Linear Algebra

Organized Results

complied by

Sarthak[1]

November 2022

To my stars,
Giuseppe, and the Doctor...

[1] vijaysarthak@iitgn.ac.in

Contents

Chapter I

Matrices

1 Row echelon

August 27, 2022

Remark. *We'll take the matrix entries from a field.*

Lemma 1.1.

(i) Deleting rightmost column or a non-pivot column preserves row reduced echelon form.

(ii) A row reduced echelon matrix in which each column contains a pivot is of the form

$$\begin{bmatrix} I_n \\ 0 \end{bmatrix},$$

i.e., its diagonal entries are 1 *and rest are* 0*.*

(iii) Deleting corresponding columns preserves row equivalence.

Theorem 1.2. *The row reduced echelon form of a matrix is unique.*

2 LU decomposition

Lemma 2.1 (Triangular matrices)**.**

(i) Product of lower (respectively upper) triangular square matrices is lower (respectively upper) triangular.

(ii) *The inverse of a lower (respectively upper) triangular square is lower (respectively upper) triangular.*

(iii) *The diagonal entries of the product of lower (respectively upper) triangular square matrices is the product of their diagonal entries.*

Definition 2.2 (LU decomposition)**.** Expressing a square matrix as a product of a lower and respectively an upper triangular matrix is called an LU decomposition of it.

Proposition 2.3 (Uniqueness of LU decomposition)**.** *Let A be an invertible matrix with an LU decomposition. Then its LU decomposition in which all diagonal the entries of the upper triangular matrix are 1, is unique.*

Remark. *An invertible matrix needn't have an LU decomposition: Consider* $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$.

A non-invertible matrix can have more than one "standard" LU decompositions: Consider $\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$.

3 Determinants

Give a general formalism for (commutative) rings and prove all the things! Define $\det\colon \bigcup_{n\in\mathbb{N}} R^{n\times n} \to R$. *Also define* $R^{n\times n}$.

Definition 3.1 (Inversions)**.** Let $n \geq 2$ and $\sigma \in S_n$. Then (i, j), for $1 \leq i, j \leq n$ is called an inversion of σ iff

$$i < j \text{ and } \sigma(i) > \sigma(j).$$

Definition 3.2 (Odd or even permutations)**.** A permutation is said to be odd (respectively even) if it has odd (respectively even) number of inversions.

Lemma 3.3. *Let A be a finite set and $f\colon A \to A$ such that $f \circ f = \mathrm{id}$ and $f(a) \neq a$ for all $a \in A$. Then $|A|$ is even.*

Theorem 3.4. *A transposition changes the parity of permutation.*

Definition 3.5 (Determinant)**.** Let A be an $n \times n$ square matrix for $n \geq 1$. Then we define

$$\det A := \sum_{\sigma \in S_n} a_{1,\sigma(1)} \cdots a_{n,\sigma(n)}.$$

Corollary 3.6 (Immediate facts)**.**

(i) Determinant of of triangular matrices is the product of its diagonal entries.

(ii) We have

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

(iii) We have

$$\det(I) = 1.$$

(iv) If a row or a column of a matrix is zero, then its determinant is 0*.*

Theorem 3.7 (Determinant under elementary row operations)**.** *Let A be a square matrix. Then the following hold:*

(i) $i \to i + cj$ leaves determinant unchanged.

(ii) $i \leftrightarrow j$ negates the determinant for $i \neq j$.

(iii) $i \to ci$ scales the determinant by c.

Corollary 3.8 (Determinant of elementary matrices)**.** *The determinant of the elementary matrix corresponding to the row operations $i \to i + cj$, $i \leftrightarrow j$ (for $i \neq j$), $i \to ci$ are respectively* 1, -1, c*.*

Corollary 3.9. *For an elementary matrix E, we have*

$$\det(EA) = (\det E)(\det A).$$

Lemma 3.10. *The reduced row echelon form R of a square matrix A is either I or has the last row as zero. Further, A is invertible $\iff$ $R = I$.*

Theorem 3.11. *We have*

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

Theorem 3.12 (Characterizing invertibility)**.** *A square matrix A is invertible $\iff$* $\det A \neq 0$*.*

Corollary 3.13. *If A is invertible, then* $\det(A) \neq 0$*, and*

$$\det(A^{-1}) = (\det A)^{-1}.$$

Definition 3.14 (Determinant-like functions)**.** A function δ that assigns to each square matrix a scalar is called a determinant-like function iff]tfh:

(i) We have

$$\delta(I) = 1.$$

(ii) δ gets negated if two rows are interchanged.

(iii) δ is "linear" in the first row, *i.e.*, for any $1 \times n$ row vectors $r_1, \ldots, r_n, r_1'$ and scalars k, l, we have

$$\delta \begin{bmatrix} kr_1 + lr_1' \\ r_2 \\ \vdots \\ r_n \end{bmatrix} = k\,\delta \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_n \end{bmatrix} + l\,\delta \begin{bmatrix} r_1' \\ r_2 \\ \vdots \\ r_n \end{bmatrix}.$$

Theorem 3.15 (Characterizing det). *The determinant given by Definition 3.5 is the unique determinant-like function.*

Remark. *Having defined determinants, Now we are in a shape to analyze Atul's parity definition.*

Theorem 3.16 (Equivalence of the parity definitions for permutations). *Parity of the number of inversions in a permutation is the same as the parity of the number of the transpositions that it can be decomposed into.*

Chapter II

Vector spaces

1 Spaces and subspaces

October 11, 2022

Definition 1.1 (Vector spaces)**.** Let V be an additive abelian group and F be a field, along with a scalar multiplication operation $F \times V \to V$. Then V is called a vector space over F iff the following hold:

(i) $1v = v$.

(ii) $(ab)v = a(bv)$.

(iii) $(a + b)v = av + bv$, and $a(u + v) = au + av$.

Remark. *We have followed the usual convention that any "multiplicative" operation (here, the scalar multiplication) precedes over the "additive" operation (here, the vector addition).*

We'll call the elements of V as vectors, and the group operation of V as vector addition.

