

2.12 Invariant Subspaces and Reducing Pairs

2.13 Projection Operators

2.14 Topological Vector Spaces

2.15 Linear Operators on $V^{\mathbb{C}}$

Chapter 3

The Isomorphism Theorems

3.1 Quotient Spaces

3.2 The Universal Property of Quotients and the First Isomorphism Theorem

3.3 Quotient Spaces, Complements, and Codimension

3.4 Additional Isomorphism Theorems

3.5 Linear Functionals

3.6 Dual Basis

3.7 Reflexivity

3.8 Annihilators

3.9 Operator Adjoints

Chapter 4

Modules I: Basic Properties

4.1 Modules

4.2 Submodules

4.3 Spanning Sets

4.4 Linear Independence

4.5 Torsion Elements

4.6 Annihilators

4.7 Free Modules

4.8 Homomorphisms

4.9 Quotient Modules

4.10 The Correspondence and Isomorphism Theorems

4.11 Direct Sums and Direct Summands

4.12 Modules Suck

Chapter 5

Modules II: Free and Noetherian Modules

5.1 The Rank of a Free Module

5.2 Free Modules and Epimorphisms

5.3 Noetherian Modules

5.4 The Hilbert Basis Theorem

Chapter 6

Modules over a Principal Ideal Domain

6.1 Annihilators and Orders

6.2 Cyclic Modules

6.3 Free Modules over a Principal Ideal Domain

6.4 Torsion-Free and Free Modules

6.5 The Primary Cyclic Decomposition Theorem

6.6 The Invariant Factor Decomposition

6.7 Characterizing Cyclic Modules

6.8 Indecomposable Modules

Chapter 7

The Structure of Linear Operators

7.1 The Module Associated with a Linear Operator

7.2 The Primary Cyclic Decomposition of V_τ

7.3 The Characteristic Polynomial

7.4 Cyclic and Indecomposable Modules

7.5 The Big Picture

7.6 The Rational Canonical Form

Chapter 8

Eigenvalues and Eigenvectors

8.1 Eigenvalues and Eigenvectors

8.2 Geometric and Algebraic Multiplicities

8.3 The Jordan Canonical Form

8.4 Triganularizability and Schur's Theorem

8.5 Diagonalizable Operators

Chapter 9

Real and Complex Inner Product Spaces

9.1 Norm and Distance

9.2 Isometries

9.3 Orthogonality

9.4 Orthogonal and Orthonormal Sets

9.5 The Projection Theorem and Best Approximations

9.6 The Riesz Representation Theorem

Chapter 10

Structure Theory for Normal Operators

10.1 The Adjoint of a Linear Operator

10.2 Orthogonal Projections

10.3 Unitary Diagonalizability

10.4 Normal Operators

10.5 Special Types of Normal Operators

10.6 Self-Adjoint Operators

10.7 Unitary Operators and Isometries

10.8 The Structure of Normal Operators

10.9 Functional Calculus

10.10 Positive Operators

10.11 The Polar Decomposition of an Operator