

Linear Algebra

REU 2004. Info:

<http://people.cs.uchicago.edu/~laci/reu04>.

Instructor: László Babai

07/18/04

Contents

1		5
1.1	Vector spaces, linear independence	5
1.2	Basis	8
2		9
2.1	Subspaces	9
2.2	All bases are equal	9
2.3	Coordinates	10
2.4	Linear maps, isomorphism of vector spaces	11
2.5	Vector spaces over number fields	12
2.6	Elementary operations	13
3		15
3.1	Rank	15
3.2	Number Fields, Roots of Unity	16
3.3	Irreducible Polynomials	18
4		21
4.1	Linear maps	21
5		27
5.1	Fields	27
5.2	Polynomials, rational functions	30

6	33
6.1 Inverse matrix, rank	33
6.2 Similarity of matrices, characteristic polynomial	36
7	39
7.1 Review	39
7.1.1 Matrices	39
7.1.2 Change of basis	39
7.2 Characteristic Polynomials	40
7.2.1 Similar matrices	40
7.2.2 Characteristic Polynomial	40
7.3 The Determinant	43
7.3.1 Upper Triangular Matrices	43
7.3.2 Determinant	44
7.3.3 Permutations and Parity	44
7.4 Eigenvectors and Eigenvalues	47
7.4.1 Diagonalizability	47
7.4.2 Eigenvalues	47
8	49
8.1 Cauchy-Hilbert matrix	49
8.2 Eigenvectors, eigenvalues	50
9	53
9.1 Spectral Theorem	53
10	57
10.1 Vandermonde Matrix	57
10.2 Real Euclidean Spaces, bilinear forms	57
11	63
11.1 Complex vector spaces, sesquilinear forms	63

<i>CONTENTS</i>	5
12	67
12.1 Complex Euclidean (unitary) spaces	67
12.2 Unitary transformations	68
12.3 Hermitian forms and self-adjoint transformations	69
12.4 Normal matrices	70
12.5 Gramian	72
13	75
13.1 Explicit form of the inverse matrix	75
13.2 Gram-Schmidt orthogonalization	76
13.3 Algebraic numbers, minimal polynomials	77
13.4 The minimal polynomial of a matrix	78
14	81
14.1 Rank inequalities	81
14.2 Rings	81
14.3 Ideals	82
14.4 Minimal polynomial	84

Chapter 1

1st day, Monday 6/28/04 (Scribe: Daniel Štefankovič)

1.1 Vector spaces, linear independence

Definition 1.1. A **vector space** is a set V with

- (a) addition $V \times V \rightarrow V$, $(x, y) \mapsto x + y$, and
- (b) scalar multiplication $\mathbb{R} \times V \rightarrow V$, $(\alpha, x) \mapsto \alpha x$,

satisfying following axioms

- (a) $(V, +)$ is an abelian group, i. e.
 - (a1) $(\forall x, y \in V)(\exists! x + y \in V)$,
 - (a2) $(\forall x, y \in V)(x + y = y + x)$ (commutative law),
 - (a3) $(\forall x, y, z \in V)((x + y) + z = x + (y + z))$ (associative law),
 - (a4) $(\exists 0 \in V)(\forall x)(x + 0 = 0 + x = x)$ (existence of zero),
 - (a5) $(\forall x \in V)(\exists(-x) \in V)(x + (-x) = 0)$,
- b)
 - (b1) $(\forall \alpha, \beta \in \mathbb{R})(\forall x \in V)(\alpha(\beta x)) = (\alpha\beta)x$ ("associativity" linking two operations),
 - (b2) $(\forall \alpha, \beta \in \mathbb{R})(\forall x \in V)((\alpha + \beta)x = \alpha x + \beta x)$ (distributivity over scalar addition),
 - (b3) $(\forall \alpha \in \mathbb{R})(\forall x, y \in V)(\alpha(x + y) = \alpha x + \alpha y)$ (distributivity over vector addition),
- (c) $(\forall x \in V)(1 \cdot x = x)$ (normalization).

Exercise 1.2. Show $(\forall x \in V)(0x = 0)$. (The first 0 is a number, the second a vector.)

Exercise 1.3. Show $(\forall \alpha \in \mathbb{R})(\alpha 0 = 0)$.

Exercise 1.4. Show $(\forall \alpha \in \mathbb{R})(\forall x \in V)(\alpha x = 0 \Leftrightarrow (\alpha = 0 \text{ or } x = 0))$

Definition 1.5. A **linear combination** of vectors $v_1, \dots, v_k \in V$ is a vector $\alpha_1 v_1 + \dots + \alpha_k v_k$ where $\alpha_1, \dots, \alpha_k \in \mathbb{R}$. The **span** of $v_1, \dots, v_k \in V$ is the set of all linear combinations of v_1, \dots, v_k , i. e., $\text{Span}(v_1, \dots, v_k) = \{\alpha_1 v_1 + \dots + \alpha_k v_k \mid \alpha_1, \dots, \alpha_k \in \mathbb{R}\}$.

Remark 1.6. We let $\text{Span}(\emptyset) = \{0\}$.

Remark 1.7. A linear combination of an infinite set of vectors $S \subseteq V$ is a linear combination of a finite subset of S .

Note that 0 is always in $\text{Span}(v_1, \dots, v_k)$ because the trivial linear combination $(\forall i)\alpha_i = 0$ is $0 \cdot v_1 + \dots + 0 \cdot v_k = 0$.

Definition 1.8. Vectors $v_1, \dots, v_k \in V$ are **linearly independent** if only the trivial linear combination gives 0, i. e., $\alpha_1 v_1 + \dots + \alpha_k v_k = 0 \Rightarrow \alpha_1 = \dots = \alpha_k = 0$.

Exercise 1.9. Which one element sets of vectors are linearly independent?

Exercise 1.10. Show that if $T \subseteq S \subseteq V$ and S is linearly independent then T is linearly independent.

We say that vectors $u, v \in V$ are **parallel** if $u, v \neq 0$ and $\exists \alpha \in \mathbb{R}$ such that $u = \alpha v$.

Exercise 1.11. Show that vectors $u, v \in V$ are linearly dependent if and only if $a = 0$ or $b = 0$ or a, b are parallel.

Exercise 1.12. An infinite set of vectors is linearly independent if and only if all finite subsets are linearly independent.

Remark 1.13. We say that a property P is a **finitary property** if a set S has the property P if and only if all finite subsets of S have property P .

Exercise 1.14. * (**Erdős – deBruijn**) Show that 3-colorability of a graph is a finitary property. (The same holds for 4-colorability, etc.)

The set of all polynomials with real coefficients is a vector space $\mathbb{R}[x]$.

Exercise 1.15. Show that $1, x, x^2, \dots$ are linearly independent.

Definition 1.16. The polynomial $f(x) = \sum a_i x^i$ has **degree** k if $a_k \neq 0$, but $(\forall j > k)(a_j = 0)$. Notation: $\deg(f) = k$. We let $\deg(0) = -\infty$. Note: the nonzero constant polynomials have degree 0.

Exercise 1.17. Prove: $\deg(fg) = \deg(f) + \deg(g)$. (Note that this remains true if one of the polynomials f, g is the zero polynomial.)

Exercise 1.18. Prove: $\deg(f + g) \leq \max\{\deg(f), \deg(g)\}$.

Exercise 1.19. Prove that if f_0, f_1, f_2, \dots is a sequence of polynomials, $\deg(f_i) = i$ then f_0, f_1, f_2, \dots are linearly independent.

Exercise 1.20. Let $f(x) = (x - \alpha_1)\dots(x - \alpha_k)$ where $\alpha_i \neq \alpha_j$ for $i \neq j$. Let $g_i(x) = f(x)/(x - \alpha_i)$. Show that g_1, \dots, g_k are linearly independent.

Exercise 1.21. Prove: for all $\alpha, \beta \in \mathbb{R}$, $\sin(x), \sin(x + \alpha), \sin(x + \beta)$ are linearly dependent functions $\mathbb{R} \rightarrow \mathbb{R}$.

Exercise 1.22. Prove: $1, \sin(x), \sin(2x), \sin(3x), \dots, \cos(x), \cos(2x), \dots$ are linearly independent functions $\mathbb{R} \rightarrow \mathbb{R}$.

Definition 1.23. A **maximal** linearly independent subset of a set $S \subseteq V$ is a subset $T \subseteq S$ such that

- (a) T is linearly independent, and
- (b) if $T \subsetneq T' \subseteq S$ then T' is linearly dependent.

Definition 1.24. A **maximum** linearly independent subset of a set $S \subseteq V$ is a subset $T \subseteq S$ such that

- (a) T is linearly independent, and
- (b) if $T' \subseteq S$ is linearly independent then $|T| \geq |T'|$.

Exercise 1.25. (Independence of vertices in a graph.) Show that 6-cycle, there exists a maximum independent set of vertices which is not maximal.

We shall see that this cannot happen with linear independence: every maximal linearly independent set is maximum.

Exercise 1.26. Let $S \subseteq V$. Then there exists $T \subseteq S$ such that T is a maximal independent subset of S .

Exercise 1.27. Let $L \subseteq S \subseteq V$. Assume L is linearly independent. Then there exists a maximal linearly independent subset $T \subseteq S$ such that $L \subseteq T$. (Every linearly independent subset of S set can be extended to a maximal linearly independent subset of S .)

Remark 1.28. This is easy to prove Ex. 1.26 by successively adding vectors until our set becomes maximal as long as all linearly independent subsets of S are finite. For the infinite case, we need an axiom from set theory called Zorn's Lemma (a version of the Axiom of Choice).

Definition 1.29. A vector $v \in V$ **depends** on $S \subseteq V$ if $v \in \text{Span}(S)$, i.e. v is a linear combination of S .

Definition 1.30. A set of vectors $T \subseteq V$ **depends** on $S \subseteq V$ if $T \subseteq \text{Span}(S)$.

Exercise 1.31. Show that dependence is transitive: if $R \subseteq \text{Span}(T)$ and $T \subseteq \text{Span}(S)$ then $R \subseteq \text{Span}(S)$.

Exercise 1.32. Suppose that $\sum \alpha_i v_i$ is a nontrivial linear combination. Then $(\exists i)$ such that v_i depends on the rest (i. e. on $\{v_j \mid j \neq i\}$). Indeed, this will be the case whenever $\alpha_i \neq 0$.

Exercise 1.33. If v_1, \dots, v_k are linearly independent and v_1, \dots, v_k, v_{k+1} are linearly dependent then v_{k+1} depends on v_1, \dots, v_k .

Theorem 1.34 (Fundamental Fact of Linear Algebra). *If v_1, \dots, v_k are linearly independent and $v_1, \dots, v_k \in \text{Span}(w_1, \dots, w_\ell)$ then $k \leq \ell$.*

Corollary 1.35. *All maximal independent sets are maximum.*

Exercise 1.36. If $T \subseteq S$, T is a maximal independent subset of S then $S \subseteq \text{Span}(T)$.

Exercise 1.37. Prove Corollary 1.35 from Theorem 1.34 and Exercise 1.36.

Definition 1.38. For $S \subseteq V$, the **rank** of S is the common cardinality of all the maximal independent subsets of S . Notation: $\text{rk}(S)$.

Definition 1.39. The **dimension** of a vector space is $\dim(V) := \text{rk}(V)$.

Exercise 1.40. Show that $\dim(\mathbb{R}^n) = n$.

Exercise 1.41. Let P_k be the space of polynomials of degree $\leq k$. Show that $\dim(P_k) = k + 1$.

Exercise 1.42. Let $T = \{\sin(x + \alpha) \mid \alpha \in \mathbb{R}\}$. Prove $\text{rk}(T) = 2$.

1.2 Basis

Definition 1.43. A **basis** of V is a linearly independent set which spans V .

Definition 1.44. A **basis** of $S \subseteq V$ is a linearly independent subset of S which spans S . In other words, a basis B of S is a linearly independent set satisfying $B \subseteq S \subseteq \text{Span}(B)$.

Exercise 1.45. B is a basis of S if and only if B is a maximal independent subset of S .

Exercise 1.46. Prove: if B is a basis of V then $\dim(V) = |B|$.

Exercise 1.47. A “Fibonacci-type sequence” is a sequence (a_0, a_1, a_2, \dots) such that $(\forall n)(a_{n+2} = a_{n+1} + a_n)$.

- Prove that the Fibonacci-type sequences form a 2-dimensional vector space.
- Find a basis in this space consisting of two geometric progressions.
- Express the Fibonacci sequence $(0, 1, 1, 2, 3, 5, 8, 13, \dots)$ as a linear combination of the basis found in item (b).

Chapter 2

2nd day, Tuesday 6/29/04 (Scribe: Justin Noel)

2.1 Subspaces

Definition 2.1. A subset $U \subseteq V$ is a **subspace** (written $U \leq V$) if

- (a) $0 \in U$
- (b) $(\forall \alpha \in \mathbb{R})(\alpha u \in U)$
- (c) $(\forall u, v \in U)(u + v \in U)$

Exercise 2.2. Show $\forall S \subseteq V$, $\text{Span}(S) \leq V$, i.e. that $\text{Span}(S)$ is a subspace of V . (Recall that $\text{Span}(\emptyset) = \{0\}$.)

Exercise 2.3. Show that $\text{Span}(S)$ is the smallest subspace containing S , i.e.

- (a) $\text{Span}(S) \leq V$ and $\text{Span}(S) \supseteq S$.
- (b) If $W \leq V$ and $S \subseteq W$ then $\text{Span}(S) \leq W$.

Corollary 2.4. $\text{Span}(S) = \bigcap_{S \subseteq W \leq V} W$.

Exercise 2.5. Show that the intersection of any set of subspaces is a subspace.

Remark 2.6. Note that this doesn't hold true for unions.

Exercise 2.7. If $U_1, U_2 \leq V$ then $U_1 \cup U_2$ is a subspace if and only if $U_1 \subseteq U_2$ or $U_2 \subseteq U_1$.

2.2 All bases are equal

Theorem 2.8 (Fundamental Fact of Linear Algebra). *If $L, M \subseteq V$ with L is a linearly independent set of vectors, and $L \leq \text{Span}(M)$ then $|L| \leq |M|$.*

Lemma 2.9 (Steinitz Exchange Principle). *If v_1, \dots, v_k are linearly independent and $v_1, \dots, v_k \in \text{Span}(w_1, \dots, w_\ell)$ then $\exists j, 1 \leq j \leq \ell$ such that w_j, v_2, \dots, v_k are linearly independent. (Note in particular that $w_j \neq v_2, \dots, v_k$.)*

Exercise 2.10. Prove the Steinitz Exchange Principle.

Exercise 2.11. Prove the Fundamental Fact using the Steinitz Exchange Principle.

Definition 2.12. An $m \times n$ **matrix** is an $m \times n$ array of numbers $\{\alpha_{ij}\}$ which we write as

$$\begin{pmatrix} \alpha_{11} & \cdots & \alpha_{1n} \\ \vdots & \ddots & \vdots \\ \alpha_{m1} & \cdots & \alpha_{mn} \end{pmatrix}$$

Definition 2.13. The **row (respectively column) rank** of a matrix is the rank of the set of row (respectively column) vectors.

Theorem 2.14 (The Most Amazing Fact of Basic Linear Algebra). *The row rank of a matrix is equal to its column rank. (To be proven later in today.)*

Definition 2.15. Let $S \subseteq V$, then a subset $B \subseteq S$ is a **basis** of S if

- (a) B is linearly independent.
- (b) $S \leq \text{Span}(B)$.

Exercise 2.16. Show that the statement: “**All bases for S have equal size**” is equivalent to Theorem 2.8.

Definition 2.17. We call the common size of all bases of S the **rank** of S , denoted $\text{rk}(S)$.

2.3 Coordinates

Exercise 2.18. Show that $B \subseteq S$ is a basis if and only if B is a maximal linearly independent subset of S .

Exercise 2.19. If B is a basis of S then $\forall x \in S$ there exists a unique linear combination of elements in B that sums to x . In other words for all x there are unique scalars β_i such that

$$x = \sum_{i=1}^k \beta_i b_i.$$

Definition 2.20. For a basis B , regarded as an ordered set of vectors, we associate to each $x \in S$ the column vector

$$[x]_B := \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_k \end{pmatrix}$$

called the **coordinates** of x , where the β_i are as above.

2.4 Linear maps, isomorphism of vector spaces

Definition 2.21. Let V and W be vector spaces. We say that a map $f : V \rightarrow W$ is a **homomorphism** or a **linear map** if

$$(a) (\forall x, y \in V)(f(x + y) = f(x) + f(y))$$

$$(b) (\forall x \in V)(\forall \alpha \in \mathbb{R})(f(\alpha x) = \alpha f(x))$$

Exercise 2.22. Show that if f is a linear map then $f(0) = 0$.

Exercise 2.23. Show that $f(\sum_{i=1}^k \alpha_i v_i) = \sum_{i=1}^k \alpha_i f(v_i)$.

Definition 2.24. We say that f is an **isomorphism** if f is a bijective homomorphism.

Definition 2.25. Two spaces V and W are **isomorphic** if there exists an isomorphism between them.

Exercise 2.26. Show the relation of being isomorphic is an equivalence relation.

Exercise 2.27. Show that an isomorphism maps bases to bases.

Theorem 2.28. If $\dim(V) = n$ then $V \cong \mathbb{R}^n$.

Proof: Choose a basis, B of V , now map each vector to its coordinate vector, i.e. $v \mapsto [v]_B$.

Definition 2.29. We denote the **image** of f as the set

$$\text{im}(f) = \{f(x) : x \in V\}$$

Definition 2.30. We denote the **kernel** of f as the set

$$\text{ker}(f) = \{x \in V : f(x) = 0\}$$

Exercise 2.31. For a linear map $f : V \rightarrow W$ show that $\text{im}(f) \leq W$ and $\text{ker}(f) \leq V$.

Theorem 2.32. For a linear map $f : V \rightarrow W$ we have

$$\dim \text{ker}(f) + \dim \text{im}(f) = \dim V.$$

Lemma 2.33. If $U \leq V$ and A is a basis of U then A can be extended to a basis of V .

Exercise 2.34. Prove Theorem 2.32. HINT: apply Lemma 2.33 setting $U = \text{ker}(f)$.

2.5 Vector spaces over number fields

Definition 2.35. A subset $\mathbb{F} \subseteq \mathbb{C}$ is a **number field** if \mathbb{F} is closed under the four arithmetic operations, i.e. for $\alpha, \beta \in \mathbb{F}$

- (a) $\alpha \pm \beta \in \mathbb{F}$
- (b) $\alpha\beta \in \mathbb{F}$
- (c) $\frac{\alpha}{\beta} \in \mathbb{F}$ (assuming $\beta \neq 0$).

Exercise 2.36. Show that if \mathbb{F} is a number field then $\mathbb{Q} \subseteq \mathbb{F}$.

Exercise 2.37. Show that $\mathbb{Q}[\sqrt{2}]$ is a number field.

Exercise 2.38. Show that $\mathbb{Q}[\sqrt[3]{2}] = \{a + b\sqrt[3]{2} + c\sqrt[3]{4} : a, b, c \in \mathbb{Q}\}$ is a number field.

Exercise 2.39 (Vector Spaces over Number Fields). Convince yourself that all of the things we have said about vector spaces remain valid if we replace \mathbb{R} and \mathbb{F} .

Exercise 2.40. Show that if \mathbb{F}, G are number fields and $\mathbb{F} \subseteq G$ then G is a vector space over \mathbb{F} .

Exercise 2.41. Show that $\dim_{\mathbb{R}}\mathbb{C} = 2$.

Exercise 2.42. Show that $\dim_{\mathbb{Q}}\mathbb{R}$ has the cardinality of “continuum,” that is, it has the same cardinality as \mathbb{R} .

Exercise 2.43 (Cauchy’s Equation). We consider functions $f : \mathbb{R} \rightarrow \mathbb{R}$ satisfying Cauchy’s Equation: $f(x + y) = f(x) + f(y)$ with $x, y \in \mathbb{R}$. For such a function prove that

- (a) If f is continuous then $f(x) = cx$.
- (b) If f is continuous at a point then $f(x) = cx$.
- (c) If f is bounded on some interval then $f(x) = cx$.
- (d) If f is measurable in some interval then $f(x) = cx$.
- (e) There exists a $g : \mathbb{R} \rightarrow \mathbb{R}$ such that $g(x) \neq cx$ but $g(x + y) = g(x) + g(y)$. (HINT: Use the fact that \mathbb{R} is a vector space over \mathbb{Q} . Use a basis of this vector space. Such a basis is called a **Hamel basis**.)

Exercise 2.44. Show that $1, \sqrt{2}$, and $\sqrt{3}$ are linearly independent over \mathbb{Q} .

Exercise 2.45. Show that $1, \sqrt{2}, \sqrt{3}, \sqrt{5}, \sqrt{6}, \sqrt{10}, \sqrt{15}$ and $\sqrt{30}$ are linearly independent over \mathbb{Q} .

Exercise 2.46. * Show that the set of square roots of all of the square-free integers are linearly independent over \mathbb{Q} . (An integer is **square free** if it is not divisible by the square of any prime number. For instance, 30 is square free but 18 is not.)