We'll often omit specifying F, and just call the elements of xF as scalars.

Example 1.2 (Matrices)**.** The set of $m \times n$ matrices with scalar entries, $F^{m\times n}$ for $m, n \geq 1$ forms a vector space over F with the usual operations.

Notation. *We'll sometimes denote $F^{m\times 1}$ by simply F^m.*

Proposition 1.3. *Let V be a vector space, $v_1, \ldots, v_n \in V$ and $a_1, \ldots, a_n$ be a scalars with $m, n \geq 0$. Then*

$$\Big(\sum_{i=1}^{m} a_i\Big)\Big(\sum_{j=1}^{n} v_j\Big) = \sum_{i=1}^{m}\sum_{j=1}^{n} a_i v_j.$$

Definition 1.4 (Subspaces)**.** Let V be a vector space over a field F. Then $W \subseteq V$ is called a subspace of V iff the operations of vector addition and scalar multiplication can be inherited to W such that W itself forms a vector space over F under these inherited operations.

Proposition 1.5 (Characterizing subspaces)**.** *Let V be a vector space and $W \subseteq V$. Then W is a subspace of $V \iff W \neq \emptyset$ and W is closed under vector addition and scalar multiplication.*

Proposition 1.6.

(i) Subspaces of a subspace are subspaces of the parent space.

(ii) Nonempty intersections of subspaces are subspaces.

(iii) If U, W are subspaces of a vector space V such that $W \subseteq U$, then W is a subspace of U.

2 Sums of subspaces

October 11, 2022

Definition 2.1 (Sums of subspaces)**.** Let V be a vector space and W_i's be subspaces of V for $i \in I$. Then we define

$$\sum_{i\in I} W_i := \Big\{\text{finite sums in } \bigcup_{i\in I} W_i\Big\}.$$

Proposition 2.2.

(i) Sums of subspaces are subspaces.

(ii) $\sum_i W_i$ is the smallest subspace containing $\cup_i W_i$.

Definition 2.3 (Sum of two subspaces)**.** Let U, W be subspaces of a vector space V. Then we define

$$U + W := \sum_{X\in\{U,W\}} X.$$

Proposition 2.4. *For subspaces U, W, X of a vector space V, we have*

$$\begin{aligned} U + W &= W + U, \\ (U + W) + X &= U + (W + X), \text{ and} \\ U + \{0\} &= U. \end{aligned}$$

Proposition 2.5 (No notational collision)**.** *The inductive definition of finite sums of subspaces via the binary operation $+$ gives the same subspace as the one given by Definition 2.1.*

Proposition 2.6 (Finite sums of subspaces)**.** *Let $n \geq 0$ and $W_1, \ldots, W_n$ be subspaces of a vector space V. Then*

$$W_1 + \cdots + W_n = \{w_1 + \cdots + w_n : w_i \in W_i\}.$$

Definition 2.7 (Direct sums)**.** Let V be a vector space and W_i's be subspaces of V for $i \in I$. Then V is called the direct sum of W_i's iff for each $v \in V$, there exists a unique $w \in \prod_{i \in I} W_i$ such that the set $J := \{i \in I : w_i \neq 0\}$ is finite, and

$$v = \sum_{i \in J} w_i.$$

We call this w as the *decomposition* of w in the direct sum.

Remark. *We'll denote this fact by*

$$V = \bigoplus_{i \in I} W_i.$$

Proposition 2.8 (Characterizing direct sums)**.** *Let W_i's be subspaces of a vector space V, for $i \in I$ such that $V = \sum_{i \in I} W_i$. Then the following are equivalent:*

(i) 0 vector admits a unique finite sum of nonzero vectors in $\bigcup_{i \in I} W_i$ (i.e., the empty sum).

(ii) $V = \bigoplus_{i \in I} W_i$.

Proposition 2.9 (Finite direct sums)**.** *Let $n \geq 0$ and $W_1, \ldots, W_n$ be subspaces of a vector space V such that $V = W_1 + \cdots + W_n$. Then the following are equivalent:*

(i) For each $v \in V$, there exists a unique $w \in \prod_{i=1}^{n} W_i$ such that $v = w_1 + \cdots + w_n$.

(ii) The above holds for $v = 0$.

(iii) $V = W_1 \oplus \cdots \oplus W_n$

Proposition 2.10 (Characterizing direct sums of two subspaces)**.** *Let U and W be subspaces of a vector space V such that $V = U + W$. Then the following are equivalent:*

(i) $U \cap W = \{0\}$.

(ii) $V = U \oplus W$.

3 Algebra to simplify the later work

October 15, 2022

Definition 3.1 (Simplifying notation for linear combinations)**.** Let V be a vector space, $v \in V^m$ and a be an $m \times n$ matrix of scalars, for $m, n \geq 0$. Then we define $va \in V^n$ as

$$(va)_j := \sum_{i=1}^{m} a_{i,j}\, v_i.$$

Proposition 3.2 (Algebra of V^n over matrices)**.** *Let V be a vector space and $n \geq 1$. Then V^n forms an abelian group under slot-wise addition, and for $v \in V^n$ and matrices a, b of scalars of appropriate sizes, the following hold:*

$$(va)b = v(ab)$$
$$(v + w)a = va + wa$$
$$v(a + b) = va + vb$$

4 Spans, independence, bases...

October 11, 2022

Definition 4.1 (Span)**.** Let V be a vector space and $S \subseteq V$. Then we define

$$\operatorname{span} S := \text{ smallest subspace of } V \text{ containing S.}$$

Corollary 4.2. *For subspaces W_i's of a vector space V, we have*

$$\sum\nolimits_i W_i = \operatorname{span}\Big(\bigcup\nolimits_i W_i\Big).$$

Definition 4.3 (Linear combinations)**.** Let V be a vector space and $S \subseteq V$. Then a linear combination of vectors in S is a vector of the form

$$a_1 v_1 + \cdots + a_n v_n$$

where $v_1, \ldots, v_n \in S$ for an $n \geq 0$ and $a_1, \ldots, a_n$ are scalars.