Definition 2.47. A **rational function** f over \mathbb{R} is a fraction of the form

$$f(x) = \frac{g(x)}{h(x)}$$

where $g, h \in \mathbb{R}[x]$ (that is g, h are real polynomials) and $h(x) \neq 0$ (h is not the identically zero polynomial). More precisely, a rational function is an equivalence class of fractions of polynomials, where the fractions $\frac{g_1(x)}{h_1(x)}$ and $\frac{g_2(x)}{h_2(x)}$ are equivalent if and only if $g_1 \cdot h_2 = g_2 \cdot h_1$. (This is analogous to the way fractions of integers represent rational numbers; the fractions $3/2$ and $6/4$ represent the same rational number.) We denote the set of all rational functions as $\mathbb{R}(x)$.

Note that a rational function is not a function; it is an equivalence class of formal quotients.

Exercise 2.48. Prove that the rational functions $\{\frac{1}{x-\alpha} : \alpha \in \mathbb{R}\}$ are linearly independent set over $\mathbb{R}(x)$.

Corollary 2.49. $\dim_{\mathbb{R}[x]} \mathbb{R}(x)$ has the cardinality of “continuum” (the same cardinality as \mathbb{R}).

2.6 Elementary operations

Definition 2.50. The following actions on a set of vectors $\{v_1, \dots, v_k\}$ are called **elementary operations**:

- (a) Replace v_i by $v_i - \alpha v_j$ where $i \neq j$.
- (b) Replace v_i by αv_i where $\alpha \neq 0$.
- (c) Switch v_i and v_j .

Exercise 2.51. Show that the rank of a list of vectors doesn't change under elementary operations.

Exercise 2.52. Let $\{v_1, \dots, v_k\}$ have rank r . Show that by a sequence of elementary operations we can get from $\{v_1, \dots, v_k\}$ to a set $\{w_1, \dots, w_k\}$ such that w_1, \dots, w_r are linearly independent and $w_{r+1} = \dots = w_k = 0$.

Consider a matrix. An **elementary row-operation** is an elementary operation applied to the rows of the matrix. Elementary column operations are defined analogously. Exercise 2.51 shows that **elementary row-operations** do not change the **row-rank** of A .

Exercise 2.53. Show that elementary **row-operations** do not change the **column-rank** of a matrix.

Exercise 2.54. Use Exercises 2.51 and 2.53 prove the “amazing” Theorem 2.14.

Chapter 3

3rd day, Wednesday 6/30/04 (Scribe: Richard Cudney)

3.1 Rank

Let V be a vector space.

Definition 3.1. Let $S \subseteq V$, and $T \subseteq S$. T is a set of **generators** of S if $S \subseteq \text{Span}(T)$.

Definition 3.2. A **basis** of S is a linearly independent set of generators of S .

Definition 3.3. The **rank** of a set S , $\text{rk}(S) := |B|$ for any basis B of S . Note that this is well-defined because of the Fundamental Fact.

Definition 3.4. Let $U \leq V$. The **dimension** of U , $\dim(U) := \text{rk}(U)$ (=maximal number of linearly independent vectors in U).

Exercise 3.5. $\dim \text{Span}(T) = \text{rk}(T)$ Hint: One direction is immediate and the other follows from the fundamental fact.

Given a matrix, there are a priori two different ranks associated to it, the rank of the set of column vectors, and the rank of the set of row vectors (column-rank and row-rank respectively). It is an Amazing Fact that row-rank=column-rank. The following exercise gives a proof of this fact using Gaussian elimination.

Exercise 3.6. (a) Elementary row operations **change** which **sets** of rows are linearly independent, while the maximum number of linearly independent rows remains the same.

(b) Elementary column operations do not affect the linear independence of any given set of rows.

(c) By applying elementary row and column operations any matrix can be made to have zeroes everywhere outside a square sub-matrix in the upper left hand corner, in which it will have ones down the diagonal and zero elsewhere.

Exercise 3.7 (Rank invariance under field extensions). If \mathbb{F}, G are fields, $\mathbb{F} \leq G$ and A is a matrix over \mathbb{F} , then $\text{rk}_{\mathbb{F}}(A) = \text{rk}_G(A)$.

3.2 Number Fields, Roots of Unity

Exercise 3.8. If \mathbb{F} is a number field then $\mathbb{F} \supseteq \mathbb{Q}$.

Recall the previous exercise that asked to show that $\mathbb{Q}[\sqrt{2}] = \{a + b\sqrt{2} : a, b \in \mathbb{Q}\}$ is a number field. The only non-obvious part was that we could divide by non-zero elements. We can accomplish division by multiplying by the conjugate:

$$\frac{1}{a + b\sqrt{2}} = \frac{1}{a + b\sqrt{2}} \cdot \frac{a - b\sqrt{2}}{a - b\sqrt{2}} = \frac{a - b\sqrt{2}}{a^2 - 2b^2}$$

Exercise 3.9. Let $a, b \in \mathbb{Q}$. If $a^2 - 2b^2 = 0$ then $a = b = 0$.

Now how can we generalize this trick to show that $\mathbb{Q}[\sqrt[3]{2}]$ is a number field? What are the conjugates of $\sqrt[3]{2}$? Previously we used both roots of $x^2 - 2$, now we want to use all roots of $x^3 - 2$. What are the other roots beside $\sqrt[3]{2}$? Let $\omega = \cos(\frac{2\pi}{3}) + i\sin(\frac{2\pi}{3})$ so that $\omega^2 = \cos(\frac{4\pi}{3}) + i\sin(\frac{4\pi}{3})$. These are the non-real roots of $x^3 = 1$, so $\omega\sqrt[3]{2}, \omega^2\sqrt[3]{2}$ are the other cube roots of 2.

Note:

$$\omega = -\frac{1}{2} + i\frac{\sqrt{3}}{2}$$

and

$$\omega^2 = -\frac{1}{2} - i\frac{\sqrt{3}}{2}.$$

This example leads us to study the n -th roots of unity-the complex solutions of

$$x^n = 1.$$

We can calculate the n -th roots as follows, writing x in polar form:

$$x = r(\cos(\alpha) + i\sin(\alpha))$$

$$x^n = r^n(\cos(n\alpha) + i\sin(n\alpha)) = 1 + i0$$

So $r = 1$, and $\alpha = \frac{2k\pi}{n}$.

Exercise 3.10. Let S_n be the sum of all n th roots of unity. Show that $S_0 = 1$ and $S_n = 0$ for $n \geq 1$.

Definition 3.11. z is a **primitive** n -th root of unity if $z^n = 1$ and $z^j \neq 1$ for $1 \leq j \leq n - 1$.

Let

$$\zeta_n := \cos\left(\frac{2\pi}{n}\right) + i\sin\left(\frac{2\pi}{n}\right) = e^{2\pi i/n}$$

Exercise 3.12. $1, \zeta_n, \dots, \zeta_n^{n-1}$ are all of the n -th roots of unity.

Exercise 3.13. Let $z^n = 1$. Then the powers of z give all n -th roots of unity iff z is a primitive n -th root of unity.

Exercise 3.14. Suppose z is a primitive n -th root of unity. For what k is z^k also a primitive n -th root of unity?

Exercise 3.15. If z is an n -th root of unity then z^k is also an n -th root of unity.

Definition 3.16. The **order** of a complex number is the smallest positive n such that $z^n = 1$. (If no such n exists then we say z has infinite order.)

$$\text{ord}(-1) = 2, \text{ord}(\omega) = 3, \text{ord}(i) = 4, \text{ord}(1) = 1, \text{ord}(\pi) = \infty.$$

Exercise 3.17. $\text{ord}(z) = n$ iff z is a primitive n -th root of unity.

Exercise 3.18. Let $\mu(n)$ be the sum of all primitive n -th roots of unity.

- a) Prove that for every n , $\mu(n) = 0, 1$, or -1 .
- b) Prove $\mu(n) \neq 0$ iff n is square free.
- c) Prove if $\text{g.c.d.}(k, \ell) = 1$ then $\mu(k\ell) = \mu(k)\mu(\ell)$.
- d) If $n = p_1^{t_1} \dots p_k^{t_k}$, find an explicit formula for $\mu(n)$ in terms of the t_i .

Exercise 3.19. Show that the number of primitive n -th roots of unity is equal to Euler's phi function. $\varphi(n) :=$ number of k such that $1 \leq k \leq n$, $\text{g.c.d.}(k, n) = 1$.

n	$\varphi(n)$
1	1
2	1
3	2
4	2
5	4
6	2
7	6
8	4
9	6
10	4

Definition 3.20. $f : \mathbb{N}^+ \rightarrow \mathbb{C}$ is **multiplicative** if $(\forall k, \ell)(\text{if } \text{g.c.d.}(k, \ell) = 1 \text{ then } f(k\ell) = f(k)f(\ell))$.

Definition 3.21. f is **totally multiplicative** if $(\forall k, \ell)(f(k\ell) = f(k)f(\ell))$.

Exercise 3.22. μ function is multiplicative.

Exercise 3.23. φ function is multiplicative.

Exercise 3.24. Neither μ nor φ are totally multiplicative.

Exercise 3.25. Prove that
$$\sum_{d|n, 1 \leq d \leq n} \varphi(d) = n.$$

Remark 3.26. We call

$$g(n) = \sum_{d|n, 1 \leq d \leq n} f(d)$$

the **summation function** of f .

Exercise 3.27. f is multiplicative if and only if g is.

Now, using the preceding ideas, we can apply in $\mathbb{Q}[\sqrt[3]{2}]$ the same construction we used in $\mathbb{Q}[\sqrt{2}]$:

$$\frac{1}{a + \sqrt[3]{2}b + \sqrt[3]{4}c} \cdot \frac{a + \omega \sqrt[3]{2}b + \omega^2 \sqrt[3]{4}c}{a + \omega \sqrt[3]{2}b + \omega^2 \sqrt[3]{4}c} \cdot \frac{a + \omega^2 \sqrt[3]{2}b + \omega \sqrt[3]{4}c}{a + \omega^2 \sqrt[3]{2}b + \omega \sqrt[3]{4}c}.$$

Exercise 3.28. Show that the denominator in the above expression is rational and non-zero.

3.3 Irreducible Polynomials

Definition 3.29. Let \mathbb{F} be a number field. $\mathbb{F}[x]$ is the ring of all univariate polynomials with coefficients in \mathbb{F} .

Definition 3.30. f is **irreducible** over \mathbb{F} if

- (i) $\deg(f) \geq 1$ and
- (ii) $(\forall g, h) (\in \mathbb{F}[x], f = gh \rightarrow \deg(f) = 0 \text{ or } \deg(g) = 0)$.

Remark 3.31. If $\deg(f) = 1$, then f is irreducible because degree is additive.

Theorem 3.32 (Fundamental Theorem of Algebra). *If $f \in \mathbb{C}[x]$ and $\deg(f) \geq 1$ then $(\exists \alpha \in \mathbb{C})(f(\alpha) = 0)$.*

Exercise 3.33. Over \mathbb{C} a polynomial is irreducible iff it is of degree 1. HINT: Follows from the FTA and the exercise 3.35 that lets you pull out roots.

Definition 3.34. Let $f, g \in \mathbb{F}[x]$. We say $f | g$ if $(\exists h)(fh = g)$.

Exercise 3.35. $(\forall \alpha)(x - \alpha) | (f(x) - f(\alpha))$.

Corollary 3.36. α is a root of f iff $(x - \alpha) \mid f(x)$.

Remark 3.37. If $f(x) = x^n$, then

$$x^n - \alpha^n = (x - \alpha)(x^{n-1} + \alpha x^{n-2} + \dots + \alpha^{n-1}).$$

Exercise 3.38. $f(x) = ax^2 + bx + c$ is irreducible over \mathbb{R} iff $b^2 - 4ac < 0$.

Remark 3.39. Odd degree polynomials over \mathbb{R} always have real roots.

Exercise 3.40. If $f \in \mathbb{R}[x]$, and $z \in \mathbb{C}$, then $f(\bar{z}) = \overline{f(z)}$.

Consequence: If z is a root of f then so is \bar{z} .

Theorem 3.41. Over \mathbb{R} , all irreducible polynomials have $\deg \leq 2$.

Proof: Suppose $f \in \mathbb{R}[x]$, $\deg(f) \geq 3$. We want to show that f is not irreducible over \mathbb{R} .

- 1) If f has a real root α , then $(x - \alpha) \mid f$.
- 2) Otherwise by FTA f has a complex root z which is not real, so that $z \neq \bar{z}$. Thus $(x - z)(x - \bar{z}) = x^2 - 2ax + a^2 + b^2$ divides f , where $z = a + bi$.

□

Theorem 3.42 (Gauss Lemma). If $f = gh$, $f \in \mathbb{Z}[x]$, $g, h \in \mathbb{Q}[x]$ then $\exists \alpha \in \mathbb{Q}$ such that $\alpha g \in \mathbb{Z}[x]$ and $\frac{h}{\alpha} \in \mathbb{Z}[x]$.

Exercise 3.43. If a_1, \dots, a_n are distinct integers, then $\prod_{i=1}^{i=n} (x - a_i) - 1$ is irreducible over \mathbb{Q} .

Exercise 3.44. $\forall n, x^n - 2$ is irreducible over \mathbb{Q} .

Definition 3.45. The n -th **cyclotomic polynomial** is $\Phi_n(x) = \prod (x - \zeta)$, where ζ ranges over the primitive n -th roots of unity.

Remark 3.46. $\deg \Phi_n(x) = \varphi(n)$.

$$\begin{aligned} \Phi_1(x) &= x - 1 \\ \Phi_2(x) &= x + 1 \\ \Phi_3(x) &= x^2 + x + 1 \\ \Phi_4(x) &= x^2 + 1 \\ \Phi_5(x) &= x^4 + x^3 + x^2 + x + 1 \\ \Phi_6(x) &= x^2 - x + 1 \\ \Phi_7(x) &= x^6 + x^5 + x^4 + x^3 + x^2 + x + 1 \\ \Phi_8(x) &= x^4 + 1 \end{aligned}$$

Exercise 3.47. $x^n - 1 = \prod_{d|n, 1 \leq d \leq n} \Phi_d(x)$.

Exercise 3.48. $\Phi_n(x) \in \mathbb{Z}[x]$.

Theorem 3.49. * $\Phi_n(x)$ is irreducible over \mathbb{Q} .

Definition 3.50. $\alpha \in \mathbb{C}$ is **algebraic** if $(\exists f)(f \in \mathbb{Q}[x], f \neq 0, f(\alpha) = 0)$.

Definition 3.51. A **minimal polynomial** for α is a nonzero polynomial $m_\alpha(x) \in \mathbb{Q}[x]$ such that

$$m_\alpha(\alpha) = 0 \tag{3.1}$$

and $m_\alpha(x)$ has minimal degree, among polynomials satisfying (3.1).

Remark 3.52. The minimal polynomial is well-defined up to a constant multiple.

Exercise 3.53. $m_\alpha(x)$ is irreducible.

Definition 3.54. $\deg(\alpha) = \deg(m_\alpha)$.

Exercise 3.55. If ζ is a primitive n th root of unity then $\deg(\zeta) = \varphi(n)$.

Exercise 3.56. $(\forall f \in \mathbb{Q}[x])(f(\alpha) = 0 \iff m_\alpha | f)$

Definition 3.57. The **algebraic conjugates** of α are the roots of m_α .

Exercise 3.58. If $\deg(\alpha) = n$ then the set

$$\mathbb{Q}[a] := \{a_0 + a_1\alpha + a_{n-1}\alpha^{n-1} \mid a_i \in \mathbb{Q}\}$$

is a field.

Chapter 4

4th day, Thursday 7/1/04 (Scribe: Nick Gurski)

4.1 Linear maps

We shall study the notion of maps between vector spaces, called linear maps or homomorphisms. A function $f : V \rightarrow W$ between vector spaces defined over the same field \mathbb{F} is a **linear map** if

$$\begin{aligned} (1) \quad & f(\underline{x} + \underline{y}) = f(\underline{x}) + f(\underline{y}) \\ (2) \quad & f(\alpha \underline{x}) = \alpha f(\underline{x}) \end{aligned}$$

for all $\underline{x}, \underline{y} \in V$ and all $\alpha \in \mathbb{F}$; these two conditions are equivalent to the single condition that f distributes over linear combinations

$$f\left(\sum_i \alpha_i \underline{x}_i\right) = \sum_i \alpha_i f(\underline{x}_i).$$

Now assume that we have a basis

$$\mathbf{e} = \{\underline{e}_1, \underline{e}_2, \dots, \underline{e}_n\}$$

for V and a basis \mathbf{e}' for W . We would like to assign a matrix to the linear map f which depends on the two bases we have chosen and represents the function f completely. We do this by defining the matrix $[f]_{(\mathbf{e}, \mathbf{e}')}$ to have as its i -th column the vector $[f(\underline{e}_i)]_{\mathbf{e}'}$. Thus we can write this matrix as

$$[f]_{(\mathbf{e}, \mathbf{e}')} = [[f(\underline{e}_1)]_{\mathbf{e}' } \cdots [f(\underline{e}_n)]_{\mathbf{e}' }];$$

if W is of dimension k , this matrix is a $k \times n$ matrix. Our next goal is to show that this matrix $[f]_{(\mathbf{e}, \mathbf{e}')}$ gives the same geometric information as the linear map f .

We begin this task with a theorem.

Theorem 4.1. *Let $\mathbf{e} = \{\underline{e}_1, \dots, \underline{e}_n\}$ be a basis for V . Then for any set of vectors $\underline{w}_1, \dots, \underline{w}_n \in W$, there exists a unique linear map $f : V \rightarrow W$ with the property that $f(\underline{e}_i) = \underline{w}_i$ for all i .*

Proof: We first show uniqueness, and existence follows easily afterwards. Assuming there is such an f , then its value on a vector $\underline{x} \in V$ is determined since \mathbf{e} is a basis. Writing $\underline{x} = \sum_{i=1}^n \alpha_i \underline{e}_i$, we have that

$$f(\underline{x}) = f\left(\sum_{i=1}^n \alpha_i \underline{e}_i\right) = \sum_{i=1}^n \alpha_i f(\underline{e}_i);$$

this shows uniqueness.

To show the existence of such a map, we define

$$f(\underline{x}) = \sum_{i=1}^n \alpha_i \underline{w}_i$$

and check that this is a linear map. This is left as a simple exercise.

The theorem means that the matrix $[f]$ is determined by the images $f(\underline{e}_i)$ of each of the basis vectors in V , and by the definition of basis this determines the linear map f as well. Conversely, if we know where f sends each basis vector, then there is a unique linear map that agrees with f on \mathbf{e} . Thus the matrix $[f]$ and the linear map f give the same information.

Example 4.2 ((The derivative of polynomials)). Remember that the collection of all polynomials of degree n , denoted P_n , is a vector space of dimension $n + 1$ which is spanned by the basis $1, x, x^2, \dots, x^n$. There is a linear map D which sends the polynomial $p(x)$ to its derivative polynomial $p'(x)$; thus we have a linear map $D : P_n \rightarrow P_{n-1}$. We shall compute the matrix $[D]$ with respect to the natural bases. We know that $D(x^k) = kx^{k-1}$, so $D(\underline{e}_k) = k\underline{e}_{k-1}$ where k ranges from 0 to n . This shows that the matrix $[D]$ has, as its k -th column the vector which is all zeros except that its $(k - 1)$ -st entry is $(k - 1)$; this is an $n \times (n + 1)$ matrix. For $n = 3$, the matrix for $D : P_3 \rightarrow P_2$ is shown below.

$$[D] = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}$$

Example 4.3. Projection onto the plane:

We have a linear map $\pi : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ which forgets about the z -coordinate. Using the standard basis, we have

$$\begin{aligned} \pi(\underline{e}_1) &= \underline{e}_1 \\ \pi(\underline{e}_2) &= \underline{e}_2 \\ \pi(\underline{e}_3) &= \mathbf{0} \end{aligned}$$

and so the 2×3 matrix $[\pi]$ is

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$

Example 4.4. Rotation of the plane by the angle α : We denote by ρ_α the map which rotates \mathbb{R}^2 by the fixed angle α . Using the standard basis, we find that

$$[\rho_\alpha] = \begin{bmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{bmatrix}.$$

If we use a different basis, however, the matrix for ρ_α changes. Let \underline{e}_1 be the vector $(1, 0)$ as before, but now let \underline{e}_3 be the vector obtained by rotating \underline{e}_1 by α . In general, these vectors are linearly independent, and thus form a basis for \mathbb{R}^2 . Now we will compute $[\rho_\alpha]$ with respect to this basis. By definition, $\rho_\alpha(\underline{e}_1) = \underline{e}_3$; using some basic geometry we can determine $\rho_\alpha(\underline{e}_3)$ and thus find that the matrix $[\rho_\alpha]$ is now

$$\begin{bmatrix} 0 & -1 \\ 1 & 2 \cos \alpha \end{bmatrix}.$$

This examples shows that when computing the matrix associated to a linear map, it is important to remember what bases are involved.

We can now make two definitions that will be useful later.

Definition 4.5. Let A be an $n \times n$ matrix (it is important that A is a square matrix here). Then we define the **trace** of A , denoted $\text{tr}A$, to be the sum of the diagonal entries of A . If A is the 2×2 matrix

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix},$$

then we define the **determinant** of A to be

$$\det A = ad - bc.$$

We will later see how to define determinant for $n \times n$ matrices. If we are given a linear map $f : V \rightarrow V$, then the trace and determinant of the matrix $[f]$ do not depend on what basis we have chosen. In the examples above of the two matrices associated to the rotation map, it is easy to check that both matrices have trace $2 \cos \alpha$ and determinant 1.