We'll denote this by

Proposition 4.4 (Characterizing spans)**.** *Let V be a vector space and $S \subseteq V$. Then*

$$\operatorname{span} S = \{\textit{linear combinations of vectors in } S\}.$$

Proposition 4.5 (Characterizing spans of finite sets)**.** *Let V be a vector space and $v \in V^n$ for $n \geq 0$. Then*

$$\operatorname{span}\big(\{v_1, \ldots, v_n\}\big) = \{va : a \in F^n\}.$$

Definition 4.6 (Independence)**.** Let V be a vector space. Then a set $L \subseteq V$ is called independent iff for any $v \in L^n$ for $n \geq 0$ with distinct v_i's, we have that

$$va = 0 \implies a = 0$$

for all $a \in F^n$.

Proposition 4.7 (Independence of finite sets)**.** *Let V be a vector space and $v \in V^n$ for $n \geq 0$ with distinct v_i's. Then the following are equivalent:*

(i) $\{v_1, \ldots, v_n\}$ *is independent.*

(ii) $va = 0 \implies a = 0$ *for any* $a \in F^n$.

Proposition 4.8.

(i) Any subset of an independent set is independent as well.

(ii) Independence in a subspace is the same as that in the parent space.

Definition 4.9 (Bases)**.** A subset B of a vector space V is called a basis iff it is independent and $\operatorname{span} B = V$.

Lemma 4.10. *Let L be an independent set in a vector space V and $v \in V \setminus \operatorname{span} L$. Then $L \cup \{v\}$ is independent too.*

Theorem 4.11 (Extending independent sets to bases[1])**.** *Let V be a vector space and $L, S \subseteq V$ such that L is independent and $\operatorname{span} S = V$ with $|S| < \infty$. Then there exists a subset $T \subseteq S$ such that $L \cup T$ is a basis for V.*

[1] For $|S| = \infty$, the same result can be proven using Zorn's lemma.

Definition 4.12 (Finite-dimensional vector spaces)**.** A vector space is called finite-dimensional iff it can be spanned by some finite subset of it.

Corollary 4.13 (Existence of bases)**.** *Every finite-dimensional vector space has a basis.*

Theorem 4.14 (Independence, span and cardinality)**.** *Let V be a vector space and $L, S \subseteq V$ such that L is independent and* $\operatorname{span} S = V$ *with $|S| < \infty$. Then*

$$|L| \leq |S|.$$

Corollary 4.15 (Dimension of finite-dimensional vector spaces)**.** *Let V be a finite-dimensional vector space. Then there exists a unique natural* $\dim V \geq 0$ *such that any basis for V has* $\dim V$ *number of vectors.*

Corollary 4.16. *Let V be a finite-dimensional vector space and $L, S \subseteq V$ such that L is independent and* $\operatorname{span} S = V$. *Then the following hold:*

(i) $|L| \leq \dim V \leq |S|$.

(ii) $|L| = \dim V \implies L$ *is a basis.*

(iii) $|S| = \dim V \implies S$ *is a basis.*

Proposition 4.17 (Dimension of subspaces)**.** *Let W be a subspace of a finite-dimensional vector space V. Then the following hold:*

(i) W is finite-dimensional.

(ii) $\dim W \leq \dim V$.

(iii) $\dim W = \dim V \iff W = V$.

Proposition 4.18. *Let $W_1, \ldots, W_n$ be finite-dimensional subspaces of a vector space V for $n \geq 0$. Then $\sum_{i=1}^{n} W_i$ is finite-dimensional too.*

Proposition 4.19 (Dimension of sum of two finite-dimensional subspaces)**.** *Let U, W be subspaces of a finite-dimensional vector space V such that $V = U + W$. Then*

$$\dim V = \dim U + \dim W - \dim U \cap W.$$

Proposition 4.20 (Dimension of finite sum of finite-dimensional subspaces)**.** *Let $W_1, \ldots, W_n$ be subspaces of a finite-dimensional vector space V for $n \geq 0$ such that $V = W_1 + \cdots + W_n$. Then the following hold:*

(i) $\dim V \leq \dim W_1 + \cdots + \dim W_n$.

(ii) $\dim V = \dim W_1 + \cdots + \dim W_n \iff V = W_1 \oplus \cdots \oplus W_n$.

Corollary 4.21. *Let U, V be subspaces of a finite-dimensional vector space V such that $U \cap V = \{0\}$. Then* $\dim V = \dim U + \dim W \iff V = U \oplus W$.

5 Subspaces associated with a matrix

October 16, 2022

Remark. *Here, the matrices are over F.*

Definition 5.1 (Row, column and null spaces and their dimensions)**.** Let A by an $m \times n$ matrix. Then we define the following spaces:

$$\begin{aligned} \operatorname{row}(A) &:= \operatorname{span}\{\text{rows}\} &&\subseteq F^{1\times n} \\ \operatorname{col}(A) &:= \operatorname{span}\{\text{columns}\} &&\subseteq F^{m\times 1} \\ \operatorname{null}(A) &:= \{X : AX = 0\} &&\subseteq F^{n\times 1} \end{aligned}$$

We further define

$$\begin{aligned} \text{row rank} &:= \dim \operatorname{row}(A), \\ \text{column rank} &:= \dim \operatorname{col}(A), \text{ and} \\ \text{nullity} &:= \dim \operatorname{null}(A). \end{aligned}$$

Proposition 5.2. *For a square matrix A of size n, the following are equivalent:*

(i) Row rank is n

(ii) A is invertible.

(iii) Column rank is n.

Lemma 5.3. *For matrices $A = BC$, we have that*

$$\begin{aligned} \operatorname{row}(A) &\subseteq \operatorname{row}(C), \textit{ and} \\ \operatorname{col}(A) &\subseteq \operatorname{col}(B). \end{aligned}$$

Theorem 5.4. *For any matrix, we have*

$$\textit{row rank} = \textit{column rank.}$$

Remark. *This allows to talk of the "rank" of matrices.*

Also, prove the above in two ways: first by Gauss elimination, and second by using the above lemma.

Corollary 5.5 (Somme immediate consequences)**.**

(i) Rank of a matrix is bounded by the its number of rows and columns.

(ii) Rank of AB is bounded by those of A and B.

(iii) If $A_{m\times n}B_{n\times m} = I_m$, then $m \leq n$.

Proposition 5.6. *Row operations on matrices preserve the span of rows and the independence of columns.*

Proposition 5.7. *Complex conjugation of a complex matrix preserves the independence of rows and columns.*

Corollary 5.8 (Rank preserving operations)**.** *The following operations on a matrix preserve its rank:*

(i) Row operations.

(ii) Transpositions.

(iii) Complex conjugation for complex matrices.

Lemma 5.9. *For the linear map $F^n \to F^m$ given by $X \mapsto AX$ for $A \in F^{m\times n}$, we have that*

$$\ker = \operatorname{null}(A), \textit{ and}$$
$$\operatorname{im} = \operatorname{col}(A).$$

Theorem 5.10 (Rank-nullity)**.** *For any matrix, we have*

$$\textit{rank} + \textit{nullity} = \#(\textit{columns}).$$

Chapter III

Linear maps

Remark. *Again, we'll fix a field F, and call its elements, scalars.*

1 Basics

October 15, 2022

Definition 1.1 (Linear maps and isomorphisms)**.** Let V, W be vector spaces over a common field. Then a function $\phi\colon V \to W$ is called a linear maps iff

(i) $\phi(u + v) = \phi(u) + \phi(v)$, and

(ii) $\phi(au) = a\,\phi(u)$.

If ϕ is a bijection too, then we call it a *(linear) isomorphism*, and call V and W, *isomorphic*.

Notation. *We'll write "$T\colon V \to W$ is linear" to mean "V, W are vector spaces over a common field and $T\colon V \to W$ is a linear map".*

We'll also, for a linear map $T\colon V \to W$, write Tv for $T(v)$.

Example 1.2 (Matrix operations)**.** Let $m, n, k \geq 1$.

(i) An $m \times n$ matrix A induces a linear map $F^{n\times k} \to F^{m\times k}$ given by $X \mapsto AX$. (Similarly, another map due to right-multiplication is also induced.)

(ii) Matrix transposition $X \mapsto X^t$ gives another linear map $F^{m\times n} \to F^{n\times m}$.

Proposition 1.3 (Properties of linear maps)**.**

(i) Composition of linear maps (respectively isomorphisms) is a linear map (respectively an isomorphism).

(ii) A linear map is injective $\iff$ its kernel is $\{0\}$.

(iii) The kernel of a linear map is a subspace of the domain space.

(iv) Restriction of a linear map to a subspace of the domain space is linear.

(v) Inclusion map from a subspace is linear.

(vi) Inverse of an isomorphism is linear too.

(vii) "Being isomorphic" is an equivalence relation.

Proposition 1.4 (Properties preserved by isomorphisms)**.** *Let $T\colon V \to W$ be an isomorphism and $S \subseteq V$. Then the following hold:*

(i) S is independent in V $\iff$ $T(S)$ is independent in W.

(ii) $\operatorname{span}(S) = V \iff \operatorname{span}(T(S)) = W$.

(iii) S is a basis of V $\iff$ $T(S)$ is a basis of W.

(iv) V (or equivalently, W) is finite-dimensional $\implies$ W (and equivalently V) is finite-dimensional and $\dim V = \dim W$.

Proposition 1.5 (Algebra of linear maps)**.** *Let V, W be vector spaces over a common field F. Then the set $\mathcal{L}(V,W)$ of linear maps $V \to W$, forms a vector space over F under the following operations:*

$$(T+S)(v) := Tv + Sv$$
$$(aT)(v) := a(Tv)$$

Further, if $V = W$, then we can also define the products

$$TS := T \circ S.$$

This makes $\mathcal{L}(V,V)$ into an associative F-algebra[1] *with identity id_V and the corresponding homomorphism is given by*

$$a \mapsto a\,\mathrm{id}_V.$$

Result 1.6 (Projections on component spaces in direct sums)**.** Let V be a vector space and U, W be subspaces such that $V = U \oplus W$. For each $v \in V$, define $\mathcal{P}_U v \in U$ and $\mathcal{P}_W v \in W$ so that

$$\mathcal{P}_U v + \mathcal{P}_W v = v.$$

Then the following hold:

[1] See §4.1.

(i) $\mathcal{P}_U, \mathcal{P}_W : V \to V$ are linear.
(ii) $\mathcal{P}_U^2 = \mathcal{P}_U$ and $\mathcal{P}_W^2 = \mathcal{P}_W$.
(iii) $\mathcal{P}_U + \mathcal{P}_W = \mathrm{id}_V$.

Remark. *Note that P_U depends not just on U, but the entire direct product decomposition of V, i.e., on U as well as W.*

Result 1.7 (Characterizing such projections)**.** Let $T: V \to V$ be linear with $T^2 = T$. Then
(i) $V = \operatorname{im} T \oplus \ker T$.
(ii) T is precisely the projection on $\operatorname{im} T$ in the direct sum $V = \operatorname{im} T \oplus \ker T$.

Theorem 1.8 (A linear map is uniquely determined by its action of basis)**.** *Let V, W be vector spaces over a common field, B be a basis of V and $f: B \to W$ be any function. Then there exists a unique linear map $T: V \to W$ such that for all $u \in B$, we have*

$$Tu = f(u).$$

Theorem 1.9 (Fundamental theorem of linear maps)**.** *For a linear map T from a finite-dimensional domain space V, we have that $\operatorname{im} T$ is finite-dimensional, and*

$$\dim V = \dim(\ker T) + \dim(\operatorname{im} T).$$

Corollary 1.10. *Let $T: V \to W$ be linear with V, W being finite-dimensional. Then the following hold:*
(i) T *is surjective* $\implies \dim V \geq \dim W$.
(ii) T *is injective* $\implies \dim V \leq \dim W$.
(iii) If $\dim V = \dim W$, *then* T *is surjective* $\iff T$ *is injective.*

2 Making linear maps act on tuples of vectors

October 15, 2022

Remark. *The elements of V^n should be viewed as n-tuples. For us, tuples and matrices are different things.*

This collides with the earlier notation of F^n which contained column vectors. Thus clarification will be needed when not clear from context.