Assume that we are given two matrices, $A = (a_{ij})$ and $B = (b_{jk})$ where A is an $r \times s$ matrix and B is an $s \times t$ matrix. Then we can multiply them to get a matrix $C = AB$, where $C = (c_{ik})$ and the entry c_{ik} is equal to

$$\sum_{j=1}^s a_{ij} b_{jk}.$$

Note that A must have the same number of rows as B has columns or the indices will not match as they should.

Let $f : V \rightarrow W$ be a linear map, with bases \mathbf{e} for V and \mathbf{e}' for W . Given a vector $\underline{x} \in V$, we would like to determine the coordinates of the vector $[f(\underline{x})]$ with respect to \mathbf{e}' in terms of the coordinates of \underline{x} with respect to \mathbf{e} and the matrix $[f]_{(\mathbf{e}, \mathbf{e}')}$.

Exercise 4.6. Using the definition of matrix multiplication above, we have the formula

$$[f]_{(\mathbf{e}, \mathbf{e}')} [\underline{x}]_{\mathbf{e}} = [f(\underline{x})]_{\mathbf{e}'}$$

We can use this exercise to study the relationship between the composite of two linear maps and the product of their matrices. Let $f : V \rightarrow W$ and $g : W \rightarrow Z$ be two linear maps, and let V, W, Z have bases $\mathbf{e}, \mathbf{e}', \mathbf{e}''$, respectively. We have the composite map gf which is defined on vectors as $(gf)(\underline{x}) = g(f(\underline{x}))$; we also have matrices $[f]_{(\mathbf{e}, \mathbf{e}'')}$ and $[g]_{(\mathbf{e}', \mathbf{e}''')}$, as well as a matrix for the composite gf with respect to \mathbf{e} and \mathbf{e}'' . The goal is to prove that

$$[gf]_{(\mathbf{e}, \mathbf{e}'')} = [g]_{(\mathbf{e}', \mathbf{e}'')} [f]_{(\mathbf{e}, \mathbf{e}')}$$

To do this, a lemma is required.

Lemma 4.7. *If A is a $k \times n$ matrix, then $A = 0$ if and only if for all $\underline{x} \in \mathbb{F}^n$, $A\underline{x} = 0$.*

Proof: If $A = 0$, it is obvious that $A\underline{x} = 0$ for all \underline{x} . Assume that $A \neq 0$; we will produce a vector \underline{x} such that $A\underline{x} \neq 0$, and that will complete the proof. Since A is nonzero, it has a nonzero entry; call it a_{ij} . Then $A\underline{e}_j = \sum_l a_{lj} \underline{e}_l$, where the vector \underline{e}_l is zero except for entry l which is 1. This vector is nonzero, since a_{ij} is nonzero.

Corollary 4.8. *If for all vectors \underline{x} , $A\underline{x} = B\underline{x}$, then $A = B$ as matrices.*

Proof: If $A\underline{x} = B\underline{x}$, then $(A - B)\underline{x} = 0$ for all \underline{x} and therefore $A - B = 0$ by the above.

Now we can show that $[gf] = [g][f]$ by showing that these two matrices are equal when applied to any vector \underline{x} . On the left side, we have

$$[gf][\underline{x}] = [(gf)(\underline{x})]$$

by the exercise above. On the right, we have

$$[g][f][\underline{x}] = [g][f(\underline{x})] = [g(f(\underline{x}))],$$

but $[(gf)(\underline{x})] = [g(f(\underline{x}))]$ since that is exactly how the composition of two functions is defined. This proves that $[gf] = [g][f]$.

Returning to our example of the linear map which rotates by α , we can use that $[\rho_\alpha][\rho_\beta] = [\rho_\alpha \rho_\beta]$ to prove the angle-sum formula from trigonometry. Since rotating by β and then rotating by α is the same as rotating by $\alpha + \beta$, we get that

$$[\rho_\alpha \rho_\beta] = [\rho_{\alpha+\beta}] = \begin{bmatrix} \cos(\alpha + \beta) & -\sin(\alpha + \beta) \\ \sin(\alpha + \beta) & \cos(\alpha + \beta) \end{bmatrix}.$$

Multiplying the matrices for ρ_α and ρ_β together, the result is

$$\begin{bmatrix} \cos \alpha \cos \beta - \sin \alpha \sin \beta & -(\sin \alpha \cos \beta + \cos \alpha \sin \beta) \\ \sin \alpha \cos \beta + \cos \alpha \sin \beta & \cos \alpha \cos \beta - \sin \alpha \sin \beta \end{bmatrix};$$

comparing these two matrices gives the desired identity.

Now we turn to using matrices to solve systems of linear equations. If $A = [\underline{a}_1 \underline{a}_2 \cdots \underline{a}_n]$ is a $k \times n$ matrix with columns \underline{a}_i , and \underline{x} is any vector, then $A\underline{x} = x_1\underline{a}_1 + x_2\underline{a}_2 + \cdots + x_n\underline{a}_n$. Therefore the equation $A\underline{x} = \underline{0}$, where \underline{x} is the variable, always has the trivial solution $\underline{x} = \underline{0}$.

Exercise 4.9. $A\underline{x} = \underline{0}$ has a nontrivial solution if and only if the columns of A are linearly dependent if and only if the rank of A is less than n .

Exercise 4.10. If A is a $k \times n$ matrix, then $A\underline{x} = \underline{0}$ has a nontrivial solution if $k < n$ (HINT: use Exercise 4.9).

Exercise 4.11 (Showing the existence of a nontrivial solution). Using the corollary above, we can show that the system of equations

$$\begin{aligned} 2x + 3y - z &= 0 \\ 7x + 8y + 5z &= 0 \end{aligned}$$

has a nontrivial solution. Solving this system is equivalent to solving $A\underline{x} = \underline{0}$ where

$$A = \begin{bmatrix} 2 & 3 & -1 \\ 7 & 8 & 5 \end{bmatrix}$$

and

$$\underline{x} = \begin{bmatrix} 7 \\ 8 \\ 5 \end{bmatrix}.$$

In this case, $k = 2$ and $n = 3$, so there must be a nontrivial solution.

If we try to solve the general equation $A\underline{x} = \underline{b}$, this is equivalent to solving

$$\sum_i x_i \underline{a}_i = \underline{b},$$

or finding out if \underline{b} is in the span of the vectors \underline{a}_i .

Exercise 4.12. The equation $A\underline{x} = \underline{b}$ is solvable if and only if $\underline{b} \in \text{Span}\{\underline{a}_i\}$ if and only if the rank of A is the same as the rank of $[A|\underline{b}]$.

Let $U = \{\underline{x} : A\underline{x} = \underline{0}\}$; this is called the solution space. It is easy to see that U is a subspace of \mathbb{F}^n . It is obvious that U contains the zero vector. If $\underline{x}, \underline{y} \in U$, then $(\underline{x} + \underline{y}) \in U$ since

$$A(\underline{x} + \underline{y}) = A\underline{x} + A\underline{y} = \underline{0} + \underline{0} = \underline{0}.$$

Similarly, if $\underline{x} \in U$, then $\alpha\underline{x} \in U$.

Theorem 4.13. The dimension of U is $n - \text{rk}(A)$.

Proof: Define $\varphi : \mathbb{F}^n \rightarrow \mathbb{F}^k$ by $\varphi(\underline{x}) = A\underline{x}$; φ is linear because A is a matrix. The kernel of φ is exactly the subspace U defined above. We know that $\dim(\ker \varphi) = n - \dim(\text{im} \varphi)$, so we only need to show that $\dim(\text{im} \varphi) = \text{rk}(A)$. But the image of φ is the set of all vectors of the form $A\underline{x}$, and this subspace is spanned by the columns of A , hence the dimension of the image is necessarily the rank of A .

It is important to note that $A\underline{x} = \underline{b}$ may or may not have a solution. Assume that it does have a solution, and call that solution \underline{x}_0 . To find all solutions to $A\underline{x} = \underline{b}$, we can subtract

$$\begin{array}{rcl} A\underline{x} & = & \underline{b} \\ - A\underline{x}_0 & = & \underline{b} \\ \hline A(\underline{x} - \underline{x}_0) & = & \underline{0} \end{array}$$

to find that all solutions of $A\underline{x} = \underline{b}$ are translates of the vector \underline{x}_0 by elements of U . Another way to say this is that $\underline{x} - \underline{x}_0$ is a solution to $A\underline{x} = \underline{0}$ (where \underline{x} really means two different things in this sentence), so $\underline{x} - \underline{x}_0 \in U$; this is the same as saying that $\underline{x} \in \underline{x}_0 + U$ or that \underline{x} has the form $\underline{x}_0 + \underline{w}$ for $\underline{w} \in U$.

Chapter 5

5th day, Friday 7/2/04 (Scribe: Shreya Amin)

5.1 Fields

Definition 5.1. A **number field** is a subset $\mathbb{F} \subseteq \mathbb{C}$ such that \mathbb{F} is closed under arithmetic operations.

Example 5.2. $\mathbb{Q}, \mathbb{R}, \mathbb{C}, \mathbb{Q}[\alpha]$ where α is such that $f(\alpha) = 0$ for $f \in \mathbb{Z}[x]$.

Definition 5.3. A **field** is a set \mathbb{F} with 2 operations (addition $+$ and multiplication \bullet), $(\mathbb{F}, +, \bullet)$ such that $(\mathbb{F}, +)$ is an abelian group:

- (a1) $(\forall \alpha, \beta \in \mathbb{F})(\exists! \alpha + \beta \in \mathbb{F})$,
- (a2) $(\forall \alpha, \beta \in \mathbb{F})(\alpha + \beta = \beta + \alpha)$ (commutative law),
- (a3) $(\forall \alpha, \beta, \gamma \in \mathbb{F})(\alpha + (\beta + \gamma) = (\alpha + \beta) + \gamma)$ (associative law),
- (a4) $(\exists 0 \in \mathbb{F})(\forall \alpha)(\alpha + 0 = 0 + \alpha = \alpha)$ (existence of zero),
- (a5) $(\forall \alpha \in \mathbb{F})(\exists(-\alpha) \in \mathbb{F})(\alpha + (-\alpha) = 0)$,

and for (\mathbb{F}, \bullet) have the following. $\mathbb{F}^\times = \mathbb{F} \setminus \{0\}$ is an abelian group with respect to multiplication:

- (b1) $(\forall \alpha, \beta \in \mathbb{F})(\exists! \alpha \bullet \beta \in \mathbb{F})$,
- (b2) $(\forall \alpha, \beta \in \mathbb{F})(\alpha \bullet \beta = \beta \bullet \alpha)$ (commutative law),
- (b3) $(\forall \alpha, \beta, \gamma \in \mathbb{F})(\alpha \bullet (\beta \bullet \gamma) = (\alpha \bullet \beta) \bullet \gamma)$ (associative law),
- (b4) $(\exists 1 \in \mathbb{F})(\forall \alpha)(\alpha \bullet 1 = 1 \bullet \alpha = \alpha)$ (existence of identity),
- (b5) $(\forall \alpha \in \mathbb{F}^\times)(\exists(\alpha^{-1} \in \mathbb{F}^\times)(\alpha \bullet (\alpha^{-1}) = (\alpha^{-1}) \bullet \alpha = 1)$,

(b6) $1 \neq 0$

(b7) $(\forall \alpha, \beta, \gamma \in \mathbb{F})(\alpha \bullet (\beta + \gamma) = \alpha \bullet \beta + \alpha \bullet \gamma)$ (distributive law)

Example 5.4. Examples of fields:

(1) Number fields (Every number field is a field)

(2) $\mathbb{R}(x)$ (Recall $\mathbb{R}[x]$ = polynomials over \mathbb{R} . This is not a field since reciprocal of a polynomial is not necessarily a polynomial. Thus, consider $\mathbb{R}(x) = \{\text{rational functions}\} = \{\frac{f(x)}{g(x)} \mid g(x) \neq 0, f, g \in \mathbb{R}[x]\}$.)

(3) Finite fields: mod p residue classes, p prime. Denote it by $\mathbb{Z}/p\mathbb{Z}$.

Example 5.5. $\mathbb{Z}/p\mathbb{Z}$:

$$\begin{array}{c|cc} + & 0 & 1 \\ \hline 0 & 0 & 1 \\ 1 & 1 & 0 \end{array} \quad \begin{array}{c|cc} \bullet & 0 & 1 \\ \hline 0 & 0 & 0 \\ 1 & 0 & 1 \end{array}$$

$$\begin{array}{c|ccc} + & 0 & 1 & 2 \\ \hline 0 & 0 & 1 & 2 \\ 1 & 1 & 2 & 0 \\ 2 & 2 & 0 & 1 \end{array} \quad \begin{array}{c|ccc} \bullet & 0 & 1 & 2 \\ \hline 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 2 \\ 2 & 0 & 2 & 1 \end{array}$$

$$\begin{array}{c|ccccc} + & 0 & 1 & 2 & 3 & 4 \\ \hline 0 & 0 & 1 & 2 & 3 & 4 \\ 1 & 1 & 2 & 3 & 4 & 0 \\ 2 & 2 & 3 & 4 & 0 & 1 \\ 3 & 3 & 4 & 0 & 1 & 2 \\ 4 & 4 & 0 & 1 & 2 & 3 \end{array} \quad \begin{array}{c|ccccc} \bullet & 0 & 1 & 2 & 3 & 4 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 2 & 3 & 4 \\ 2 & 0 & 2 & 4 & 1 & 3 \\ 3 & 0 & 3 & 1 & 4 & 2 \\ 4 & 0 & 4 & 3 & 2 & 1 \end{array}$$

Part of the group structure of $(\mathbb{Z}/p\mathbb{Z}, +)$ is reflected in the facts that each row and each column is a permutation. Similarly for $((\mathbb{Z}/p\mathbb{Z})^\times, \bullet)$.

Latin Squares: $n \times n$ square filled with n symbols. Every symbol appears in each row and each column exactly once. So, $(\mathbb{Z}/p\mathbb{Z}, +)$ is 5×5 Latin square and $((\mathbb{Z}/p\mathbb{Z})^\times, \bullet)$ is a 4×4 Latin square.

Note: Commutativity \longleftrightarrow symmetric matrix

Definition 5.6. Let $A \in \mathbb{F}^{k \times n}$. **Transpose of A :** flip matrix over main diagonal. Let $A = (\alpha_{ij})_{(i=1)(j=1)}^{(k)(n)}$, then $A^T = (\beta_{ij})_{(i=1)(j=1)}^{(n)(k)}$. The matrix A is **symmetric** if $A = A^T$.

Axiom (d)

$$(\forall \alpha, \beta \in \mathbb{F})(\alpha\beta = 0 \Leftrightarrow \alpha = 0 \text{ or } \beta = 0)$$

Exercise 5.7. Prove that Axiom (d) holds in every field.

Exercise 5.8. Show that Axiom (d) fails in $\mathbb{Z}/6\mathbb{Z}$. So $\mathbb{Z}/6\mathbb{Z}$ is not a field.

Exercise 5.9. If \mathbb{F} is finite and satisfies all field axioms except possibly (b5), then (b5) \iff (d). Note: (d) does **not** necessarily imply (b5) if \mathbb{F} is infinite: \mathbb{Z} is a counterexample.

Theorem 5.10. $\mathbb{Z}/m\mathbb{Z}$ is a field $\iff m$ is prime.

Proof:

- (1) If m is composite, i.e., $m = ab$ where $a, b > 1$, then $\mathbb{Z}/m\mathbb{Z}$ is not a field: it violates axiom (d) because $ab = 0$.
- (2) $\mathbb{Z}/p\mathbb{Z}$ is finite, thus need to show that it satisfies axiom (d): $ab = 0$ in $\mathbb{Z}/p\mathbb{Z}$. **Prime property:** $p \mid ab \implies p \mid a$ or $p \mid b$.

This property plays a key role in proving the **Fundamental Theorem of Arithmetic:** Every number has a unique prime decomposition.

Exercise 5.11. Use exercises below and the prime property to prove the Fundamental Theorem of Arithmetic.

Definition 5.12 (Greatest Common Divisor). Let $a, b, d \in \mathbb{Z}$. Then $d = \text{g.c.d.}(a, b)$ if

- (i) $d \mid a, d \mid b$ (d is a common divisor),
- (ii) $(\forall e)(e \mid a \text{ and } e \mid b \implies e \mid d)$ (d is a multiple of every common divisor).

(This definition cannot distinguish between numbers and its negative: Example: $\text{g.c.d.}(4, 6) = ?$; 2 is a g.c.d. and so is -2).

Theorem 5.13. $(\forall a, b)(\exists d = \text{g.c.d.}(a, b))$ and d is unique up to factor of ± 1 .

What is the greatest common divisor of $(0, 0)$? Divisors of 0: $(\forall x)(x \mid 0)$, so every integer is a divisor of 0. Every divisor of 0 satisfies (i) above, but only 0 satisfies (ii) so $\text{g.c.d.}(0, 0) = 0$. (With respect divisibility, everything is a divisor of 0; and 1 is the divisor of everything).

Exercise 5.14. $(\forall a, b \in \mathbb{Z})(\exists x, y \in \mathbb{Z})(d = ax + by)$ where $d = \text{g.c.d.}(a, b)$.

Example 5.15. $\text{g.c.d.}(7, 11) = 1$: $7x + 11y = 1$, for $x = -3$ and $y = 2$.

Exercise 5.16. Prove; do not use prime factorization: if $\text{g.c.d.}(a, b) = d$ then $\text{g.c.d.}(ac, bc) = cd$.

5.2 Polynomials, rational functions

Let \mathbb{F} be a field and let $\mathbb{F}[x]$ be the set of polynomials over \mathbb{F} . Let $f, g \in \mathbb{F}[x]$.

Definition 5.17 (Divisibility of polynomials). $f \mid g$ if $\exists h \in \mathbb{F}[x]$ s.t. $g = fh$.

Definition 5.18 (Greatest common divisor of polynomials). Let $d, f, g \in \mathbb{F}[x]$. $d = \text{g.c.d.}(f, g)$ is the **greatest common divisor** of f and g if

- (i) $d \mid f$, $d \mid g$, and
- (ii) d is a multiple of all common divisors: $(\forall e)(e \mid f \text{ and } e \mid g \implies e \mid d)$.

g.c.d. is unique up to non-zero constant factors.

Example 5.19. $\text{g.c.d.}(x^2 - 1, (x - 1)^2) = x - 1$ (there is no distinction between $x - 1$ and $73(x - 1)$).

Definition 5.20. f is irreducible if

- (i) $\deg f \geq 1$
- (ii) $(\forall g, h)$ (if $gh = f$ then $\deg g = 0$ or $\deg h = 0$)

Theorem 5.21 (Division Theorem for Integers). $(\forall a, b)$ (if $|b| \geq 1$, then $(\exists q, r)(a = bq + r$ and $0 \leq r < |b|$).

Theorem 5.22 (Division Theorem for polynomials over a field). $(\forall f, g \in \mathbb{F}[x])$ (if $g \neq 0$, then $\exists q, r \in \mathbb{F}[x]$ s.t. $f = gq + r$ and $\deg(r) < \deg(g)$).

Theorem 5.23. Every polynomial has a unique factorization (up to constant multiples) into irreducible factors.

Example 5.24. $(x^2 - 1) = (x - 1)(x + 1) = (3x - 3)(\frac{1}{3}x + \frac{1}{3})$

Everything done so far in linear algebra is true for all fields (not just number fields). Denote “order of field \mathbb{F} ” by $|\mathbb{F}|$.

Theorem 5.25 (Galois). The orders of finite fields are the prime powers; \forall prime power q , $\exists!$ field \mathbb{F}_q of order q (unique up to isomorphism).

Note: $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$ when p is a prime, but $\mathbb{F}_q \neq \mathbb{Z}/q\mathbb{Z}$ when q is not a prime $q = p^k$, $k \geq 2$.

Definition 5.26. The field \mathbb{F} has **characteristic** m ($m \geq 1$) if $\underbrace{1 + 1 + \dots + 1}_m = m \cdot 1 = 0$ and m is the smallest positive integer with this property. If no such m exists, we say that the field has characteristic 0.

What is the characteristic of $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$? Answer: $\text{char } \mathbb{F}_p = p$. What is the characteristic of \mathbb{C} ? Answer: $\text{char } \mathbb{C} = 0$. The same is true for number fields.

Consider the field of rational functions $\mathbb{F}(x)$ over the field \mathbb{F} .

Exercise 5.27. $\text{char } \mathbb{F}(x) = \text{char } \mathbb{F}$. This gives examples of infinite fields of finite (i. e., non-zero characteristic).

Exercise 5.28. If $\mathbb{F} \subset G$ is a subfield, then $\text{char } \mathbb{F} = \text{char } G$.

Example 5.29. $\mathbb{Z}/5\mathbb{Z} \subset \mathbb{Q}$. Is this a subfield? Answer: No, the operations are different; e. g., $3 + 4 = 7$ in \mathbb{Q} but $3 + 4 = 2$ in $\mathbb{Z}/5\mathbb{Z}$.

Exercise 5.30. $\text{char } \mathbb{F} = 0$ or prime.

Exercise 5.31. If $\text{char } \mathbb{F} = 0$, then \mathbb{Q} is a subfield of \mathbb{F} . If $\text{char } \mathbb{F} = p$, then $\mathbb{Z}/p\mathbb{Z}$ is a subfield of \mathbb{F} .

Theorem 5.32. If \mathbb{F} is a finite field of characteristic p , then $|\mathbb{F}| = p^k$.

$\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$, $\mathbb{F}_p \subset \mathbb{F}$. If we have a field extension \implies bigger field is a vector space over the smaller field. Thus, we can treat \mathbb{F} as a vector space over \mathbb{F}_p . Let We can find a basis of \mathbb{F} over \mathbb{F}_p , i.e., every vector can be written as a linear combination of these basis vectors. Let $\dim_{\mathbb{F}_p} \mathbb{F} = k$. We, therefore, have a bijection: $\mathbb{F} \longleftrightarrow \mathbb{F}_p^k$ so $|\mathbb{F}| = |\mathbb{F}_p^k| = p^k$, where $x \mapsto [x]_B$.

Exercise 5.33. Find an irreducible quadratic polynomial over \mathbb{F}_2 .

Solution: What are quadratic polynomials: $ax^2 + bx + c$, $a, b, c \in \mathbb{F}_2 = \{0, 1\}$. Since we want quadratic polynomials, $a \neq 0$ so $a = 1$. Thus, we have $x^2 + bx + c$ with $b, c \in \{0, 1\}$. Thus, there are 4 quadratic polynomials:

- (1) $x^2 = x \cdot x$ not irreducible,
- (2) $x^2 + x = x(x + 1)$ not irreducible,
- (3) $x^2 + 1 = x^2 + 2x + 1 = (x + 1)^2$ not irreducible,
- (4) $x^2 + x + 1$ irreducible.

Another way to see this: the $\text{deg} = 1$ polynomials are x , $x + 1$, so $x \cdot x$, $x(x + 1)$, and $(x + 1)^2$ are the only reducible polynomials $\implies x^2 + x + 1$ is irreducible).

Theorem 5.34. $(\forall p), (\forall k), (\exists \text{ irreducible polynomials of deg } k \text{ over } \mathbb{F}_p.)$

Corollary 5.35. $(\forall p)(\forall k)(\exists \mathbb{F}_{p^k}).$

Once we have irreducible polynomials of $\text{deg } k$ over \mathbb{F}_p , we can immediately construct \mathbb{F}_{p^k} .

Chapter 6

6th day, Tuesday 7/6/04 (Scribe: Jeff Clouse)

6.1 Inverse matrix, rank

Today we will be discussing inverse matrices and representation of linear maps with respect to different bases. To begin, let A be an $n \times n$ matrix.

Definition 6.1. Write $B = A^{-1}$ if $AB = I_n$.

Exercise 6.2. If $AB = I_n$, then $BA = I_n$. We call such a B an **inverse** for A .

Now suppose that $A = (\underline{a}_1, \dots, \underline{a}_n)$ is a $k \times n$ matrix over the field \mathbb{F} , that is $A \in \mathbb{F}^{k \times n}$. If $\underline{b} \in \mathbb{F}^k$, then $\exists \underline{x} \in \mathbb{F}^n$ such that $A\underline{x} = \underline{b} \iff \underline{b} \in \text{Span}(\underline{a}_1, \dots, \underline{a}_n) = \text{column space of } A$. If R is an $n \times k$ matrix, then $AR = A[\underline{r}_1, \dots, \underline{r}_k] = [A\underline{r}_1, \dots, A\underline{r}_k]$. This leads us to the following definition.

Definition 6.3. R is a **right inverse** of A if $AR = I_k$.

We are naturally led to wonder when a right inverse should exist. Letting $\{e_1, \dots, e_k\}$ be the standard basis for \mathbb{F}^k , we see that R exists $\iff \forall i \exists \underline{r}_i$ such that $A\underline{r}_i = \underline{e}_i$. This condition is equivalent to $\underline{e}_1, \dots, \underline{e}_k \in \text{column space of } A \subseteq \mathbb{F}^k$. This means that the column space of A is equal to \mathbb{F}^k ; in other words, A has rank k . So, a $k \times n$ matrix A has a right inverse $\iff \text{rk } A = k$. We restate our findings as a theorem.

Theorem 6.4. *The following are equivalent for a $k \times n$ matrix A :*

1. A has a right inverse.
2. $\text{rk } A = k$.
3. A has full row rank.
4. A has linearly independent rows.

Remember that if $A = (a_{ij})$, then the transpose of A is the matrix given by $A^T = (a_{ji})$.

Exercise 6.5. We have the formula $(AB)^T = B^T A^T$.

Using the exercise, we see that $AR = I \iff R^T A^T = I$. Thus, A has a right inverse iff A^T has a left inverse. We therefore have a new theorem.

Theorem 6.6. *The following are equivalent for a $k \times n$ matrix A :*

1. A has a left inverse.
2. $\text{rk } A = n$.
3. A has full column rank.
4. A has linearly independent columns.

The set of $n \times n$ matrices over a field \mathbb{F} is very important and has its own notation, $M_n(\mathbb{F})$. In this case, the previous two theorems coincide.

Corollary 6.7. *The following are equivalent for $A \in M_n(\mathbb{F})$:*

1. $\text{rk } A = n$.
2. A has a right inverse.
3. A has a left inverse.
4. A has an inverse.
5. $\det A \neq 0$ (A is **nonsingular**).

Exercise 6.8. Let $\alpha_1, \dots, \alpha_n, \beta_1, \dots, \beta_n \in \mathbb{F}$ be all distinct. Now let $\gamma_{i,j} = 1/(\alpha_i - \beta_j)$. The matrix $H = (\gamma_{i,j})$ has full rank. Matrices such as H are called Cauchy-Hilbert matrices.

The set of matrices considered in the previous corollary pervades the study of linear algebra, so we give it a name.

Definition 6.9. The set of nonsingular $n \times n$ matrices over \mathbb{F} is called the **General Linear Group**, and is denoted by $GL(n, \mathbb{F})$.

To justify the name, notice that $GL(n, \mathbb{F})$ is a group under matrix multiplication. Only two axioms require effort to check. First, see that $(A^{-1})^{-1} = A$, so $GL(n, \mathbb{F})$ is closed under taking inverses. Second, see that if $A, B \in GL(n, \mathbb{F})$, then $(AB)^{-1} = B^{-1}A^{-1}$. Therefore $GL(n, \mathbb{F})$ is closed under the group operation. Associativity and the existence of an identity element are clear, so we see that the general linear group is indeed a group.

Now we can use our understanding of matrix inversion to learn about changes of basis. Let $\varphi : V^n \rightarrow W^k$ be a linear map, and suppose we have two bases for each vector space: $\underline{e}, \underline{e}'$; $\underline{f}, \underline{f}'$. Now consider the basis change transformations

$$\sigma : V \rightarrow V, \quad \sigma(\underline{e}_i) = \underline{e}'_i \quad (6.1)$$

$$\tau : W \rightarrow W, \quad \tau(\underline{f}_i) = \underline{f}'_i \quad (6.2)$$

Define $S := [\sigma]_{\underline{e}} = [[\underline{e}'_1]_{\underline{e}}, \dots, [\underline{e}'_n]_{\underline{e}}]$ and $T := [\tau]_{\underline{f}} = [[\underline{f}'_1]_{\underline{f}}, \dots, [\underline{f}'_k]_{\underline{f}}]$. Similarly, let $S' := [\sigma]_{\underline{e}'}$ and $T' := [\tau]_{\underline{f}'}$. Notice that all four of these matrices are nonsingular because their columns are vector space bases. Now define the matrices $A = [\varphi]_{\underline{e}, \underline{f}}$ and $A' = [\varphi]_{\underline{e}', \underline{f}'}$. Note that if x is a column vector in V , then $[\varphi x]_{\underline{f}} = [\varphi]_{\underline{e}, \underline{f}}[x]_{\underline{e}}$.

Our first goal is to compare $\underline{u} = [x]_{\underline{e}}$ with $\underline{u}' = [x]_{\underline{e}'}$. Let's write $\underline{u} = u_1 \underline{e}_1 + \dots + u_n \underline{e}_n$. Now consider the following simple and surprising calculation:

$$\underline{u}' = \sigma x = \sigma\left(\sum u_i \underline{e}_i\right) = \sum u_i \sigma(\underline{e}_i) = \sum u_i \underline{e}'_i.$$

This tells us that

$$\underline{u} = [x]_{\underline{e}} = [\sigma x]_{\underline{e}'} = [\sigma]_{\underline{e}'}[x]_{\underline{e}} = S'[x]_{\underline{e}} = S'\underline{u}'.$$

So, $\underline{u} = S'\underline{u}'$ and $\underline{u}' = (S')^{-1}\underline{u}$, accomplishing our first goal.

Now we can turn to our second goal, which is to compare A with A' . Define $\underline{v} = A\underline{u}$ and $\underline{v}' = A'\underline{u}'$. Now we can see that

$$T'\underline{v}' = \underline{v} = A\underline{u} = AS'\underline{u}'.$$

In other words,

$$(T')^{-1}AS'\underline{u}' = \underline{v}' = A'\underline{u}'.$$

Therefore, we have the formula

$$A' = (T')^{-1}AS'.$$

We can actually clean this formula up a bit by considering the case where $A = S$ and $A' = S'$. In this case, $\tau = \sigma$, so what above were T and T' are now S and S' . So the formula now reads: $S' = (S')^{-1}SS'$. Multiplying on the right by $(S')^{-1}$ then on the left by S' , we find that $S' = S$. We could do the same thing with T to find that $T' = T$, so our nicer formula has the form:

$$A' = T^{-1}AS.$$

Exercise 6.10. If A is nonsingular, then $\text{rk}(AB) = \text{rk } B$ and $\text{rk}(CA) = \text{rk}(C)$.

Exercise 6.11. $\text{rk}(AB) \leq \max\{\text{rk } A, \text{rk } B\}$.

Exercise 6.12. $\text{rk}(A + B) \leq \text{rk } A + \text{rk } B$.

6.2 Similarity of matrices, characteristic polynomial

Let A be an $n \times n$ matrix representing a linear map $V \rightarrow V$. Such a linear map is called a **linear transformation**. A change of basis matrix $S \in GL(n, \mathbb{F})$ represents a linear transformation. If A and A' represent the same linear transformation with respect to the two bases between which S changes, then we have $A' = S^{-1}AS$. This important concept leads us to a definition.

Definition 6.13. If $A, B \in M_n(\mathbb{F})$, then they are **similar** (or **conjugate**) if $\exists S \in GL(n, \mathbb{F})$ such that $B = S^{-1}AS$. This is denoted by $A \sim B$.

Theorem 6.14. Let V be a vector space and φ a linear transformation. Then for any two bases $(\underline{e}, \underline{e}')$, $[\varphi]_{\underline{e}} \sim [\varphi]_{\underline{e}'}$.

Exercise 6.15. Similarity of matrices is an equivalence relation.

Recall the determinant function $\det : M_n(\mathbb{F}) \rightarrow \mathbb{F}$.

Exercise 6.16. $\det(AB) = \det A \det B$

We have a neat formula for the determinant of an inverse matrix. Consider

$$AA^{-1} = I \Rightarrow \det(AA^{-1}) = \det I = 1.$$

Then, $\det(AA^{-1}) = \det A \det A^{-1} \Rightarrow \det A^{-1} = 1/\det A$.

Exercise 6.17. If $A \sim B$, then $\det A = \det B$.

Now recall that for an $n \times n$ matrix A , the trace of A is given by the formula $\text{tr } A = \sum_{i=1}^n a_{ii}$.

Exercise 6.18. For $A \in \mathbb{F}^{k \times n}$ and $B \in \mathbb{F}^{n \times k}$, we have $\text{tr}(AB) = \text{tr}(BA)$.

Now let $A \in M_n(\mathbb{F})$ and x be a variable in \mathbb{F} .

Definition 6.19. The **characteristic matrix** of A is the matrix $xI - A$. The **characteristic polynomial** of A is the polynomial $f_A(x) := \det(xI - A)$.

Example 6.20. Let $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$, so $\det A = -2$ and $\text{tr } A = 5$. Then the characteristic matrix of A is $xI - A = \begin{pmatrix} x-1 & -2 \\ -3 & x-4 \end{pmatrix}$. Then the characteristic polynomial of A is $f_A(x) = \begin{vmatrix} x-1 & -2 \\ -3 & x-4 \end{vmatrix} = (x-1)(x-4) - 6 = x^2 - 5x - 2 = x^2 - \text{tr } A + \det A$.

Exercise 6.21. The characteristic polynomial of A is actually given by

$$f_A(x) = x^n - \text{tr } Ax^{n-1} + \cdots + (-1)^n \det A.$$

Exercise 6.22. If $A \sim B$, then $f_A(x) = f_B(x)$.

Since matrices which represent the same linear map with respect to different bases are similar, we can make the following definition.

Definition 6.23. Let $\varphi : V \rightarrow V$ be linear. The **characteristic polynomial** of φ is given by $f_\varphi(x) := f_A(x)$, where $A = [\varphi]$ in some basis.

Finally, let $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$.

Exercise 6.24. Calculate $f_A(A) = A^2 - (a + d)A + (ad - bc)I$ to find a curious result.

Chapter 7

6th day, Wednesday 7/7/04 (Scribe: Ben Lee)

7.1 Review

7.1.1 Matrices

Let $\varphi : V \rightarrow W$ be a linear map of vector spaces over a field F . Choose a basis $\underline{d} = (d_1, \dots, d_n)$ of V and a basis $\varphi = (\varphi_1, \dots, \varphi_k)$ of W . We can form the matrix

$$[\varphi]_{\underline{d}, \varphi} = [[\varphi d_1]_{\varphi} \quad [\varphi d_2]_{\varphi} \quad \cdots \quad [\varphi d_n]_{\varphi}]$$

(the columns of the matrix are the images of the elements d_i under the map φ , expressed in terms of the basis φ .)

This is of course only well defined up to choice of bases.

7.1.2 Change of basis

Let $\underline{d}, \underline{d}'$ and φ, φ' be bases for V and W , respectively. Set

$$A = [\varphi]_{\underline{d}, \varphi}, A' = [\varphi]_{\underline{d}', \varphi'}.$$

Then

$$A' = T^{-1}AS$$

where

$$T = [[\varphi'_1]_{\varphi} \quad \cdots \quad [\varphi'_k]_{\varphi}], S = [[d'_1]_{\underline{d}} \quad \cdots \quad [d'_n]_{\underline{d}}].$$

(expressing the new bases φ', \underline{d}' in terms of the old bases φ, \underline{d} .) These are called “change of basis” matrices. Note these are automatically invertible¹: we can always express φ in terms of φ' , etc.

¹e.g. non-singular.

7.2 Characteristic Polynomials

7.2.1 Similar matrices

Most important case: $V = W$; $\varphi : V \rightarrow V$ a “linear transformation.” For a two bases $\underline{d}, \underline{d}'$ we get

$$A' = S^{-1}AS.$$

Definition 7.1. Two square matrices A, B in $M_n(F)$ ($= n \times n$ matrices with coefficients in F) are called **similar** if there is an invertible matrix S so that $B = S^{-1}AS$. We write $A \sim B$ in this case.

Recall

Theorem 7.2. For an $n \times n$ matrix A , the following are equivalent:

1. $\text{rk}(A) = n$;
2. A has an inverse;
3. $\det(A) \neq 0$.

Theorem 7.3. $\det(AB) = \det(A) \det(B)$.

Consequence: $A \sim B \Rightarrow \det(A) = \det(B)$ because

$$\det(S^{-1}) = \frac{1}{\det(S)}$$

.

7.2.2 Characteristic Polynomial

Definition 7.4. The characteristic polynomial of a matrix A is

$$f_A(x) = \det(xI - A).$$

Example 7.5. $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. The “characteristic matrix” is $xI - A = \begin{bmatrix} x - a & b \\ -cx & -d \end{bmatrix}$.

$$f_A(x) = \begin{vmatrix} x - a & b \\ -cx & -d \end{vmatrix} = x^2 - (a + d)x + (ad - bc) = x^2 - \text{tr}(A) + \det(A).$$

You can apply matrices to polynomials:

$$g(x) = 3x^7 - 5x^2 + x + 8$$

$$g(A) := 3A^7 - 5A^2 + A + 8I.$$

Note for our 2×2 matrix A

$$f_A(A) = A^2 - (a+d) \begin{bmatrix} a & b \\ c & d \end{bmatrix} + (ad - bc)I = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

On this scanty evidence (and limited authority) we assert:

Theorem 7.6 (Caley-Hamilton). $f_A(A) = 0$ for any matrix A .

Note

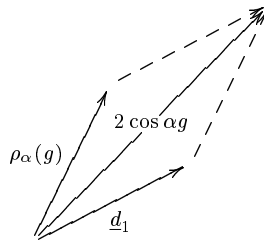
Theorem 7.7. If $A \sim B$ then $f_A(x) = f_B(x)$.

Example 7.8. ρ_α = counter-clockwise rotation in the plane around the origin by angle α . Choosing the standard basis $\underline{d} = (\underline{d}_1, \underline{d}_2)$ we get

$$A = [\rho_\alpha]_{\underline{d}} = \begin{bmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{bmatrix}.$$

e.g. $\rho_\alpha(\underline{d}_1) = \cos \alpha \underline{d}_1 + \sin \alpha \underline{d}_2$.

Another basis: $\underline{d}_1, \underline{g}$ where \underline{g} is the unit vector angle α counter-clockwise from \underline{d}_1 . Clearly $\rho_\alpha(\underline{d}_1) = \underline{g}$. To compute $\rho_\alpha(\underline{g})$ draw the rhombus with edges $\underline{d}_1, v = \rho_\alpha(\underline{g})$ and long diagonal $\underline{d}_1 + v = 2 \cos \alpha \underline{g}$.



Under this basis φ we have

$$A' = [\rho_\alpha]_{\varphi} = \begin{bmatrix} 0 & -1 \\ 1 & 2 \cos \alpha \end{bmatrix}$$

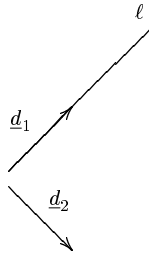
Note that $\det(A) = \det(A') = \cos^2 \alpha + \sin^2 \alpha = 1$ and $\text{tr}(A) = \text{tr}(A') = 2 \cos \alpha$. The characteristic polynomials are

$$f_A(x) = f_{A'}(x) = x^2 - 2 \cos \alpha x + 1.$$

Consequence (of Theorem 7.7): characteristic polynomials are invariant under change of basis, so we can define the characteristic polynomial of a linear transformation (not just a matrix.)

Definition 7.9. For a linear transformation $\varphi : V \rightarrow V$, define its characteristic polynomial $f_\varphi(x)$ as follows: choose any basis \underline{d} of V so that we can write the matrix of φ in terms of \underline{d} as A , and let $f_\varphi(x)$ be the characteristic polynomial of A . This doesn't depend on the choice of \underline{d} since another choice will give a similar matrix.

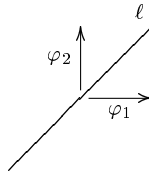
Example 7.10. Reflection: τ_ℓ through a line ℓ . One convenient basis: \underline{d}_1 is some vector on ℓ , and \underline{d}_2 is a vector perpendicular to ℓ .



Then $\tau_\ell(\underline{d}_1) = \underline{d}_1$, $\tau_\ell(\underline{d}_2) = -\underline{d}_2$, and we have

$$f_{\tau_\ell}(x) = \begin{vmatrix} x-1 & 0 \\ 0 & x+1 \end{vmatrix} = x^2 - 1.$$

Another convenient basis: φ_1, φ_2 are equal angles from ℓ :



We have $\tau_\ell(\varphi_1) = \varphi_2$, $\tau_\ell(\varphi_2) = \varphi_1$ so the matrix is

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

and $\text{tr}(A) = 0$, $\det(A) = -1$.

Finally, for the basis \underline{g}_1 is on ℓ , \underline{g}_2 is α degrees off ℓ , we have (as in the rhombus computation) $\tau_\ell(\underline{g}_1) = \underline{g}_1$, $\tau_\ell(\underline{g}_2) = 2 \cos \alpha \underline{g}_1 - \underline{g}_2$. Thus

$$A = \begin{bmatrix} 1 & 2 \cos \alpha \\ 0 & -1 \end{bmatrix}$$

and $\text{tr}(A) = 0$, $\det(A) = -1$.

Proof: [Proof that $A \sim B \Rightarrow f_A(x) = f_B(x)$] There is an S, S^{-1} so that $B = S^{-1}AS$. Thus

$$S^{-1}(xI - A)S = S^{-1}xS - S^{-1}AS = xI - B$$

and so

$$xI - B \sim xI - A$$

therefore

$$f_B = \det(xI - B) = \det(xI - A) = f_A.$$

Exercise 7.11. 3×3 : Let

$$A = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{bmatrix}.$$

Show $f_A(x) = x^3 - \text{tr}(A)x^2 + a_1x - \det(A)$ where

$$a_1 = \begin{vmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{vmatrix} + \begin{vmatrix} \alpha_{11} & \alpha_{13} \\ \alpha_{31} & \alpha_{33} \end{vmatrix} + \begin{vmatrix} \alpha_{22} & \alpha_{23} \\ \alpha_{32} & \alpha_{33} \end{vmatrix}.$$

7.3 The Determinant

7.3.1 Upper Triangular Matrices

Definition 7.12. An upper triangular matrix looks like

$$\begin{bmatrix} * & \cdots & * \\ & \ddots & \vdots \\ 0 & & * \end{bmatrix}.$$

Similarly for lower triangular matrices. A diagonal matrix looks like

$$D = \text{diag}(\lambda_1, \dots, \lambda_n) = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}.$$

Note that $f_D(x) = \prod(x - \lambda_i)$. This is also true for upper triangular matrices, but we'll first need to define the determinant.