Definition 2.1 (T acting on V^n)**.** Let $T\colon V \to W$ be a linear map and $v \in V^n$ for $n \geq 0$. Then we define $Tv \in W^n$ so that

$$(Tv)_i := Tv_i.$$

Proposition 2.2 (T's action V^n is "linear")**.** *Let $T\colon V \to W$ be a linear map, $v, w \in V^m$ and a be an $m \times n$ matrix of scalars for $m, n \geq 0$. Then*

$$T(v + w) = Tv + Tw$$
$$T(va) = (Tv)a$$

3 Studying matrices of linear maps

October 15, 2022

Definition 3.1 (Ordered bases for finite-dimensional spaces)**.** Let V be a finite-dimensional vector space. Then an n-tuple of distinct vectors $(u_1, \ldots, u_n)$ for $n \geq 0$ is called an ordered basis for V iff $\{u_1, \ldots, u_n\}$ is a basis for V.

Theorem 3.2 (Finite-dimensional spaces isomorphic to F^n's)**.** *Let V be a finite-dimensional vector space and B be an ordered basis. Then for each $v \in V$, there exists a unique column vector $[v]_B \in F^n$ so that*

$$v = B[v]_B.$$

The map $V \to F^n$ given by $v \mapsto [v]_B$ is an isomorphism.

Corollary 3.3. *Any finite-dimensional vector space V is isomorphic to $F^{\dim V \times 1}$.*

Notation. *Unless stated otherwise, the vectors e_i's will be reserved for the standard basis of F^n (for the specified n's).*

Theorem 3.4. *Any linear map $T\colon F^n \to F^m$ for $m, n \geq 1$ is due to the left-multiplication by a unique $m \times n$ matrix $[T]$, which is given by*

$$[T] := \begin{bmatrix} | & & | \\ Te_1 & \cdots & Te_n \\ | & & | \end{bmatrix}.$$

Remark. *It'll turn out (in Theorem 3.5) that this* $[T]$ *is precisely* $[T]_{C\leftarrow B}$ *where* C *and* B *are the standard ordered bases for* F^m *and* F^n *respectively.*

Theorem 3.5 (Matrices of linear maps)**.** *Let* $T\colon V \to W$ *be a linear map with* V*,* W *finite-dimensional. Let* B *and* C *respectively be ordered bases for* V *and* W*. Then there exists a unique* $\dim W \times \dim V$ *matrix* $[T]_{C\leftarrow B}$ *such that*

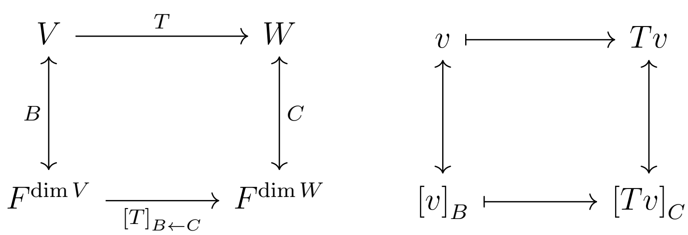
$$[Tv]_C = [T]_{C\leftarrow B}\,[v]_B.$$

This is given by

$$[T]_{C\leftarrow B} = \begin{bmatrix} | & & | \\ [Tu_1]_C & \cdots & [Tu_n]_C \\ | & & | \end{bmatrix},$$

where $(u_1, \ldots, u_n) = B$*.*

We have the following commutative diagrams:



Further, the map $\mathcal{L}(V, W) \to F^{\dim W \times \dim V}$ *given by*

$$T \mapsto [T]_{C\leftarrow B}$$

is an isomorphism.

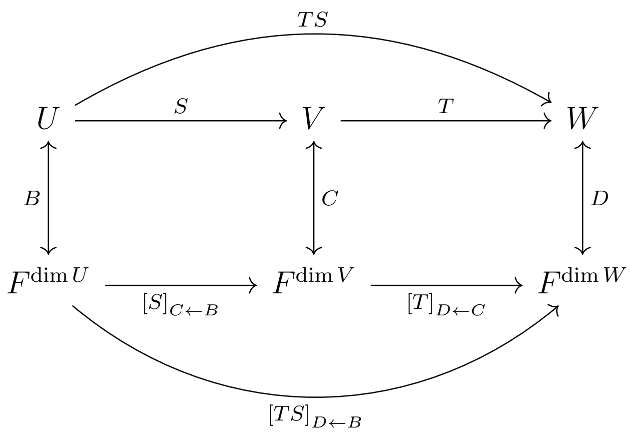
Remark. *Strictly speaking, we must require* $\dim V, \dim W \geq 1$ *since we normally don't have* $0 \times n$ *or* $m \times 0$ *matrices. We'll not care in the future to make this remark.*

Notation. *For composable linear maps* S*,* T*, we'll denote* $T \circ S$ *by* TS*.*

Theorem 3.6 (Matrix of compositions)**.** *Let* $S\colon U \to V$ *and* $T\colon V \to W$ *be linear maps with* U*,* V*,* W *being finite-dimensional. Let* B*,* C*,* D *be their respective ordered bases. Then*

$$[TS]_{D\leftarrow B} = [T]_{D\leftarrow C}\,[S]_{C\leftarrow B}.$$

We have the following commutative diagram:



Remark. *This says that the map on $\mathcal{L}(V, V) \to F^{n \times n}$ given by $T \mapsto [T]_{B \leftarrow B}$ in Theorem 3.5 is an algebra isomorphism!*[2]

Corollary 3.7 (Change of basis)**.** *Let V be a finite-dimensional vector space and B, C be ordered bases of V. Then $[\mathrm{id}_V]_{C \leftarrow B}$ is such that*

$$[v]_C = [\mathrm{id}_V]_{C \leftarrow B} [v]_B.$$

This is given by

$$[\mathrm{id}_V]_{C \leftarrow B} \begin{bmatrix} | & & | \\ [u_1]_C & \cdots & [u_n]_C \\ | & & | \end{bmatrix}$$

where $(u_1, \ldots, u_n) = B$.