7.3.2 Determinant

The determinant of a square matrix

$$A = \begin{bmatrix} \alpha_{11} & & \alpha_{1n} \\ & \ddots & \\ \alpha_{n1} & & \alpha_{nn} \end{bmatrix}$$

is a sum of $n!$ terms, each term a product of n entries from the matrix, one from each row and column.

“Rook arrangements”: put n rooks on an $n \times n$ chessboard in a non-interfering way, e.g. no two on the same row or column. This is a decision tree with $n!$ leaves.

$$\text{Permanent} = \sum_{\text{rook arrangements}} \prod \text{terms in the rook arrangement.}$$

Determinant is the permanent with signs.

Exercise 7.13. Using this (partial) definition, show upper triangular matrices have

$$\det(A) = \prod \text{diagonal entries.}$$

How to determine the signs of the terms? Accept that the sign of the diagonal is $+1$.

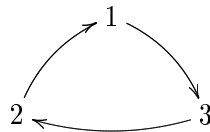
Remark. The Permanent is difficult to compute: equivalent to not just computing 3-coloring of a graph, but finding all 3-colorings!

7.3.3 Permutations and Parity

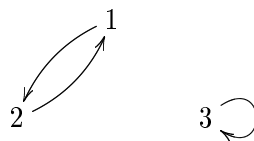
Definition 7.14. A permutation of a set S is a bijection $S \rightarrow S$ (1-to-1 correspondence.) Sometimes we write it in a table

$$\begin{array}{c|ccc} \mathbf{x} & 1 & 2 & 3 \\ \hline \mathbf{f(x)} & 3 & 1 & 2 \end{array}$$

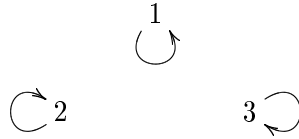
or “musical chairs”



Another example:



Permutations correspond to rook arrangements: put the rook on the i th row on the j th column, where i goes to j under the permutation (and vice versa). The identity permutation



corresponds to the diagonal rook arrangement.

S_n = all permutations of the set $\{1, \dots, n\}$ = symmetric group of degree n .

$$\text{Permanent} = \sum_{f \in S_n} \prod_{i=1}^n \alpha_{i, f(i)}.$$

e.g. $\alpha_{12}\alpha_{21}\alpha_{33} = \prod_{i=1}^3 \alpha_{i, f(i)}$ where

$$\begin{array}{c|ccc} \mathbf{x} & 1 & 2 & 3 \\ \hline \mathbf{f(x)} & 2 & 1 & 3 \end{array}.$$

$$\det(A) = \sum_{f \in S_n} \text{sign}(f) \prod_{i=1}^n \alpha_{i, f(i)}.$$

Certain permutations are even: positive sign. Others are odd: negative sign. Odd or even is parity of the number of transpositions of columns to get to the identity.

Theorem 7.15. *Parity is invariant. (Any two ways of getting to the identity by transpositions have the same parity.)*

Corollary 7.16. *Sign of a permutation is well-defined.*

Theorem is same as

Theorem 7.17. *A product of an odd number of transpositions is never the identity.*

because if you had 2 sets of transpositions which took you to a permutation and back to the identity, with different parities, their composition would be an odd parity set of transpositions giving the identity. “If we keep switching pairs of people an odd number of times, it is impossible to get everyone back to their original positions.”

Row parity is the same as column parity: consider the following permutation

$$f = \begin{bmatrix} \cdot & & & & \\ \cdot & \cdot & & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & & \cdot \end{bmatrix} = \frac{\begin{array}{c|ccccc} \mathbf{x} & 1 & 2 & 3 & 4 & 5 \\ \hline \mathbf{f(x)} & 2 & 3 & 1 & 5 & 4 \end{array}}{=} = \begin{array}{c} 1 \\ \curvearrowright \\ 2 \\ \curvearrowleft \\ 3 \end{array} \quad \begin{array}{c} 5 \\ \curvearrowright \\ 4 \\ \curvearrowleft \\ 5 \end{array}$$

Exercise 7.18. Odd cycle is an even permutation; even cycle is an odd permutation.

For example, to get the 3-cycle (231):

$$\begin{array}{ccc} 1 & \xleftrightarrow{2} & 3 \\ & & 3 \xleftrightarrow{2} 1 \\ & & 2 \quad 3 \quad 1 \end{array}$$

Now take the “transpose” of f

$$g = \begin{bmatrix} \cdot & & & & \\ \cdot & \cdot & & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & & \cdot \end{bmatrix} = \frac{\begin{array}{c|ccccc} \mathbf{x} & 1 & 2 & 3 & 4 & 5 \\ \hline \mathbf{g(x)} & 3 & 1 & 2 & 5 & 4 \end{array}}{=} = \begin{array}{c} 1 \\ \curvearrowright \\ 2 \\ \curvearrowleft \\ 3 \end{array} \quad \begin{array}{c} 5 \\ \curvearrowright \\ 4 \\ \curvearrowleft \\ 5 \end{array}$$

Note $g f x = (12345) =$ the identity. So $g = f^{-1}$. If

$$p = t_1 \cdots t_7$$

is a product of transpositions, then

$$p^{-1} = t_7 \cdots t_1.$$

Remark. Transposing 2 columns of a matrix changes the sign of the determinant. Therefore, if two columns are equal, the determinant is zero.

Remark. If you have a “block upper triangular matrix” (where A, B, D are each submatrices) then

$$\det \left[\begin{array}{c|c} A & B \\ \hline 0 & D \end{array} \right] = \det(A) \det(D).$$

Therefore $f \left(\begin{array}{c|c} A & B \\ \hline 0 & D \end{array} \right) = f_A \cdot f_D$.

7.4 Eigenvectors and Eigenvalues

7.4.1 Diagonalizability

Definition 7.19. A is diagonalizable if A is similar to a diagonal matrix.

Example 7.20.

$$\begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}$$

Exercise 7.21. $\text{diag}(\lambda_1, \dots, \lambda_n) \sim \text{diag}(\lambda_{f(1)}, \dots, \lambda_{f(n)})$ for all $f \in S_n$. (Characteristic polynomials are the same.)

7.4.2 Eigenvalues

Definition 7.22. \underline{x} is an **eigenvector** of $A \in M_n(F)$ if $\underline{x} \neq 0$ and there is a λ so that

$$A\underline{x} = \lambda\underline{x}.$$

Similar for an eigenvector of a linear transformation.

Example 7.23. Reflection has eigenvectors with $\lambda = 1$ (on the line) and -1 (perpendicular.)

Definition 7.24. λ is an **eigenvalue** of A if there is an eigenvector \underline{d} so that

$$A\underline{d} = \lambda\underline{d}.$$

If λ is an eigenvalue, we can solve the equations

$$\begin{aligned} \lambda\underline{x} - A\underline{x} &= 0 \\ \lambda I\underline{x} - A\underline{x} &= 0 \\ (\lambda I - A)\underline{x} &= 0 \end{aligned}$$

which is a system of linear equations. We want a non-trivial solution (a nonzero eigenvector), so λ is an eigenvalue if and only if

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

In other words, λ is a root of the characteristic polynomial

$$f_A(x) = \det(xI - A).$$

We say λ is a characteristic root.

Theorem 7.25 (Spectral Theorem, Part 1). *If A is a real ($F = \mathbb{R}$) symmetric matrix ($\alpha_{ij} = \alpha_{ji}$, or $A^T = A$) then A is diagonalizable.*

If $A = \text{diag}(\lambda_i)$ then $f_A(x) = \prod(x - \lambda_i)$, the eigenvalues are λ_i . Eigenvectors of a diagonal matrix are the standard basis.

Chapter 8

8th day, Thursday 7/8/04 (Scribe: D. Jeremy Copeland)

8.1 Cauchy-Hilbert matrix

Recall

$$\mathbb{F}[x] = \{\text{polynomials over a field, } \mathbb{F}\},$$
$$\mathbb{F}(x) = \{\text{rational functions over } \mathbb{F}\} = \left\{ \frac{f}{g} \mid f, g \in \mathbb{F}[x] \text{ and } g \neq 0 \right\}.$$

$\mathbb{F}(x)$ is a field as well as a vector space over \mathbb{F} .

Theorem 8.1. *If $\alpha_j \in \mathbb{F}$ are all distinct, then the functions $\frac{1}{x-\alpha_1}, \dots, \frac{1}{x-\alpha_n}$ are all linearly independent.*

Theorem 8.2. *Let $f(x) = a_0 + a_1x + \dots + a_nx^n \in \mathbb{F}[x]$. If f has more than n roots, then $f(x) \equiv 0$.*

Lemma 8.3. *If $f \in \mathbb{F}$, there is some polynomial $g \in \mathbb{F}$ such that $f(x) = (x - \alpha)g(x) + f(\alpha)$.*

Lemma 8.4. $(x - \alpha) \mid f(x) - f(\alpha)$.

Corollary 8.5. *If $f(\alpha) = 0$, then $(x - \alpha) \mid f(x)$.*

Let $\alpha_1, \dots, \alpha_n, \beta_1, \dots, \beta_n$ be distinct elements of \mathbb{F} . Recall that the Cauchy-Hilbert matrix is defined as

$$H = \begin{bmatrix} \frac{1}{\alpha_1 - \beta_1} & \cdots & \frac{1}{\alpha_1 - \beta_n} \\ \vdots & & \vdots \\ \frac{1}{\alpha_n - \beta_1} & \cdots & \frac{1}{\alpha_n - \beta_n} \end{bmatrix}.$$

Theorem 8.6. *The Cauchy-Hilbert matrix has full rank.*

Remark 8.7. The Cauchy-Hilbert matrix is not defined if $\alpha_i = \beta_j$ for some i, j and is not of full rank if $\alpha_i = \alpha_j$ or $\beta_i = \beta_j$ for some $i \neq j$.

8.2 Eigenvectors, eigenvalues

Recall that we have the standard inner product on \mathbb{F}^n : if $\underline{a} = (\alpha_1, \dots, \alpha_n)$ and $\underline{b} = (\beta_1, \dots, \beta_n)$, then

$$\underline{a} \cdot \underline{b} = \sum_{i=1}^n \alpha_i \beta_i.$$

Definition 8.8. \underline{a} and \underline{b} are **perpendicular**, $\underline{a} \perp \underline{b}$, if $\underline{a} \cdot \underline{b} = 0$.

Definition 8.9. For $S \subset \mathbb{F}^n$, $S^\perp := \{\underline{x} \in \mathbb{F}^n \mid (\forall \underline{s} \in S) \underline{x} \perp \underline{s}\}$.

Exercise 8.10. $S^\perp \leq \mathbb{F}^n$ (S^\perp is a subspace of \mathbb{F}^n).

Exercise 8.11. $S \subset S^{\perp\perp}$.

Corollary 8.12. $\text{Span}S \subset S^{\perp\perp}$.

Theorem 8.13. If $U \subset \mathbb{F}^n$, then $\dim U + \dim U^\perp = n$.

Definition 8.14. $\underline{x} \in \mathbb{F}^n$ is an **eigenvector** of $A \in M_n(\mathbb{F})$ if $\exists \lambda \in \mathbb{F}$ such that $A\underline{x} = \lambda\underline{x}$ and $\underline{x} \neq 0$.

Definition 8.15. $\lambda \in \mathbb{F}$ is an **eigenvalue** of $A \in M_n(\mathbb{F})$ if $\exists \underline{x} \in \mathbb{F}^n$ such that $A\underline{x} = \lambda\underline{x}$ and $\underline{x} \neq 0$.

Theorem 8.16. λ is an eigenvalue of $A \iff \lambda$ is a root of the characteristic polynomial of A , i. e., $f_A(\lambda) = 0$, where $f_A(x) = \det(xI - A)$.

Let ρ_α be rotation by α in the plane \mathbb{R}^2 . In the standard basis,

$$A = [\rho_\alpha] = \begin{bmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{bmatrix}.$$

It has the characteristic polynomial $f_{[\rho_\alpha]}(x) = x^2 - 2\cos(\alpha)x + 1$. Its eigenvalues are:

$$\lambda_1 = \cos(\alpha) + i\sin(\alpha) \quad \lambda_2 = \cos(\alpha) - i\sin(\alpha)$$

The eigenvectors are:

$$\underline{f}_1 = \begin{bmatrix} 1 \\ -i \end{bmatrix} \quad \underline{f}_2 = \begin{bmatrix} 1 \\ i \end{bmatrix}$$

Remark 8.17. Given that $A\underline{f}_1 = \lambda_1\underline{f}_1$, we may deduce the other eigenvector by conjugating:

$$\begin{aligned} \overline{A\underline{f}_1} &= \overline{\lambda_1\underline{f}_1} \\ A\underline{f}_1 &= \overline{\lambda_1} \overline{\underline{f}_1} \\ A\underline{f}_1 &= \lambda_2 \overline{\underline{f}_1} \end{aligned}$$

(8.1)

Remark 8.18. All rotation matrices (other than $\pm I$) have the same eigenvectors, \underline{f}_1 and \underline{f}_2 .

Remark 8.19. \underline{f}_1 and \underline{f}_2 are linearly independent since

$$\begin{vmatrix} 1 & 1 \\ -i & i \end{vmatrix} = 2i \neq 0,$$

thus $B = \{\underline{f}_1, \underline{f}_2\}$ is a basis for \mathbb{C}^2 .

The basis change matrix from the standard basis to B is:

$$S = \begin{bmatrix} 1 & 1 \\ -i & i \end{bmatrix}.$$

In fact, the basis change matrix from the standard basis to another basis B' is always the matrix whose columns are the vectors in B' . Also,

$$S^{-1} = \frac{1}{2} \begin{bmatrix} 1 & -i \\ 1 & i \end{bmatrix}.$$

What is $[\rho_\alpha]$ in our new basis, B ?

$$\begin{aligned} [\rho_\alpha]_B &= S^{-1}[\rho_\alpha]_{\text{std}}S \\ &= \frac{1}{2} \begin{bmatrix} 1 & -i \\ 1 & i \end{bmatrix} \begin{bmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -i & i \end{bmatrix} \\ &= \begin{bmatrix} \lambda_2 & 0 \\ 0 & \lambda_1 \end{bmatrix}. \end{aligned} \tag{8.2}$$

It becomes the diagonal matrix with diagonal being the eigenvalues. Let $A' = [\rho_\alpha]_B$. The defining property of \underline{f}_j is that $\rho_\alpha \underline{f}_j = \lambda_j \underline{f}_j$. But A' is ρ_α in the basis $B = \{\underline{f}_2, \underline{f}_1\}$, so we should expect that

$$A' = \begin{bmatrix} \lambda_2 & 0 \\ 0 & \lambda_1 \end{bmatrix}.$$

Theorem 8.20. *In an ordered basis of eigenvectors, the matrix for a linear transformation is diagonal with diagonal entries the eigenvalues with the respective order.*

Definition 8.21. If $\varphi : V \rightarrow V$ is a linear transformation, then $\underline{x} \in V$ is an **eigenvector** if $\underline{x} \neq 0$ and $\exists \lambda$ such that $\varphi \underline{x} = \lambda \underline{x}$.

Definition 8.22. If $\varphi : V \rightarrow V$ is a linear transformation, then $\lambda \in \mathbb{F}$ is an **eigenvalue** if $\exists \underline{x} \in V$ such that $\underline{x} \neq 0$ and $\varphi \underline{x} = \lambda \underline{x}$.

Remark 8.23. If $\varphi \underline{x} = \lambda \underline{x}$ then $[\varphi][\underline{x}] = \lambda[\underline{x}]$ in any basis.

Remark 8.24. λ is an eigenvalue for $\varphi \iff \lambda$ is an eigenvalue for $[\varphi]$ in any basis $\iff f_{[\varphi]}(\lambda) = 0$ in any basis.

Recall that similar matrices have the same characteristic polynomials, and that the matrices for φ in any two bases are similar.

Definition 8.25. The **characteristic polynomial** for φ is $f_\varphi = f_{[\varphi]}$ in some (thus any) basis.

Theorem 8.26. $[\varphi]_B$ is diagonal $\iff B$ is a basis of eigenvectors.

Definition 8.27. An **eigenbasis** for A is a basis consisting of eigenvectors.

Definition 8.28. A matrix, A , is **diagonalizable** if and only if there exists an eigenbasis for A .

Remark 8.29. As we saw before, S is the expression of the eigenbasis of A in terms of the original basis.

Exercise 8.30. The matrix

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

has no eigenbasis.

Exercise 8.31. If $\{\underline{e}_1, \dots, \underline{e}_k\}$ are eigenvectors associated to distinct eigenvalues, then they are linearly independent.

Exercise 8.32. If an $n \times n$ matrix has n distinct eigenvalues, then it is diagonalizable.

Corollary 8.33.

$$\begin{bmatrix} 1 & 71 & 522 \\ 0 & 2 & \sqrt{\pi^2 + e} \\ 0 & 0 & 3 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

This follows because the matrix on the left hand side (call it B) has characteristic polynomial

$$f(x) = (x - 1)(x - 2)(x - 3),$$

therefore it has 3 distinct eigenvalues $\{1, 2, 3\}$. By Exercise 8.32 it is diagonalizable. So some diagonal matrix A is similar to B ; but A and B must have the same eigenvalues, so A must be the matrix on the right hand side.

Chapter 9

9th day, Friday 7/9/04 (Scribe: Sanjeevi Krishnan)

9.1 Spectral Theorem

Definition 9.1. The **Vandermode Matrix** $V_n(x_1, \dots, x_n)$ is defined to be the matrix

$$\begin{pmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_n \\ x_1^2 & x_2^2 & \dots & x_n^2 \\ \vdots & \vdots & \dots & \vdots \\ x_1^n & x_2^n & \dots & x_n^n \end{pmatrix}$$

The terms x_1, \dots, x_n are called the generators of $V_n(x_1, \dots, x_n)$.

Exercise 9.2. If x_1, \dots, x_n are distinct, then V_n has full rank.

Exercise 9.3. $\det(V_n(x_1, \dots, x_n)) = \prod_{1 \leq j < i \leq n} (x_i - x_j)$

Theorem 9.4. *The identity permutation cannot be represented as the product of an odd number of transpositions.*

Definition 9.5. A matrix A (with coefficients in \mathbb{R}) is a **real symmetric matrix** if $A = A^T$, i.e. if $A = (\alpha_{ij})$ then $\alpha_{ij} = \alpha_{ji}$.

Remark 9.6. Over \mathbb{C} , every matrix has an eigenvalue (by the Fundamental Theorem of Algebra.)

Theorem 9.7. *All eigenvalues of a real symmetric matrix are real.*

Proof: Suppose $Ax = \lambda x$, $x \neq 0$. We want to show that $\lambda = \bar{\lambda}$. If we let x^* denote \bar{x}^T , then note:

$$x^*Ax = x^*(\lambda x) = \lambda x^*x = \lambda(\bar{x}_1x_1 + \dots + \bar{x}_nx_n) = \lambda \left(\sum_{i=1}^n |x_i|^2 \right)$$

and so

$$\lambda \left(\sum_{i=1}^n |x_i|^2 \right) = x^* Ax = x^* Ax^{**} = (x^* Ax)^* = \left(\lambda \left(\sum_{i=1}^n |x_i|^2 \right) \right)^* = \bar{\lambda} \left(\sum_{i=1}^n |x_i|^2 \right)$$

which implies $\lambda = \bar{\lambda}$.

Example 9.8. Let

$$A = \begin{pmatrix} a & b \\ b & d \end{pmatrix}$$

Then A is a symmetric matrix with characteristic polynomial $f_A(x) = x^2 - (\operatorname{tr}(A))x + \det(A) = x^2 - (a+d)x + (ad - b^2)$, which has real roots iff the discriminant is nonnegative. This is true because the discriminant is:

$$(a+d)^2 - 4(ad - b^2) = a^2 - 2ad + d^2 + 4b^2 = (a-d)^2 + 4b^2 \geq 0$$

Recall the following fact:

Theorem 9.9. A matrix A with coefficients in a field \mathbb{F} is diagonalizable iff A has an eigenbasis, i.e. \mathbb{F}^n has a basis consisting of eigenvectors of A .

Definition 9.10. The **inner product** (on \mathbb{R}^n) is a function $\cdot : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ defined as

$$\underline{a} \cdot \underline{b} = \sum_{i=1}^n \alpha_i \beta_i = \underline{a}^T \underline{b}$$

where

$$\underline{a} = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} \quad \underline{b} = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix}$$

Definition 9.11. The **(Euclidean) norm of a vector \underline{a}** , written $\|\underline{a}\|$, is defined to be $\sqrt{\underline{a} \cdot \underline{a}}$.

Definition 9.12. The set of vectors $\{\underline{a}_1, \underline{a}_2, \dots, \underline{a}_k\}$ is **orthonormal (ON)** iff

1. $\underline{a}_i \perp \underline{a}_j$, i.e. $\underline{a}_i^T \underline{a}_j = 0$, for all $i \neq j$
2. $\|\underline{a}_i\| = 1$, i.e. $\underline{a}_i^T \underline{a}_i = 1$, for all i .

Exercise 9.13. If $\underline{a}_1, \dots, \underline{a}_k$ are orthonormal, then they are linearly independent.

Definition 9.14. An **orthonormal basis (ONB)** is an orthonormal set of n vectors $\underline{a}_1, \dots, \underline{a}_n \in \mathbb{R}^n$ (which thus forms a basis).