Further, we have that

$$[\mathrm{id}_V]_{B \leftarrow C} [\mathrm{id}_V]_{C \leftarrow B} = I.$$

Corollary 3.8 (Change of bases in maps)**.** *Let $T\colon V \to W$ be a linear map and V, W be finite-dimensional vector spaces. Let B, B' be ordered bases of V and C, C' be those of W. Then*

$$[T]_{C' \leftarrow B'} = [\mathrm{id}_W]_{C' \leftarrow C} [T]_{C \leftarrow B} [\mathrm{id}_V]_{B \leftarrow B'}.$$

[2] See §4.1. Also see Proposition 4.36 that uses this.

Chapter IV

Inner product spaces

Remark. *In this chapter, our field F would be either $\mathbb{R}$ or $\mathbb{C}$. We'll denote this fact by changing our notation to $\mathbb{K}$.*

Remark. *When writing $A \in \mathbb{K}^{m\times n}$, we'll omit saying "$m, n \geq 1$".*

1 Basics

October 15, 2022

Definition 1.1 (Inner product spaces)**.** A vector space V over $\mathbb{K}$ along with a function (called the inner product) $\langle \cdot, \cdot \rangle \colon V \times V \to \mathbb{K}$ is called an inner product space iff the following hold:

(i) $\langle v, v\rangle \geq 0$.

(ii) $\langle v, v\rangle = 0 \iff v = 0$.

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

(iv) $\langle u + v, w\rangle = \langle u, w\rangle + \langle v, w\rangle$.

(v) $\langle av, w\rangle = a\langle v, w\rangle$.

We'll also define

$$\|v\| := \sqrt{\langle v, v\rangle}.$$

Proposition 1.2 (Easy identities)**.** *Let V be an inner product space. Then the* *following hold:*

$$
\begin{aligned}
\langle u, v+w\rangle &= \langle u, v\rangle + \langle u, w\rangle \\
\langle u, av\rangle &= \overline{a}\langle u, v\rangle \\
\|av\| &= |a|\,\|v\| \\
\|x+y\|^2 + \|x-y\|^2 &= 2\big(\|x\|^2 + \|y\|^2\big) \\
\langle x, y\rangle &= \begin{cases} \frac{\|x+y\|^2-\|x-y\|^2}{4} + i\,\frac{\|x+iy\|^2-\|x-iy\|^2}{4}, & \mathbb{K}=\mathbb{C} \\ \frac{\|x+y\|^2-\|x-y\|^2}{4}, & \mathbb{K}=\mathbb{R} \end{cases}
\end{aligned}
$$

Theorem 1.3 (Cauchy-Schwarz inequality)**.** *Let V be an inner product space and $u, v \in \mathbb{V}$. Then*

$$|\langle u, v\rangle| \le \|u\|\,\|v\|.$$

with equality holding if and only if $\{u, v\}$ is dependent.

Proposition 1.4 (Triangle inequality)**.** *Let V be an inner product space and $u, v \in V$. Then*

$$\|u+v\| \le \|u\| + \|v\|$$

with equality holding if and only if $\langle u, v\rangle = \|u\|\,\|v\|$.

Proposition 1.5 (Matrix representation of inner product)**.** *Let V be a finite-dimensional inner product space with an ordered basis $B := (u_1, \dots, u_n)$ for $n \ge 1$. Then we have*

$$\langle v, w\rangle = [v]_B^t\, A\, \overline{[w]_B}$$

where $A \in \mathbb{K}^{n\times n}$ is given by

$$A := \begin{bmatrix} \langle u_1, u_1\rangle & \cdots & \langle u_1, u_n\rangle \\ \vdots & & \vdots \\ \langle u_n, u_1\rangle & \cdots & \langle u_n, u_n\rangle \end{bmatrix}.$$

Remark. *If B is orthonormal (see Definition 2.1), then this reduces to $\langle u, v\rangle = [v]_B^t\,\overline{[w]_B}$.*

1.1 The Euclidean inner product on $\mathbb{K}^{m\times n}$

October 17, 2022

Definition 1.6 (Complex conjugation of matrices)**.** Let $A \in \mathbb{K}^{m\times n}$. Then we define $\overline{A} \in \mathbb{K}^{m\times n}$ as

$$(\overline{A})_{i,j} := \overline{A_{i,j}}.$$

Remark. *Hence, we are also defining complex conjugation for real matrices, for which this will leave the matrix unchanged.*

Proposition 1.7 (Properties of complex conjugation)**.** *Let A, B be matrices over $\mathbb{K}$ and $\lambda \in \mathbb{K}$. Then whenever defined, the following hold:*

$$\begin{aligned} \overline{A+B} &= \overline{A} + \overline{B} \\ \overline{\lambda A} &= \overline{\lambda}\,\overline{A} \\ \overline{AB} &= \overline{A}\,\overline{B} \\ (\overline{A})^t &= \overline{A^t} \\ (\overline{A})^{-1} &= \overline{A^{-1}} \quad \text{if } A \text{ is invertible} \\ \det(\overline{A}) &= \overline{\det A} \qquad \text{if } A \text{ is square} \end{aligned}$$

Notation. *We denote $(\overline{A})^t$ by A^*.*

Definition 1.8 (Trace)**.** For a square matrix A over any field, we define

$$\operatorname{tr} A := \sum_i A_{i,i}.$$

Proposition 1.9. *Let A, B be matrices over any field. Then whenever defined, the following hold:*

$$\begin{aligned} \operatorname{tr}(A+B) &= \operatorname{tr} A + \operatorname{tr} B \\ \operatorname{tr}(\lambda A) &= \lambda(\operatorname{tr} A) \\ \operatorname{tr}(AB) &= \operatorname{tr}(BA) \end{aligned}$$

Proposition 1.10 (An inner product on $\mathbb{K}^{m\times n}$)**.** *On $\mathbb{K}^{m\times n}$ over $\mathbb{K}$, we can define an inner product as*

$$\langle A, B\rangle := \operatorname{tr}(A^T\,\overline{B}).$$

It follows that, whenever defined, the following equalities hold across the appropriate spaces:

$$\begin{aligned}\langle A^t, B^t\rangle &= \langle A, B\rangle\\ \langle \overline{A}, \overline{B}\rangle &= \overline{\langle A, B\rangle}\end{aligned}$$

Example 1.11 (An inner product on $C[0,1]$)**.** On $C[0,1]$, the space of continuous $\mathbb{K}$-valued functions on interval $[0,1]$,

$$\langle f, g\rangle := \int_0^1 g(t)\,\overline{f(t)}\,\mathrm{d}t$$

defines an inner product.