Theorem 9.15 (Spectral Theorem). *If A is a symmetric real matrix, then A has an orthonormal eigenbasis.*

Observation 9.16. The standard basis on \mathbb{R}^n is orthonormal. If we want to change to another ONB $\{e_1, \dots, e_n\}$, then our base-change matrix would be:

$$S = (e_1 \quad \dots \quad e_n)$$

so that

$$S^T = \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix}$$

and so $S^T S = I$ because $e_i^T e_j = 0$ for $i \neq j$, and $e_i^T e_i = 1$ for all i (by orthonormality.)

Observation 9.17. The following are equivalent for an $n \times n$ matrix S .

1. The columns of S are orthonormal.
2. $S^T S = I$

Recall the following:

Theorem 9.18 (Amazing Fact 1). $\text{rk}(S) = \text{rk}(S^T)$.

Our last observation leads us to the following:

Theorem 9.19 (Amazing Fact 2). *If the columns of a matrix S are orthonormal, then so are the rows.*

Proof: The columns of S are ON $\Rightarrow S^T S = I \Rightarrow S^T = S^{-1} \Rightarrow S S^T = I \Rightarrow (S^T)^T S^T = I \Rightarrow$ the columns of S^T are ON \Rightarrow rows of S are ON.

Definition 9.20. A real $n \times n$ matrix is an **orthogonal matrix** iff its transpose is its inverse.

Theorem 9.21 (Spectral Theorem Restated). *If A is a real symmetric matrix, then there is an orthogonal matrix S such that $S^{-1} A S = \text{diag}(\lambda_1, \dots, \lambda_n)$, where λ_i are real numbers (and thus the eigenvalues of A .)*

Definition 9.22. Two matrices A, B are **similar** (written $A \sim B$) iff there is a matrix S such that $B = S^{-1} A S$.

Observation 9.23. If A is a symmetric real matrix, then A has a real eigenvector \underline{f}_1 , where w.l.o.g. we can take $\|\underline{f}_1\| = 1$.

Definition 9.24. Suppose $U \leq V$, i.e. U is a subspace of V , and $\varphi : V \rightarrow V$ is a linear map. We say that U is an **invariant subspace** for φ iff $\varphi(U) \subset U$, i.e. for all $u \in U$, $\varphi(u) \in U$.

Theorem 9.25. *Let A be a nonsingular matrix, and suppose U is a right-invariant subspace of A , i.e. for all $u \in U$, $Au \in U$. Then U^\perp is a left-invariant subspace of A (i.e. U^\perp is a right-invariant subspace of A^T).*

Corollary 9.26. *Suppose A is a symmetric real matrix. If U is an invariant subspace for the linear map determined by A (i.e. U is a left-invariant subspace of the matrix A), then so is U^\perp .*

Chapter 10

10th day, Monday 7/12/04 (Scribe: Daniel Štefankovič)

10.1 Vandermonde Matrix

Recall that $V_n(x_1, \dots, x_n)$ is the Vandermonde matrix

$$V_n(x_1, \dots, x_n) = \begin{pmatrix} 1 & 1 & \dots & \dots & 1 \\ x_1 & x_2 & \dots & \dots & x_n \\ x_1^2 & x_2^2 & \dots & \dots & x_n^2 \\ \vdots & \vdots & & & \vdots \\ x_1^{n-1} & x_2^{n-1} & \dots & \dots & x_n^{n-1} \end{pmatrix}.$$

Exercise 10.1. Show, using Gaussian elimination, that $\det V_n(x_1, \dots, x_n) = \prod_{i < j} (x_i - x_j)$.

Exercise 10.2. Let $\varphi : V \rightarrow V$ be a linear transformation. Let v_1, \dots, v_k be eigenvectors, $\varphi v_i = \lambda_i v_i$, $v_i \neq 0$. Suppose that $\alpha_1 v_1 + \dots + \alpha_k v_k = 0$. Show that $\alpha_1 \lambda_1^\ell v_1 + \dots + \alpha_k \lambda_k^\ell v_k = 0$ for every ℓ .

Exercise 10.3. Let $\varphi : V \rightarrow V$ be a linear transformation. Let v_1, \dots, v_k be eigenvectors, $\varphi v_i = \lambda_i v_i$, $v_i \neq 0$. Suppose that the λ_i are distinct. Then v_1, \dots, v_k are linearly independent.

Definition 10.4. A set of vectors in \mathbb{R}^n is in **general position** if every n of them are linearly independent.

Exercise 10.5. Find a space curve $f : \mathbb{R} \rightarrow \mathbb{R}^n$ such that the points of the curve are in general position, i. e., for all $t_1 < t_2 < \dots < t_n$, the vectors $f(t_1), \dots, f(t_n)$ are linearly independent.

10.2 Real Euclidean Spaces, bilinear forms

Let V be a vector space over \mathbb{R} .

Definition 10.6. A **bilinear form** is a map $B : V \times V \rightarrow \mathbb{R}$ such that $B(x, \cdot)$ is linear for any $x \in V$ and $B(\cdot, y)$ is linear for any $y \in V$. I. e.,

$$(a1) \quad B(x, y_1 + y_2) = B(x, y_1) + B(x, y_2) \text{ for any } x, y_1, y_2 \in V;$$

$$(a2) \quad B(x, \alpha y) = \alpha B(x, y) \text{ for any } x, y \in V \text{ and } \alpha \in \mathbb{R};$$

$$(b1) \quad B(x_1 + x_2, y) = B(x_1, y) + B(x_2, y) \text{ for any } x_1, x_2, y \in V;$$

$$(b2) \quad B(\alpha x, y) = \alpha B(x, y) \text{ for any } x, y \in V \text{ and } \alpha \in \mathbb{R}.$$

Definition 10.7. Matrix associated with bilinear form, given a basis:

$$[B]_e = (\beta_{i,j})_{i,j=1}^n,$$

where $\beta_{i,j} = B(e_i, e_j)$.

Example 10.8. Let $n = 2$ and $B(x, y) = 3x_1y_1 + 5x_1y_2 - x_2y_1 + \sqrt{7}x_2y_2$. Then

$$[B] = \begin{pmatrix} 3 & 5 \\ -1 & \sqrt{7} \end{pmatrix}$$

Exercise 10.9. Show that for $x = \sum_{i=1}^n \alpha_i e_i$ and $y = \sum_{j=1}^n \gamma_j e_j$

$$B(x, y) = \sum_{i=1}^n \sum_{j=1}^n B(e_i, e_j) \alpha_i \gamma_j = (\alpha_1, \dots, \alpha_n) [B] (\gamma_1, \dots, \gamma_n)^T.$$

Theorem 10.10. Given a basis $e = (e_1, \dots, e_n)$, the correspondence $B \mapsto [B]_e$ is a bijection $\{\text{bilinear forms on } V\} \rightarrow M_n(\mathbb{R})$.

Definition 10.11. A bilinear form is **symmetric** if $B(x, y) = B(y, x)$ for all $x, y \in V$.

Exercise 10.12. Show that a bilinear form B is symmetric if and only if $[B]$ is symmetric.

Definition 10.13. Let B be a bilinear form. Then the function $x \mapsto B(x, x)$ is called a **quadratic form**.

Exercise 10.14. For every quadratic form $Q : V \rightarrow \mathbb{R}$ there exists unique symmetric bilinear form $B : V \times V \rightarrow \mathbb{R}$ such that $Q(x) = B(x, x)$.

Definition 10.15. A quadratic form Q is **positive semidefinite** if $(\forall x \in V)(Q(x) \geq 0)$.

Definition 10.16. A quadratic form Q is **positive definite** if $(\forall x \in V, x \neq 0)(Q(x) > 0)$.

Exercise 10.17. Show that the dot product of vectors $\underline{x} \cdot \underline{y} = \sum_{i=1}^n x_i y_i$ gives a positive definite quadratic form $\underline{x} \mapsto \underline{x} \cdot \underline{x}$.

We will use A^T to denote the transpose of matrix A .

Exercise 10.18. If $(\forall x, y \in \mathbb{F}^n)(x^T Ay = x^T By)$ then $A = B$.

Definition 10.19. Two $n \times n$ matrices A, B are **congruent** if there is a nonsingular $n \times n$ matrix S such that $A = S^T B S$.

Exercise 10.20. Show that if A and B are congruent and A is positive definite then B is positive definite.

Exercise 10.21 (Effect of change of basis on matrix of the bilinear form). Let $e = \{e_1, \dots, e_n\}$ be a basis and let $e' = \{e'_1, \dots, e'_n\}$ be another basis. Let $\sigma : e_i \mapsto e'_i$ and let $S = [\sigma]_e$ (recall that $[x]_e = S[x]_{e'}$). Show that

$$[B]_{e'} = S^T [B]_e S.$$

Definition 10.22. A real vector space V with a positive definite symmetric bilinear form $\langle \cdot, \cdot \rangle \rightarrow \mathbb{R}$ called “inner product” is an **Euclidean space**. Let $\|\cdot\| \rightarrow \mathbb{R}$ be defined by $\|x\| = \langle x, x \rangle$.

Example 10.23. $\mathbb{R}[x]$ with inner product $(f, g) = \int_0^1 f(x)g(x) dx$ is an Euclidean space.

Exercise 10.24 (Cauchy-Schwarz inequality). Let V be an Euclidean space. Show that for all $x, y \in V$

$$|\langle x, y \rangle| \leq \|x\| \cdot \|y\|.$$

Exercise 10.25 (Triangle inequality). Let V be an Euclidean space. Show that for all $x, y \in V$

$$\|x + y\| \leq \|x\| + \|y\|.$$

Definition 10.26. Vectors x, y are **perpendicular** $x \perp y$ if $\langle x, y \rangle = 0$.

Definition 10.27. The basis $e_1, \dots, e_n \in V$ is **orthonormal** if

$$(\forall i, j) \left(\langle e_i, e_j \rangle = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases} \right).$$

Theorem 10.28. Every Euclidean space of finite or countable dimension has an orthonormal basis. Moreover every orthonormal set of vectors can be extended to an orthonormal basis.

Definition 10.29. $\varphi : V \rightarrow V$ is an **orthogonal transformation** if it preserves inner products:

$$(\forall x, y \in V)(\langle \varphi x, \varphi y \rangle = \langle x, y \rangle).$$

Exercise 10.30. If φ is orthogonal then $(\forall x)(\|\varphi x\| = \|x\|)$.

Definition 10.31. **Distance** between two vectors $x, y \in V$ in an Euclidean vector space is $d(x, y) = \|x - y\|$.

Definition 10.32. A linear map $\varphi : V \rightarrow U$ between two Euclidean spaces is an **isometry** if it is a distance preserving bijection. Equivalently such a map preserves inner products.

Exercise 10.33. If e_1, \dots, e_n is an orthonormal basis and φ is orthogonal then $\varphi(e_1), \dots, \varphi(e_n)$ is also an orthonormal basis.

Exercise 10.34. If e_1, \dots, e_n is an orthonormal basis and $\varphi : V \rightarrow V$ is a linear transformation such that $\varphi(e_1), \dots, \varphi(e_n)$ is an orthonormal basis then φ is orthogonal.

Exercise 10.35. If φ maps **one** orthonormal basis to an orthonormal basis then it maps **all** orthonormal bases to orthonormal bases.

Theorem 10.36 (Uniqueness of n -dimensional real Euclidean space). *If V, W are n -dimensional real Euclidean spaces then $\exists \varphi : V \rightarrow W$ an isomorphism which is an isometry.*

Exercise 10.37. If e_1, \dots, e_n is an orthonormal basis and $x, y \in V$ then

$$\langle x, y \rangle = [x]_e^T \cdot [y]_e.$$

Note the right hand side is the *standard* inner product in \mathbb{R}^n .

Theorem 10.38.

- (1) If $\varphi : V \rightarrow V$ is a linear transformation then $B(x, y) = \langle x, \varphi y \rangle$ is a bilinear form.
- (2) If $\psi : V \rightarrow V$ is a linear transformation then $B(x, y) = \langle \psi x, y \rangle$ is a bilinear form.
- (3) If B is a bilinear form $B : V \times V \rightarrow \mathbb{R}$ then $\exists! \varphi, \psi \in V \rightarrow V$ such that

$$(\forall x, y)(B(x, y) = \langle x, \varphi y \rangle = \langle \psi x, y \rangle).$$

Corollary 10.39.

$$(\forall \varphi)(\exists! \psi)(\forall x, y \in V)(\langle x, \varphi y \rangle = \langle \psi x, y \rangle).$$

Definition 10.40. The **transpose** φ^T of a linear transformation φ is the unique transformation ψ given by Corollary 10.39.

Definition 10.41. $\varphi : V \rightarrow V$ is **symmetric** if $\varphi = \varphi^T$.

Theorem 10.42 (Spectral Theorem). *If φ is a symmetric linear transformation of a real Euclidean space then φ has an orthonormal eigenbasis.*

Definition 10.43. A subspace $U \leq V$ is **invariant** under φ if $x \in U \Rightarrow \varphi x \in U$.

Exercise 10.44. If $U \leq V$ is invariant under $\varphi : V \rightarrow V$ then U^\perp is invariant under φ^T .

Exercise 10.45. All complex eigenvalues of a symmetric transformation are real.

Exercise 10.46. If φ is symmetric then it has an eigenvector.

Proof of Spectral Theorem:

By induction on $n = \dim V$.

Exercise 10.47. Prove the base case $n = 1$.

Given $\varphi : V \rightarrow V$, $\dim V = n$, $\varphi = \varphi^T$. By Exercise 10.46 $\exists x \neq 0 : \varphi x = \lambda_1 x$. Let $e_1 = x/\|x\|$. The space $U = \text{Span}(e_1)$ is invariant under φ , $\dim U = 1$.

Hence, by Exercise 10.44, U^\perp is invariant under $\varphi^T = \varphi$. Hence the restriction of φ to U^\perp , $\varphi' = \varphi|_{U^\perp}$ is a linear transformation $U^\perp \rightarrow U^\perp$ where $\dim U^\perp = n - 1$.

We have $(\forall u, v \in U^\perp)(\langle u, \varphi'v \rangle = \langle \varphi'u, v \rangle)$ and hence φ' is symmetric. By induction hypothesis, φ' has an orthonormal eigenbasis $e_2, \dots, e_n \in U^\perp$.

e_1, \dots, e_n is an orthonormal eigenbasis for φ in V because $e_1 \perp e_2, \dots, e_n$. \square

Corollary 10.48 (Spectral Theorem, 2nd form). *If Q is a quadratic form $Q : V \rightarrow V$ on the real n -dimensional Euclidean space V then there exists an orthonormal basis \underline{e} such that*

$$Q(x) = \lambda_1 x_1^2 + \dots + \lambda_n x_n^2,$$

where $[x]_{\underline{e}} = (x_1, \dots, x_n)^T$.

Example 10.49. Let $Q(x_1, x_2) = x_1^2 + 3x_1x_2 + 2x_2^2$. We have $Q(x_1, x_2) = (x_1, x_2)B(x_1, x_2)^T$ where

$$B = \begin{pmatrix} 1 & 3/2 \\ 3/2 & 2 \end{pmatrix}.$$

The characteristic polynomial of B is $x^2 - 3x + 3/4$, which has roots $\lambda_{1,2} = (3 \pm \sqrt{6})/2$. Hence B is orthogonally similar to the diagonal matrix with diagonal λ_1, λ_2 . Thus there is an orthonormal basis \underline{f} such that $Q(x_1, x_2) = \lambda_1 u_1^2 + \lambda_2 u_2^2$ where $\underline{x} = u_1 f_1 + u_2 f_2$.

Let A be a real symmetric matrix with eigenvalues $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$.

Exercise 10.50. Prove:

$$\lambda_1 = \max_{\|x\|=1} x^T A x = \max_{x \neq 0} \frac{x^T A x}{\|x\|}.$$

Exercise 10.51. Let G be a graph. Let A be its adjacency matrix (i. e., $a_{ij} = 1$ if $i \sim j$ (i, j are adjacent); and $a_{ij} = 0$ otherwise). Let λ_1 be the largest eigenvalue of A . Prove $\deg_{\max} \geq \lambda_1 \geq \deg_{\text{avg}}$ (i. e., the largest eigenvalue of the adjacency matrix is sandwiched between the maximum and average degrees.)

Chapter 11

11th day, Tuesday 7/13/04 (Scribe: Justin Noel)

11.1 Complex vector spaces, sesquilinear forms

Unless otherwise stated, V is a complex vector space. And $\varphi, \psi : V \rightarrow V$ are linear transformations.

Definition 11.1. A function $B : V \times V \rightarrow \mathbb{C}$ is called a **sesquilinear form** if the following hold:

- (1) $(\forall x, y_1, y_2 \in V) B(x, y_1 + y_2) = B(x, y_1) + B(x, y_2)$
- (2) $(\forall \alpha \in \mathbb{C})(B(x, \alpha y) = \alpha B(x, y))$
- (3) $(\forall x_1, x_2, y \in V)(B(x_1 + x_2, y) = B(x_1, y) + B(x_2, y))$
- (4) $(\forall \alpha \in \mathbb{C})(B(\alpha x, y) = \bar{\alpha} B(x, y))$

Definition 11.2. If $A \in M_n(\mathbb{C})$ we set $A^* = \overline{A^T}$. The matrix A^* is called the **Hermitian adjoint** of A .

Definition 11.3. A sesquilinear form $B : V \times V \rightarrow \mathbb{C}$ is said to be **Hermitian** if it additionally satisfies:

$$(5) B(y, x) = \overline{B(x, y)}$$

Exercise 11.4. Show that for $A, B \in M_n(\mathbb{C})$, $(AB)^* = B^* A^*$.

Definition 11.5. A matrix $A \in M_n(\mathbb{C})$ is said to be **Hermitian** if $A = A^*$.

Remark 11.6. Note that every Hermitian matrix must have its diagonal entries be real.

Definition 11.7. As with bilinear forms over \mathbb{R} we can associate a quadratic form Q to a sesquilinear form B by $Q(x) = B(x, x)$.

Exercise 11.8. If B is sesquilinear then B is determined by the corresponding quadratic form (i.e. if $(\forall x \in V)(B_1(x, x) = B_2(x, x))$ then $(\forall x, y \in V)(B_1(x, y) = B_2(x, y))$).

Exercise 11.9. If $(\forall x \in V)(B(x, x) \in \mathbb{R})$ then $(\forall x, y \in V)(B(x, y) = \overline{B(x, y)})$ (i.e. B is sesquilinear).

Definition 11.10. Quadratic form Q on V is positive semidefinite if $(\forall x \in V)(Q(x) \geq 0)$.

Definition 11.11. Quadratic form Q on V is positive definite if $(\forall x \in V \neq 0)(Q(x) > 0)$.

Definition 11.12. We say that V is a **\mathbb{C} -Euclidean space** if it is equipped with a positive definite Hermitian form which we call it's inner product and denote $\langle x, y \rangle$. That is to say $\langle x, y \rangle$ satisfies the following properties:

- (1) $(\forall x, y_1, y_2 \in V)(\langle x, y_1 + y_2 \rangle = \langle x, y_1 \rangle + \langle x, y_2 \rangle)$
- (2) $(\forall x, y \in V, \forall \alpha \in \mathbb{C})(\langle x, \alpha y \rangle = \alpha \langle x, y \rangle)$
- (5) $(\forall x, y \in V)(\langle y, x \rangle = \overline{\langle x, y \rangle})$
- (6) $(\forall x \in V - \{0\})(\langle x, x \rangle > 0)$

Exercise 11.13. Show that (1), (2), (5) imply (3) and (4).

Proposition 11.14 (Change of Basis). *To express a bilinear form with respect to a new basis there is a matrix $S \in M_n(\mathbb{C})$ such that $[B]_{new} = S^*[B]_{old}S$.*

Definition 11.15. We say that a basis e_1, \dots, e_n of a Euclidean space is an **orthonormal basis** if $\langle e_i, e_j \rangle$ is 1 if $i = j$ and 0 otherwise.

Definition 11.16. The linear map φ is a **unitary transformation** if $\langle \varphi(x), \varphi(y) \rangle = \langle x, y \rangle$.

Theorem 11.17. *The map φ is unitary iff $[\varphi]^*[\varphi] = I$ with respect to an orthonormal basis.*

Definition 11.18. $A \in M_n(\mathbb{C})$ is a **unitary matrix** if $AA^* = I$.

Exercise 11.19. Show that $A \in M_n(\mathbb{C})$ is unitary iff the rows of A form an orthonormal basis under the standard inner product $\langle x, y \rangle = [x]^*[y]$.

Exercise 11.20. Show that A is unitary iff its columns form an orthonormal basis.

Definition 11.21. The **general linear group**, $GL(n, F)$ is the set of nonsingular matrices over the field F with the group operation multiplication.

Definition 11.22. The **unitary group**, $U(n)$ is the set of unitary matrices with the group operation multiplication.

Definition 11.23. The **orthogonal group**, $O(n)$ is the set of orthogonal matrices with the group operation multiplication.

Definition 11.24. The **special linear group**, $SL(n, F) = \{A \in M_n(\mathbb{F}) \mid \det(A) = 1\}$ with the group operation multiplication.

Definition 11.25. The **special orthogonal group**, $SO(n) = O(n) \cap SL(n, \mathbb{R})$ with the group operation multiplication.

Exercise 11.26. If $A \in O(n)$ then $\det(A) = \pm 1$.

Exercise 11.27. If $A \in U(n)$ then $|\det(A)| = 1$.

Exercise 11.28. Show that $|O(n) : SO(n)| = 2$.

Exercise 11.29. Show that $|U(n) : SU(n)| = \infty$, and that it has the cardinality of the continuum.

Exercise 11.30. Show that $O(n)/SO(n) \cong \mathbb{Z}_2$.