2 Orthogonality

October 17, 2022

Definition 2.1 (Orthogonal and orthonormal sets)**.** Let V be an inner product space. Then an $L \subseteq V$ is called *orthogonal* iff for each $u, v \in L$, we have

$$\langle u, v\rangle = 0 \text{ whenever } u \neq v.$$

If we further have

$$\|v\| = 1$$

for each $v \in L$, we call L *orthonormal.*

For $M, N \subseteq V$, we also say that M *is orthogonal to* N, written $M \perp N$, iff

$$u \in M \text{ and } v \in N \implies \langle u, v\rangle = 0.$$

Corollary 2.2 (Preservation of orthonormality)**.** *The following preserve the orthonormality and orthogonality of a set in an inner product space:*

(i) Taking subsets.

(ii) Scaling vectors by scalars of absolute value 1*.*

Proposition 2.3. *Orthogonal set of nonzero vectors is independent.*

Proposition 2.4 (Expansion in orthogonal bases)**.** *Let V be a finite-dimensional inner product space and B be an ordered orthogonal basis. Then for any $v \in V$, we have*

$$[v]_B = \left(\frac{\langle v, u_1 \rangle}{\|u_1\|^2}, \ldots, \frac{\langle v, u_n \rangle}{\|u_n\|^2} \right)$$

where $(u_1, \ldots, u_n) = B$.

Proposition 2.5 (Pythagoras)**.** *Let V be an inner product space and $v_1, \ldots, v_n \in V$ be distinct and orthogonal for $n \geq 0$. Then*

$$\|v_1 + \cdots + v_n\|^2 = \|v_1\|^2 + \cdots + \|v_n\|^2.$$

Theorem 2.6 (Gram-Schmidt[1])**.** *Let V be an inner product space and $v_1, \ldots, v_n$ be distinct independent vectors for $n \geq 1$. Define $e_1, \ldots, e_n$ as*

$$\begin{aligned} e_1 &:= \frac{v_1}{\|v_1\|}, \textit{ and} \\ e_{i+1} &:= \frac{v_{i+1} - \sum_{j=1}^{i} \langle v_{i+1}, e_j \rangle e_j}{\|v_{i+1} - \sum_{j=1}^{i} \langle v_{i+1}, e_j \rangle e_j\|} \quad \textit{for } 1 \leq i < n. \end{aligned}$$

Then $e_1, \ldots, e_n$ so obtained are orthonormal such that for each $1 \leq i \leq n$, we have

$$\operatorname{span}\big(\{e_1, \ldots, e_i\}\big) = \operatorname{span}\big(\{v_1, \ldots, v_i\}\big).$$

Corollary 2.7. *Every orthonormal set in a finite-dimensional inner product space can be extended to an orthonormal basis.*[2]

3 Orthogonal complements

October 16, 2022

Proposition 3.1 (Orthogonal complements)**.** *Let V be an inner product space and W be a subspace. Then*

$$W^\perp := \{v \in V : \langle v, w \rangle = 0 \textit{ for all } w \in W\}$$

is a subspace of W such that

[1] Countable bases can be orthonormalized using this too.
[2] Complete infinite-dimensional inner product spaces have no orthonormal basis.

(i) $W \cap W^\perp = \{0\}$, *and*

(ii) W *if finite-dimensional* $\implies V = W \oplus W^\perp$.

Remark. *This allows to talk of "orthogonal projections on a subspace W" whenever $V = W \oplus W^\perp$.*

Proposition 3.2 ($(W^\perp)^\perp$ and W)**.** *Let U, W be subspaces of an inner product space V. Then the following hold:*

(i) $U \subseteq W \implies W^\perp \subseteq U^\perp$.[3]

(ii) $U \subseteq (U^\perp)^\perp$.[4]

(iii) $V = U \oplus U^\perp \implies U = (U^\perp)^\perp$.[5]

(iv) $V = U + W$ *and* $U \perp W \implies U^\perp = W$ *and* $W^\perp = U$.

Proposition 3.3. *Let V be an inner product space and U, W be orthogonal subspaces.*

Proposition 3.4 (Orthogonal projections)**.** *Let V be an inner product space and W be a subspace such that $V = W \oplus W^\perp$. Let $P_W : V \to V$ be the orthogonal projection onto W. Then*

$$\langle P_W u, v \rangle = \langle u, P_W v \rangle.$$

Proposition 3.5 (Characterizing orthogonal projections)**.** *Let V be an inner product space and $T : V \to V$ be linear with such that $T^2 = T$ and $\langle Tu, v \rangle = \langle u, Tv \rangle$. Then in addition to the conclusion of Result 1.7, we also have*

$$(\operatorname{im} T)^\perp = \ker T \text{ and } (\ker T)^\perp = \operatorname{im} T$$

so that T is the orthogonal projection onto $\operatorname{im} T$.

Proposition 3.6 (Orthogonal projections via orthogonal bases)**.** *Let V be an inner product space and W be a finite-dimensional subspace with an ordered orthogonal basis $(u_1, \dots, u_n)$ for $n \geq 0$. Then the orthogonal projection $P_W : V \to V$ onto W is given by*

$$P_W v = \sum_{i=1}^{n} \frac{\langle v, u_i \rangle}{\|u_i\|^2} u_i.$$

---

[3] For a counterexample for converse, consider $V := \ell^2$ and $U := \{x \in \ell^2 :$ only finitely many x_i's are nonzero$\}$.

[4] Same example to show proper inclusion.