Exercise 11.31. Show that $U(n)/SU(n) \cong \{z \in \mathbb{C} \mid |z| = 1\}$.

Definition 11.32. Over any field \mathbb{F} we define the **adjoint** of $A \in M_n(\mathbb{F})$ with $A = (\alpha_{ij})$ as

$$\text{adj}(A) = \begin{pmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & \ddots & \vdots \\ A_{n1} & \cdots & A_{nn} \end{pmatrix}$$

Where

$$A_{ji} = (-1)^{i+j} \begin{vmatrix} \alpha_{11} & \cdots & \widehat{\alpha_{1j}} & \cdots & \alpha_{1n} \\ \vdots & & \ddots & & \vdots \\ \widehat{\alpha_{i1}} & \cdots & \widehat{\alpha_{ij}} & \cdots & \widehat{\alpha_{in}} \\ \vdots & & \ddots & & \vdots \\ \alpha_{n1} & \cdots & \widehat{\alpha_{nj}} & \cdots & \alpha_{nn} \end{vmatrix}$$

Where the "hats" $\widehat{\alpha_{ij}}$ indicate that we are omitting the i th row and the j th column.

Exercise 11.33. Show that

$$A \cdot \text{adj}(A) = \begin{pmatrix} \det(A) & & \\ & \ddots & \\ & & \det(A) \end{pmatrix} = \det(A) \cdot I$$

Corollary 11.34. If $\det(A) \neq 0$ then $A^{-1} = \frac{1}{\det(A)} \text{adj}(A)$.

Exercise 11.35. $A \in M_n(\mathbb{Z})$ is an integral matrix such that $A^{-1} \in M_n(\mathbb{Z})$ iff $\det(A) = \pm 1$.

Definition 11.36. $GL(n, \mathbb{Z})$ is the set of $n \times n$ integral matrices which have an inverse which is an integral matrix. Or equivalently the set of $n \times n$ matrices with determinant ± 1 .

Exercise 11.37. Show that $|GL(n, \mathbb{Z}) : SL(n, \mathbb{Z})| = 2$.

Exercise 11.38. Show that for any φ the map $B : V \times V \rightarrow V$ defined by $B(x, y) = \langle x, \varphi(y) \rangle$ is sesquilinear.

Exercise 11.39. Show that for any ψ the map $B : V \times V \rightarrow V$ defined by $B(x, y) = \langle \psi(x), y \rangle$ is sesquilinear.

Theorem 11.40. Let $\forall B : V \times V \rightarrow V$ be sesquilinear. $(\exists! \varphi, \psi)(\forall x, y \in V)(B(x, y) = \langle \psi(x), y \rangle = \langle x, \varphi(y) \rangle)$.

Corollary 11.41. $(\forall \varphi)(\exists! \psi)$ such that $(\forall x, y \in V)(\langle \psi(x), y \rangle = \langle x, \varphi(y) \rangle)$.

Definition 11.42. We define φ^* to be ψ if $(\forall x, y \in V)(\langle \psi(x), y \rangle = \langle x, \varphi(y) \rangle)$.

Exercise 11.43. Show that φ is unitary iff $\varphi^* = \varphi^{-1}$.

Exercise 11.44. With respect to an orthonormal basis show that φ is unitary iff $[\varphi] \in U(n)$.

Exercise 11.45. If λ is an eigenvalue of the unitary transformation φ then $|\lambda| = 1$.

Chapter 12

12th day, Wednesday 7/14/04 (Scribe: Richard Cudney)

12.1 Complex Euclidean (unitary) spaces

Note that $M_n(\mathbb{C})$ is a vector space of dimension n^2 : for a basis is given by the matrices which are zero in all but one entry, and one in that one.

Recall the standard inner product on \mathbb{C}^n : If x and y are two column vectors, then

$$\langle x, y \rangle = x^* y = \sum_{i=1}^n \bar{x}_i y_i.$$

It has the following properties:

It is **sesquilinear**:

- (1) $(\forall x, y_1, y_2 \in V)(\langle x, y_1 + y_2 \rangle = \langle x, y_1 \rangle + \langle x, y_2 \rangle)$
- (2) $(\forall \alpha \in \mathbb{C})(\langle x, \alpha y \rangle = \alpha \langle x, y \rangle)$
- (3) $(\forall x_1, x_2, y \in V)(\langle x_1 + x_2, y \rangle = \langle x_1, y \rangle + \langle x_2, y \rangle)$
- (4) $(\forall \alpha \in \mathbb{C})(\langle \alpha x, y \rangle = \bar{\alpha} \langle x, y \rangle)$

It is **Hermitian**:

- (5) $\langle y, x \rangle = \overline{\langle x, y \rangle}$

It is **positive definite**:

- (6) $\langle x, x \rangle > 0$ unless $x = 0$.

Definition 12.1. A **complex Euclidean space**, also called a **unitary space**, is a complex vector space V given with a bilinear form, called the “inner product” satisfying the above properties.

Definition 12.2. $e_1, \dots, e_k \in V$ is an **orthonormal** (abbreviated ON) system if $\langle e_i, e_j \rangle = 0$ if $i \neq j$ and 1 otherwise.

Exercise 12.3. Every ON system is linearly independent.

Theorem 12.4. *Every complex Euclidean space of finite or countable dimension has an orthonormal basis (abbreviated ONB) and every ON system can be extended to an ONB.*

[Proof: “Gram-Schmidt orthogonalization”]

Exercise 12.5. Cauchy-Schwarz inequality: $|\langle x, y \rangle| \leq \|x\| \|y\|$ for real and complex Euclidean spaces.

Exercise 12.6. As a consequence of the last exercise, derive the **triangle inequality** $\|x + y\| \leq \|x\| + \|y\|$.

12.2 Unitary transformations

Definition 12.7. $\varphi : V \rightarrow V$ is **unitary** if $(\forall x, y \in V)(\langle x, y \rangle = \langle \varphi x, \varphi y \rangle)$.

Exercise 12.8. φ is unitary iff $(\forall x \in V)(\|\varphi x\| = \|x\|)$.

Definition 12.9. $A \in M_n(\mathbb{C})$ is **unitary** if $AA^* = A^*A$ where here A^* is the conjugate transpose of A .

Definition 12.10. $U(V)$ is the set of unitary transformations of the complex Euclidean space V .

Definition 12.11. $U(n)$ is the set of $n \times n$ unitary matrices.

Exercise 12.12. If e is an ONB, then φ is a unitary linear transformation if and only if $[\varphi]_e$ is a unitary matrix.

Exercise 12.13. $U(V)$ is a group under composition.

Observe: If $\varphi \in U(V)$ then if e_1, \dots, e_n is an ONB, so is $\varphi e_1, \dots, \varphi e_n$.

Exercise 12.14. If e_1, \dots, e_n is an ONB, and $\varphi e_1, \dots, \varphi e_n$ is also, then $\varphi \in U(V)$.

Corollary 12.15. *If φ takes one ONB to an ONB then φ moves every ONB to an ONB.*

Theorem 12.16. *If e_1, \dots, e_n is an ONB then $\langle x, y \rangle = [x]_e^* [y]_e$.*

Theorem 12.17. $(\forall \varphi : V \rightarrow V)(\exists ! \psi : V \rightarrow V)(\forall x, y \in V)(\langle \psi x, y \rangle = \langle x, \varphi y \rangle)$

12.3 Hermitian forms and self-adjoint transformations

Definition 12.18. $\psi = \varphi^*$, the **Hermitian adjoint** of φ , so that $(\forall x, y \in V)(\langle \varphi^* x, y \rangle = \langle x, \varphi y \rangle)$.

Observation: If e_1, \dots, e_n is an ONB then $[\varphi^*]_e = [\varphi]_e^*$, where the left $*$ is the Hermitian adjoint, and the second is the conjugate transpose.

Theorem 12.19. $(\forall B : V \times V \rightarrow \mathbb{C} \text{ sesquilinear form})$
 $(\exists ! \varphi, \psi : V \rightarrow V)(\forall x, y \in V)(B(x, y) = \langle \psi x, y \rangle = \langle x, \varphi y \rangle)$.

Definition 12.20. $[B]_e = (B(e_i, e_j))$.

This matrix depends on the choice of basis as follows: $[B]_{\text{new}} = S^*[B]_{\text{old}}S$, where $S = [\text{new basis}]_{\text{old basis}}$, so that the column vectors of S are the coordinates of the new basis in terms of the old basis.

Definition 12.21. φ is a **self-adjoint** transformation if $\varphi = \varphi^*$.

Exercise 12.22. $\varphi : V \rightarrow V$ is self-adjoint iff $[\varphi]^* = [\varphi]$ with respect to an ONB.

Thus, if the latter condition is true with respect to one ONB it is true with respect to all ONBs.

Exercise 12.23. Prove: the sesquilinear form $B(x, y) = \langle x, \varphi y \rangle$ is Hermitian (i. e., $B(y, x) = \overline{B(x, y)}$) iff φ is self-adjoint.

Exercise 12.24. If φ is self-adjoint then all eigenvalues of φ are real.

Lemma 12.25. If $U \leq V$ and U is invariant under $\varphi : V \rightarrow V$ (meaning $x \in U \rightarrow \varphi(x) \in U$) then U^\perp is invariant under φ^* .

Theorem 12.26 (Spectral Theorem). If $\varphi : V \rightarrow V$ is self-adjoint then φ has an orthonormal eigenbasis (ONEB).

Proof: We will prove it by induction on $\dim(V)$.
 Let λ_1 be a root of the characteristic polynomial of φ .

$$(\exists x \neq 0)(\varphi x = \lambda_1 x)$$

Let $e_1 = \frac{1}{\|x\|} \cdot x$, so that e_1 is an orthonormal eigenvector. Let $U = \text{Span}(e_1)^\perp$. U is invariant under φ .

Exercise 12.27. $(\varphi|_U)^* = \varphi|_U$, where $\varphi|_U$ is φ restricted to U .

$\dim(U) = n - 1$ so by induction hypothesis $\varphi|_U$ has an ONEB e_2, \dots, e_n . $e_1 \perp U$ so that $e_1 \perp e_2, \dots, e_n$, so e_1, \dots, e_n is an ONEB.

Theorem 12.28. *Let V be a complex Euclidean space. $(\forall \varphi : V \rightarrow V)(\exists \text{ ONB } e)([\varphi]_e \text{ is upper triangular})$.*

Hint: quotient space.

Here is the equivalent statement about matrices:

Theorem 12.29. $(\forall A \in M_n(\mathbb{C}))(\exists S \in U(n))(S^{-1}AS \text{ is upper triangular})$.

Proof of matrix form by induction on n :

Find an eigenvector and normalize it so it is an eigenvector e_1 of unit length. Extend e_1 to an ONB. Let $T = [e_1 \dots e_n]$ be the change of basis matrix, which is unitary since e_1, \dots, e_n is an ONB.

$$T^{-1}AT = \begin{pmatrix} \lambda_1 & \text{junk} \\ 0 & B \end{pmatrix}$$

where B is an $(n-1) \times (n-1)$ matrix. So by induction there is an $R \in U(n-1)$ such that $R^{-1}BR$ is upper triangular. Let

$$Q = \begin{pmatrix} 1 & 0 \\ 0 & R \end{pmatrix}.$$

Then $S = TQ$ is the desired matrix.

12.4 Normal matrices

Definition 12.30. $\varphi : V \rightarrow V$ is **normal** if $\varphi\varphi^* = \varphi^*\varphi$.

Definition 12.31. $A \in M_n(\mathbb{C})$ is **normal** if $AA^* = A^*A$.

Exercise 12.32. With respect to an ONB e , the linear transformation φ is normal iff the matrix $[\varphi]_e$ is normal.

Theorem 12.33 (Generalized Spectral Theorem). *A transformation of a complex Euclidean space is normal if and only if it has an ONB. We shall prove this theorem in matrix form, see Theorem 12.39 below.*

Remark 12.34. Diagonal matrices are normal because all diagonal matrices commute with one another.

Definition 12.35. $A \sim B$ if $(\exists S)(B = S^{-1}AS)$.

Definition 12.36. $A \sim_u B$ if $(\exists S \in U(n))(B = S^{-1}AS)$.

Exercise 12.37. If $A \sim_u B$ then (if A is normal then B is normal).

Exercise 12.38. If A is normal and upper triangular then A is diagonal.

Theorem 12.39 (Generalized Spectral Theorem). *If A is an $n \times n$ matrix over \mathbb{C} then A is normal $\iff A \sim_u D$ where D is a diagonal matrix.*

Proof: The left hand implication follows from the above exercises. To prove the forward implication, let us be given A , a normal matrix. By the above theorem, $A \sim_u T$ where T is upper triangular. By an above exercise, T is normal since A is, and thus by another exercise it is diagonal since T is upper triangular.

Exercise 12.40. Show that this Theorem is equivalent to Theorem 12.33 above.

Remark 12.41. The spectral theorem is an immediate consequence, for if $A=A^*$ then clearly A is normal since any matrix commutes with itself. Thus A is unitarily similar to a diagonal matrix, where the entries are real because they are the eigenvalues of a self-adjoint matrix.

Exercise 12.42. If A is normal with real eigenvalues then A is self-adjoint.

Remark 12.43. If A is unitary then $A^* = A^{-1}$ so $AA^* = A^*A = I$, so A is normal.

Theorem 12.44. *If $A \in U(n)$ then all eigenvalues have unit absolute value.*

Proof: Let x be a non-zero eigenvector with eigenvalue λ . $\langle x, x \rangle = \langle Ax, Ax \rangle = \langle \lambda x, \lambda x \rangle = \lambda \bar{\lambda} \langle x, x \rangle$ so $|\lambda| = 1$.

Exercise 12.45. A diagonal matrix is unitary if and only if all of the diagonal entries have absolute value one.

Exercise 12.46. $A \in U(n) \iff A \sim_u D$, D a diagonal matrix with entries of absolute value one.

We can restate the above result on normal matrices as a purely geometric statement about normal linear transformations.

Theorem 12.47. $\varphi : V \rightarrow V$ is normal if and only if it has an ONEB.

Reminder(needed for the proof): The column vectors of the matrix associated to a linear transformation are the images of the basis vectors under that linear transformation.

Exercise 12.48. Self-adjoint transformations are not closed under multiplication.

Theorem 12.49. *Suppose A is a positive definite symmetric real matrix. Then A has a unique positive definite symmetric square root \sqrt{A} .*

Proof: A is conjugate to a diagonal matrix D with positive real entries. Let \sqrt{D} be the diagonal matrix whose entries are the positive square roots of those of D . Now un-conjugate \sqrt{D} to get a square root of A with the desired properties.

Exercise 12.50. Prove the uniqueness statement in the above theorem.

12.5 Gramian

Let V be a real or complex Euclidean space, and $v_1, \dots, v_k \in V$.

Definition 12.51. $G(v_1, \dots, v_k) = (\langle v_i, v_j \rangle)$ is called the **Gram matrix** of v_1, \dots, v_k .

Definition 12.52. The determinant of $G(v_1, \dots, v_k)$ is the **Gramian** of v_1, \dots, v_k .

Exercise 12.53. The Gram matrix is positive semi-definite. It is positive definite $\iff \det G \neq 0 \iff v_1, \dots, v_k$ are linearly independent.

Hint: For any matrix A , not necessarily square, AA^* is positive semi-definite.

Exercise 12.54. Suppose that the characteristic polynomial of A splits into linear factors $x - \lambda_i$. Then $\det A = \prod \lambda_i$. Note that in particular the hypotheses will be satisfied over \mathbb{C} .

Corollary 12.55. If A is positive definite and $A^* = A$, then $\det A > 0$.

Exercise 12.56. Assume we are over \mathbb{R} . Prove that the k -dimensional volume of the parallelepiped spanned by v_1, \dots, v_k is $\sqrt{\det G(v_1, \dots, v_k)}$. Note that the parallelepiped spanned by v_1, \dots, v_k is by definition the set of linear combinations $\sum \lambda_i v_i$ where $0 \leq \lambda_i \leq 1$ for each i .

Lemma 12.57. Elementary operations $v_i \mapsto v_i - \alpha v_j$ do not change the Gramian, the determinant, or the volume of the parallelepiped spanned by v_1, \dots, v_k .

Exercise 12.58. v_1, \dots, v_k is an orthogonal system if and only if the Gram matrix is diagonal.

Exercise 12.59. Show the **Hadamard inequality**: If $A = [a_1 \dots a_n]$ then $|\det A| \leq \prod_{i=1}^n \|a_i\|$.

Exercise 12.60. If $v_1, \dots, v_n \in \mathbb{R}^n$ then $\det(v_1, \dots, v_n) = \pm$ the volume of the parallelepiped spanned by v_1, \dots, v_n .

Exercise 12.61. Suppose $A = A^T$ is a real symmetric matrix. Then A is positive definite iff all the top left corner determinants are positive.

Theorem 12.62 (Interlacing theorem). If $A = A^T$ is a symmetric matrix, and B is the matrix we get by removing the i^{th} row and column, let

$$\lambda_1 \geq \dots \geq \lambda_n$$

be the eigenvalues of A and

$$\mu_1 \geq \dots \geq \mu_{n-1}$$

the eigenvalues of B . Then

$$\lambda_1 \geq \mu_1 \geq \lambda_2 \geq \mu_2 \geq \dots \geq \mu_{n-1} \geq \lambda_n$$

.

Exercise 12.63. Prove the interlacing theorem using the following characterization of eigenvalues.

Lemma 12.64 (Courant-Fischer theorem). *Let $A \in M_n(\mathbb{C})$, $A^* = A$. Then we have the following expressions for the eigenvalues of A :*

$$\lambda_1 = \max_{x \neq 0} \frac{x^* A x}{x^* x},$$
$$\lambda_n = \min_{x \neq 0} \frac{x^* A x}{x^* x},$$
$$\lambda_i = \max_{\dim(U)=i} \min_{x \in U, x \neq 0} \frac{x^* A x}{x^* x}.$$

Hint: By the spectral theorem, it suffices to show these relations for real diagonal matrices.

Chapter 13

13th day, Thursday 7/15/04 (Scribe: Nick Gurski)

13.1 Explicit form of the inverse matrix

Let $A = (\alpha_{ij})$ be an $n \times n$ matrix over a field F . Then the adjoint of A , written $\text{adj}(A)$, is the matrix $(a_{ij}) = ((-1)^{i+j} \det A_{\hat{j}i})$ where $A_{\hat{j}i}$ is the matrix A with row j and column i deleted (note how i and j switch roles). This allows us to prove the formula $A \cdot \text{adj}(A) = \det A \cdot I$. If we call the lefthand side $B = (\beta_{ij})$, consider β_{11} . This is equal to

$$\sum_j \alpha_{1j} (-1)^{1+j} \det A_{\hat{1}\hat{j}},$$

which is equal to the determinant of A computed by expanding along the first row; the same calculation works for β_{ii} . To show that the off-diagonal entries are zero, we compute β_{12} and note that the same technique will work for any off-diagonal. Now β_{12} is

$$\sum_j \alpha_{1j} (-1)^{2+j} \det A_{\hat{2}\hat{j}},$$

which is the determinant of the matrix A' which is the same as A except row 2 is replaced by a copy of row 1. This is zero, since the determinant of any matrix with linearly dependent rows is zero, proving the formula. Thus we have shown that if the determinant of A is nonzero, then

$$A^{-1} = \frac{1}{\det A} \text{adj } A.$$

This formula answers the question of when a matrix with integer entries has an inverse with integer entries. Since the determinant of A^{-1} is $1/\det A$, we know that if A^{-1} has integer entries then the determinant of A is ± 1 ; by the formula above we find that the converse is true as well. The adjoint of an integral matrix is integral, and if $\det A = \pm 1$, then the righthand side of the formula for the inverse is a matrix with integer entries.

13.2 Gram-Schmidt orthogonalization

Now we will consider a process called Gram-Schmidt orthogonalization. This is a machine that takes as input a sequence of vectors v_1, v_2, \dots and gives as output a sequence of orthogonal vectors b_1, b_2, \dots . The defining properties of the Gram-Schmidt process are that

1. the span of v_1, \dots, v_k is the same as the span of b_1, \dots, b_k (which we shall call U_k),
2. for all $i \neq j$, $\langle b_i, b_j \rangle = 0$, and
3. $v_k - b_k \in U_{k-1}$.

Note that $U_0 = \{0\}$; thus $v_1 - b_1 = 0$ or $v_1 = b_1$. We then get that $v_2 - b_2 = \mu_{2,1}b_1$, or rewriting we get

$$v_2 = b_2 + \mu_{2,1}b_1.$$

The general formula then becomes

$$v_k = b_k + \mu_{k,k-1}b_{k-1} + \mu_{k,k-2}b_{k-2} + \dots + \mu_{k,1}b_1.$$

Thus we must determine the coefficients $\mu_{k,j}$. To calculate $\mu_{3,2}$, we consider its equation and take the inner product of both sides with b_2 . This yields

$$\langle b_2, v_3 \rangle = 0 + \mu_{3,2}\|b_2\|^2 + 0,$$

using orthogonality and knowledge of b_1, b_2 . This gives an inductive definition of the $\mu_{k,j}$ as

$$\mu_{k,j} = \frac{\langle b_j, v_k \rangle}{\|b_j\|^2}.$$

This verifies the uniqueness of the coefficients. Assuming we have b_1, \dots, b_{k-1} , define $b_k = v_k - \sum \mu_{k,j}b_j$. We must verify the three properties above. The first follows by an inductive argument using this definition of b_k ; the second is a consequence of the definition of the coefficients. The third is immediate from the definition.

Exercise 13.1. Show that

$$\det G(b_1, \dots, b_k) = \det G(v_1, \dots, v_k),$$

where G is the Gram matrix. If the field of definition is \mathbb{R} , show that

$$\text{vol}(b_1, \dots, b_k) = \text{vol}(v_1, \dots, v_k),$$

where $\text{vol}(v_1, \dots, v_k)$ stands for the volume of the parallelepiped spanned by v_1, \dots, v_k . This parallelepiped is the set $\{\sum \alpha_i v_i \mid 0 \leq \alpha_i \leq 1\}$.

Exercise 13.2. Show that $b_k = 0$ if and only if $v_k \in U_{k-1}$.

We also have that if the vectors v_i form a basis, then so do the new vectors b_i , and thus an orthogonal basis. By scaling, we have thus shown the existence of an orthonormal basis.

Exercise 13.3. If the vectors v_1, \dots, v_k are orthogonal, then $b_i = v_i$ for $i = 1, \dots, k$.

It is easy to see that if $V = \text{Span}\{v_1, \dots, v_n\}$, then the vectors b_1, \dots, b_n consist of an orthogonal basis for V plus some number of additional zero vectors.

Example 13.4. Real polynomials.

The set of all polynomials with real coefficients, $\mathbb{R}[x]$, is a vector space over \mathbb{R} . We can give it an inner product by picking a density function $\rho(x)$ (with the properties that $\rho \geq 0$ and $0 < \int_{-\infty}^{\infty} x^{2n} \rho(x) dx < \infty$ for all n); this defines the inner product

$$\langle f(x), g(x) \rangle = \int_{-\infty}^{\infty} f(x)g(x)\rho(x)dx.$$

The standard basis of $\mathbb{R}[x]$ is $1, x, x^2, \dots$. Using Gram-Schmidt, we get a new orthogonal basis (depending on ρ), consisting of the polynomials $f_0 = 1, f_1, f_2, \dots$, where $\deg f_k = k$. These give special kinds of polynomials depending on which ρ we choose. If $\rho(x) = \sqrt{1-x^2}$ for $-1 \leq x \leq 1$ and $\rho(x) = 0$ otherwise, these are called the Chebyshev polynomials of the first kind; if instead we take $\rho(x) = 1/\sqrt{1-x^2}$ on the interval $(-1, 1)$, these are Chebyshev polynomials of the second kind. If $\rho(x) = \exp(-x^2)$, we get the Hermite polynomials.

Exercise 13.5. Regardless of ρ , each f_k has k real distinct roots, and the roots of f_{k-1} interlace those of f_k .

13.3 Algebraic numbers, minimal polynomials

We return now to more algebraic considerations. We say that a field F is algebraically closed if every nonconstant polynomial in $F[x]$ has a root. An example of such is the complex numbers, \mathbb{C} , and this result has the name of the Fundamental Theorem of Algebra. The real numbers are not algebraically closed, as $x^2 + 1$ has no root.

Exercise 13.6. There exists a countable, algebraically closed field.

Exercise 13.7. Every field F is contained in an algebraically closed field; moreover, there exists a unique (up to the appropriate isomorphism) smallest such algebraically closed field. This latter field is called the algebraic closure of F .

Exercise 13.8. Let A be the algebraic closure of the rational numbers. Then A is countable. (This is just one way to solve Exercise 5.)

Definition 13.9. We say that a complex number α is **algebraic** if there is a nonzero polynomial $f(x) \in \mathbb{Z}[x]$ such that $f(\alpha) = 0$.

Exercise 13.10. Show that the the set of all algebraic numbers forms a field.

Exercise⁺ 13.11. The field of all algebraic numbers in Exercise 13.10 is algebraically closed.

A corollary of these last two exercises is that the field A from exercise 7 is actually the same as the field of all algebraic numbers.

Let α be an algebraic number. Then the minimal polynomial $m_\alpha(x)$ of α is the polynomial with integer coefficients of least degree such that α is a root of m_α .

Exercise 13.12. The minimal polynomial is unique up to constant multiples.

Exercise 13.13. For any polynomial $f(x)$ with rational coefficients, $f(\alpha) = 0$ if and only if m_α divides f .

Example 13.14. Some minimal polynomials. The minimal polynomial of $\sqrt{2}$ is $x^2 - 2$. Let ω be the primitive cube root of unity given by $\frac{-1}{2} + \frac{\sqrt{3}}{2}i$. Then $m_\omega = x^2 + x + 1$, not $x^3 - 1$ since this last polynomial is not of minimal degree.

Exercise 13.15. The polynomial m_α is irreducible over the rational numbers.

Exercise 13.16. If f is a polynomial with integer coefficients such that $f(\alpha) = 0$ and f is irreducible, then $f = m_\alpha$.

13.4 The minimal polynomial of a matrix

Now we shall try to understand the minimal polynomial for a matrix A . By the Cayley-Hamilton theorem, we know that A satisfies its own characteristic polynomial. If we let f_A denote the characteristic polynomial of A , this means that $f_A(A) = 0$. Without using this theorem, we can easily prove that every matrix A is the root of some polynomial. If $f(x) = a_0 + a_1x + \cdots + a_nx^n$, then recall that

$$f(A) = a_0 \cdot I + a_1A + a_2A^2 + \cdots + a_nA^n.$$

Thus to show that A is the root of some f , we only need to show that there is some k such that the matrices I, A, A^2, \dots, A^k are linearly dependent. If A is an $n \times n$ matrix, then we know that $k = n^2$ is such a value, since the vector space $M_n(F)$ is only n^2 -dimensional and I, A, \dots, A^{n^2} is a collection of $n^2 + 1$ elements.

Now that we know (in two different ways) that every matrix is the root of a polynomial, we can ask about a minimal such. This is a quite different question than finding minimal polynomials for algebraic numbers, indicated in the following exercise.

Exercise 13.17. Over the real numbers, there are infinitely many $n \times n$ -matrices ($n \geq 2$) A with $A^2 = I$; the same holds for $A^2 = -I$.

We define m_A to be the minimal polynomial (over F) for which A is a root, just as we did for algebraic numbers.

Exercise 13.18. $f(A) = 0$ if and only if m_A divides f .

Exercise 13.19. Let f be any polynomial, and D a diagonal matrix with entries λ_i . Then $f(D)$ is a diagonal matrix with entries $f(\lambda_i)$.

Using this, we can compute that the minimal polynomial of

$$\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$$

is $(x - 1)(x - 2)$.

Exercise 13.20. If D is diagonal with entries λ_i , then $m_D = \prod (x - \lambda_i)$ where the product is taken over the distinct λ_i only, i.e., m_D has no repeated roots.

Exercise 13.21. λ is an eigenvalue of A if and only if $m_A(\lambda) = 0$.

Exercise 13.22. m_A divides f_A , and f_A divides $(m_A)^n$.

Exercise 13.23. If a matrix M is triangular with main diagonal entries λ_i , then $f(M)$ is triangular with main diagonal entries $f(\lambda_i)$ for any polynomial f .

Exercise 13.24. Let f be a polynomial, $S \in GL(n, F)$, and $A \in M_n(F)$. Then $f(S^{-1}AS) = S^{-1}f(A)S$.

Remember that we say that A and B are similar, $A \sim B$ if $B = S^{-1}AS$ for some matrix S . The previous exercise then says that if A and B are similar, so are $f(A)$ and $f(B)$ for any polynomial f .

Exercise 13.25. If $A \sim B$, then $m_A = m_B$.

A corollary is that if A is diagonalizable (that is, similar to a diagonal matrix), then m_A has no repeated roots. Now consider

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}.$$

The characteristic polynomial of this matrix is $(x - 1)^2$; since the minimal polynomial must divide the characteristic polynomial, the minimal polynomial must be one of 1 , $x - 1$, and $(x - 1)^2$. It is simple to check that it is not either of the first two, so the minimal polynomial must be $(x - 1)^2$; this polynomial has a repeated root, and thus is not diagonalizable. This leads us to the following theorem, to be proved later.

Theorem 13.26. *A is diagonalizable if and only if m_A has no repeated roots.*

Chapter 14

14th day, Friday 7/16/04 (Scribe: D. Jeremy Copeland)

14.1 Rank inequalities

Theorem 14.1. *If U and V are subspaces of W , then $\dim(\text{Span}(U \cup V)) \leq \dim(U) + \dim(V)$.*

Definition 14.2. If U and V are subspaces of W , then let $U + V = \{u + v \mid u \in U, v \in V\}$.

Exercise 14.3. If U and V are subspaces of W , then $\text{Span}(U \cup V) = U + V$

Exercise 14.4 (Modular identity). $\dim(U + V) + \dim(U \cap V) = \dim(U) + \dim(V)$.

Corollary 14.5 (Submodular inequality). $S, T \subset V$ (subsets), then

$$\text{rk}(S) + \text{rk}(T) \geq \text{rk}(S \cap T) + \text{rk}(S \cup T).$$

This says that “rank is submodular.”

Theorem 14.6. *If $A, B \in M_n(F)$, $\text{rk}(AB) \leq \min(\text{rk}(A), \text{rk}(B))$.*

Corollary 14.7. $\text{rk}(AA^T) \leq \text{rk}(A)$ over any field.

Corollary 14.8. $\text{rk}(AA^*) \leq \text{rk}(A)$ over \mathbb{C} .

Exercise 14.9. $\text{rk}(AA^*) = \text{rk}(A)$ over \mathbb{C} , thus $\text{rk}(AA^T) = \text{rk}(A)$ over \mathbb{R} .

Exercise 14.10. Give examples of other fields of all characteristics such that $\text{rk}(AA^T) \neq \text{rk}(A)$

14.2 Rings

Definition 14.11. A **ring** is a set $(R, +, \cdot)$ such that

- (A) $(R, +)$ is an abelian group,

- (B) $(R \setminus \{0\}, \cdot)$ is a group,
- (C1) $(a + b)c = ac + bc$,
- (C2) $a(b + c) = ab + ac$.

Exercise 14.12. $0 + 0 = 0$.

Definition 14.13. a is a **zero divisor** if $a \neq 0$ and $\exists b \neq 0$ such that $ab = 0$.

Example 14.14. \mathbb{Z} has no zero divisors.

Example 14.15. $\mathbb{Z}/m\mathbb{Z}$ has zero divisors $\iff m$ is composite.

Example 14.16. A field has no zero divisors.

Exercise 14.17. $\mathbb{F}[x]$, the ring of polynomials over a field \mathbb{F} , has no zero divisors.

Remark 14.18. $M_n(\mathbb{F})$ is a non-commutative ring with zero divisors. Let

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

Then $A^2 = 0$, so A is a zero divisor.

Exercise 14.19. $A \in M_n(\mathbb{F})$ is a zero divisor $\iff A$ is singular and non-zero.

Exercise 14.20. $a \in \mathbb{Z}/m\mathbb{Z}$ is a zero divisor $\iff a \neq 0$ and $\text{g.c.d.}(a, m) > 1$.

Definition 14.21. An **integral domain** is a commutative ring with identity and no zero divisors.

Example 14.22. \mathbb{Z} , any field \mathbb{F} , and $\mathbb{F}[x]$ are all integral domains.

14.3 Ideals

Definition 14.23. If $\mathcal{I} \subset R$ is a subset of a ring, then \mathcal{I} is an **ideal** ($\mathcal{I} \triangleleft R$) if

- A) \mathcal{I} is an additive subgroup.
- B) $(\forall a, b \in \mathcal{I})(\forall r \in R)(ar \in \mathcal{I})$.

Example 14.24. $d\mathbb{Z} \triangleleft \mathbb{Z}$, where $d\mathbb{Z} = \{dn \mid n \in \mathbb{Z}\}$.

Example 14.25. If $f \in \mathbb{F}[x]$, $f\mathbb{F}[x] = \{fg \mid g \in \mathbb{F}[x]\} \triangleleft \mathbb{F}[x]$.

Definition 14.26. If R is a commutative ring with identity, $u \in R$, then $(u) = \{ur \mid r \in R\}$ is called a **principal ideal**.

Exercise 14.27. Prove that this is an ideal.

Theorem 14.28. In \mathbb{Z} and in $\mathbb{F}[x]$, every ideal is a principal ideal.

Example 14.29. In $\mathbb{Z}[x]$, not every ideal is a principal ideal. Indeed, consider the set of polynomials with even constant term. This is an ideal which is not principal.

Lemma 14.30 (Division theorem). $(\forall f, g)(\exists q, r)(g = fq + r \text{ and } \deg(r) < \deg(f)$.

Definition 14.31. Suppose R is an integral domain, $f, g \in R$. $f | g$ if $\exists h \in R$ such that $fh = g$.

Definition 14.32. $d = \text{g.c.d.}(f, g)$ if

- $d | f$,
- $d | g$, and
- for every e such that $e | f$ and $e | g$, we have that $e | d$.

Example 14.33. Everything is a divisor of 0, so $\text{g.c.d.}(0, 0) = 0$.

Definition 14.34. Divisors of 1 are called **units**. $g | 1$ means that $(\exists h)(gh = 1)$.

Definition 14.35. R^\times is the set of units in R .

Exercise 14.36. R^\times is a multiplicative group.

Example 14.37. $\mathbb{F}[x]^\times = \mathbb{F}^\times$.

Example 14.38. $\mathbb{Z}[x]^\times = \{\pm 1\}$.

Definition 14.39. The **Gaussian integers** are the ring $\mathbb{Z}[i] = \{a + bi \mid a, b \in \mathbb{Z}\}$.

Exercise 14.40. $\mathbb{Z}[i]^\times = \{\pm 1, \pm i\}$.

Exercise 14.41. Every ideal in $\mathbb{Z}[i]$ is a principal ideal. *Hint:* invent a division theorem for Gaussian integers.

Definition 14.42. R is a **principal ideal domain** (PID) if R is an integral domain and all ideals of R are principal.

Example 14.43. \mathbb{Z} , $\mathbb{F}[x]$, and $\mathbb{Z}[x]$ are all principal ideal domains.

Exercise 14.44. In $\mathbb{Z}[\sqrt{-5}] = \{a + b\sqrt{-5} \mid a, b \in \mathbb{Z}\}$, the g.c.d. does not exist (that is not every pair of elements has a g.c.d.).

Theorem 14.45. If R is a PID, then gcd always exists and is a linear combination:

$$(\forall f, g \in R)(\exists d)d = \text{g.c.d.}(f, g) \text{ and } \exists r, s \in R \text{ such that } d = rf + sg.$$

Definition 14.46. $f \in \mathbb{F}[x]$ is **irreducible** if $f \neq 0$, f is not a unit, and whenever $f = gh$, either g or h is a unit.

14.4 Minimal polynomial

Henceforth, “irreducible polynomial” will mean irreducible over \mathbb{Q} , when no other field is indicated.

Definition 14.47. Assume that $\alpha \in \mathbb{C}$. Then $\mathcal{I}_\alpha = \{f \in \mathbb{Q}[x] \mid f(\alpha) = 0\} \triangleleft \mathbb{Q}[x]$. This is principal. Thus $\mathcal{I}_\alpha = (m_\alpha)$, and m_α is the **minimal polynomial** of α .

Corollary 14.48. $if \in \mathbb{Q}[x], f\alpha = 0 \iff m_\alpha \mid f_\alpha$.

Exercise 14.49. m_α is irreducible.

Exercise 14.50. If $f \in \mathbb{Q}[x], f\alpha = 0$, and f is irreducible, then $f = cm_\alpha$ for some unit c .

Exercise 14.51. $x^3 - 2$ is irreducible.

Corollary 14.52. $m_{\sqrt[3]{2}}(x) = x^3 - 2$.

Definition 14.53. α is **algebraic** if $m_\alpha \neq 0$. i.e. $\exists f \in \mathbb{Q}[x] \setminus \{0\}$ such that $f\alpha = 0$.

Definition 14.54. If α is not algebraic, then it is called **transcendental**.

Definition 14.55. If α is algebraic, then $\deg(\alpha) = \deg(m_\alpha)$.

Example 14.56. $\deg(\sqrt[3]{2}) = 3$.

Example 14.57. $q \in \mathbb{Q} \Rightarrow \deg(q) = 1$.

Exercise 14.58. $x^n - 2$ is irreducible.

Definition 14.59. $\mathbb{Q}[\alpha] = \{f(\alpha) \mid f \in \mathbb{Q}[\alpha]\}$.

Theorem 14.60. *If α is algebraic, $\deg(\alpha) = k \geq 1$, then*

$$\mathbb{Q}[\alpha] = \{a_0 + a_1\alpha^1 + \cdots + a_{k-1}\alpha^{k-1} \mid a_j \in \mathbb{Q}\}.$$

Exercise 14.61. $\dim_{\mathbb{Q}}(\mathbb{Q}[\alpha]) = \deg(\alpha)$.

Theorem 14.62. *If α is algebraic, then $\mathbb{Q}[\alpha]$ is a field.*

Theorem 14.63. *If \mathbb{F} is any field and α is algebraic over \mathbb{F} , then $\mathbb{F}[\alpha]$ is a field.*

Theorem 14.64. *The set of algebraic numbers is a field.*

In fact, we will prove that:

$$\deg(\alpha \pm \beta) \leq \deg(\alpha)\deg(\beta)$$

$$\deg(\alpha\beta) \leq \deg(\alpha)\deg(\beta)$$

$$\deg(\alpha/\beta) \leq \deg(\alpha)\deg(\beta)$$

Exercise 14.65. Find $m_{\sqrt{3}+\sqrt[3]{7}}$.

Theorem 14.66. If $K \subset L \subset M$ are fields, then $\dim_K(M) = \dim_L(M) \dim_K(L)$.

Exercise 14.67. If $K \subset L \subset M$ are fields, $\alpha_1, \dots, \alpha_r$ is a basis of L over K , and β_1, \dots, β_s is a basis of M over L , then $\{\alpha_j \beta_k\}$ is a basis of M over K .

Exercise 14.68. If $K \subset L$ are fields, and α is algebraic over K , then $\deg_L(\alpha) \leq \deg_K(\alpha)$.

Theorem 14.69. If $K \subset L$ are fields, and $\dim_K(L) = k < \infty$ (“ L is an *extension* of K of degree k ”), then every $\alpha \in L$ is algebraic over K and $\deg_K(\alpha) \mid k$.

Given an interval of unit length, can we construct with straightedge and compass an interval of length $\sqrt[3]{2}$? Galois showed that this is impossible. One observes that it is possible to construct the field \mathbb{Q} , and beyond \mathbb{Q} , any number constructed gives at most a degree two extension. Since to construct $\sqrt[3]{2}$, one would need to form an extension of degree a multiple of three, this number is not constructible.

Definition 14.70.

$$J(k, \lambda) = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \cdots & 0 \\ \vdots & & \ddots & \ddots & 0 \\ \vdots & & & \ddots & 1 \\ 0 & \cdots & 0 & \lambda & \end{pmatrix}$$

is the $k \times k$ **Jordan block with eigenvalue λ** .

Definition 14.71. A block diagonal matrix is a matrix

$$\text{diag}(A_1, \dots, A_t) = \begin{pmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_t \end{pmatrix}$$

such that the diagonal entries, A_j are all square matrices of varied sizes, $n_j \times n_j$. It has dimension $n \times n$, where $n = \sum n_j$.

Exercise 14.72. The characteristic polynomial has the form:

$$f_{\text{diag}(A_1, \dots, A_k)}(x) = \prod_{j=1}^k f_{A_j}(x).$$

Theorem 14.73. If $f \in \mathbb{F}[x]$, then $f(\text{diag}(A_1, \dots, A_k)) = \text{diag}(fA_1, \dots, fA_k)$.

Corollary 14.74. $m_{\text{diag}(A_1, \dots, A_k)} = \text{l.c.m.}(f_{A_1}, \dots, f_{A_k})$.

Theorem 14.75. $\mathcal{I}_A = \{f \in \mathbb{F}[x] \mid fA = 0\} \triangleleft \mathbb{F}[x]$. Thus $\mathcal{I}_A = (m_A)$.

Definition 14.76. A **Jordan normal form matrix** is a block diagonal matrix with each block a Jordan block.

Theorem 14.77. $f_{J(\lambda,k)}(x) = (x - \lambda)^k$.

Exercise 14.78. $m_{J(\lambda,k)}(x) = (x - \lambda)^k$.

Corollary 14.79. A Jordan block is diagonalizable if and only if $k = 1$.

Definition 14.80. A is **similar** to B ($A \sim B$) if $\exists S$ such that $S^{-1}AS = B$.

Exercise 14.81. Two Jordan normal form matrices are similar if and only if they consist of the same blocks (possibly in a different order).

Theorem 14.82. If \mathbb{F} is algebraically closed, then $\forall A \exists B$ such that B is Jordan normal form and $B \sim A$.

Observe that if $f(A) = 0$, and $B \sim A$, then

$$f(B) = f(S^{-1}AS) = S^{-1}f(A)S = 0.$$

Exercise 14.83. If $f \in \mathbb{F}[x] \setminus \{0\}$ and \mathbb{F} is algebraically closed, then for all n , the number of pairwise dissimilar solutions, $A \in M_n(\mathbb{F})$, to $fA = 0$ is finite (bounded by a function in n and $\deg(f)$).

Theorem 14.84. If $A, B \in M_n(K)$, $L \supset K$ is a field extension, and $A \sim B$ over L , then $A \sim B$ over K .

Corollary 14.85. We may drop the condition of algebraic closure in Exercise 14.83.