[5] For a counterexample to the converse, see this.

Lemma 3.7. *Let V be a vector space and U, W be subspaces such that $V = U \oplus W$. Let $\mathcal{P}_U \colon V \to V$ be the projection onto U in the direct sum. Let $u \in U$ and $v \in V$. Then*

$$u = \mathcal{P}_U v \iff v - u \in W.$$

Theorem 3.8 (Orthogonal projections as approximations)**.** *Let V be an inner product space and W be a subspace such that $V = W \oplus W^\perp$. Let $v_0 \in V$ and $w \in W$ and $P_W \colon V \to V$ be the orthogonal projection onto W. Then*

$$\|w - v_0\| \geq \|P_W v_0 - v_0\|$$

with equality holding if and only if $w = P_W v_0$, or equivalently, $w - v_0 \in W^\perp$.

4 Applying to the matrices over $\mathbb{K}$

October 17, 2022

Lemma 4.1. *Let $A \in \mathbb{K}^{m\times n}$. Then under the Euclidean inner product in $\mathbb{K}^{n\times 1}$, we have*

$$\mathrm{col}(A^*) \perp \mathrm{null}(A).$$

Remark. *A^* is the usual complex conjugate of the complex matrix A.*

Theorem 4.2. *Let $A \in \mathbb{K}^{m\times n}$. Then in $\mathbb{K}^{n\times 1}$ with the Euclidean inner product, the following hold:*

$$\mathbb{K}^{n\times 1} = \mathrm{col}(A^*) \oplus \mathrm{null}\, A$$
$$\mathrm{col}(A^*) = (\mathrm{null}\, A)^\perp$$

Remark. $\mathrm{col}(A^*)$ *is just (the transpose of)* $\mathrm{row}(\overline{A})$.

Theorem 4.3 (Least squares in $\mathbb{K}^n$)**.** *Let $A \in \mathbb{K}^{m\times n}$, and $x \in \mathbb{K}^{n\times 1}$ and $b \in \mathbb{K}^{m\times 1}$. Let $P_{\mathrm{col}(A)} \colon \mathbb{K}^{m\times 1} \to \mathbb{K}^{m\times 1}$ be the orthogonal projection onto $\mathrm{col}(A)$. Then the following are equivalent:*

(i) $\|Ax - b\| = \|P_{\mathrm{col}(A)} b - b\|$.

(ii) $Ax = P_{\mathrm{col}(A)} b$.

(iii) $A^* A x = A^* b$.

Corollary 4.4 (On A^*A)**.** *For $A \in \mathbb{K}^{m\times n}$, the following hold:*

(i) $A^*Ax = 0 \iff Ax = 0$.

(ii) A's columns are independent $\iff$ A^*A is invertible.

Corollary 4.5 (Orthogonal projections in $\mathbb{K}^n$)**.** *Let $A \in \mathbb{K}^{m\times n}$ with independent columns. Then the orthogonal projection $P_{\mathrm{col}(A)} \colon \mathbb{K}^{m\times 1} \to \mathbb{K}^{m\times 1}$ onto $\mathrm{col}(A)$ is given by*

$$P_{\mathrm{col}(A)}\, b = A(A^*A)^{-1}A^*\, b.$$

5 Orthogonal matrices

October 17, 2022

Definition 5.1 (Orthogonal matrices)**.** An square matrix A over $\mathbb{K}$ is called orthogonal iff

$$A^t\,\overline{A} = I.$$

Remark. *We could have equivalently demanded $A^*A = I$.*

Remark. *"Unitary" vs "orthogonal": Unitary is stuck because $U^*U = I$ is like an extension of complex units.*

Remark. *(This remark is imprecise!) Given a linear map $T \colon V \to W$ where V, W are inner product spaces, it's possible to define an "adjoint operator" T^* so that $[T^*] = [T]^*$. This allows to define "orthogonal operators".*

It's possible to generalize these even further so that we don't require V and W to be inner product spaces.

Proposition 5.2 (Properties of orthogonal matrices)**.** *Let $A \in \mathbb{K}^{n\times n}$ be orthogonal. Then*

$$|\det A| = 1$$

rendering A invertible with

$$A^{-1} = A^*.$$

Further, A^t, $\overline{A}$, A^* are orthogonal too.

Also, product of orthogonals is orthogonal.

Proposition 5.3 (Characterizing orthogonal matrices)**.** *Let $A \in \mathbb{K}^{n\times n}$. Then the following are equivalent:*

(i) A is orthogonal.

(ii) Rows of A are orthonormal.

(iii) Columns of A are orthonormal.

(iv) A preserves $\|\cdot\|$.

(v) A preserves $\langle\cdot,\cdot\rangle$.

Remark. *The eigenvalues of an orthogonal matrix have absolute value 1. (Result 2.7.)*

Result 5.4. Let V be a finite-dimensional inner product space and E, F be ordered orthonormal bases. Then $[\mathrm{id}_V]_{F\leftarrow E}$ is orthogonal.

Lemma 5.5. *An upper-triangular matrix orthogonal matrix with positive diagonal entries is necessarily I.*

Theorem 5.6 (QR decomposition)**.** *Let $A \in \mathbb{K}^{m\times n}$ with independent columns. Then there exist unique $Q \in \mathbb{K}^{m\times n}$ and $R \in \mathbb{K}^{n\times n}$ such that*

(i) $A = QR$,

(ii) Q's columns are orthonormal, i.e., $Q^*Q = I_n$, and

(iii) R is upper-triangular with positive diagonal entries.

Further, this R is invertible, and if $A = [a_1, \ldots, a_n]$ and $Q = [q_1, \ldots, q_n]$, then R is given by

$$R = \begin{bmatrix} \langle a_1, q_1\rangle & \cdots & \langle a_n, q_1\rangle \\ & \ddots & \vdots \\ & & \langle a_n, q_n\rangle \end{bmatrix}.$$

Chapter V

Eigenvalues and eigenvectors

1 Basics

October 25, 2022

Definition 1.1 (Linear operators)**.** A linear operator is a linear map from one vector space to itself.

Definition 1.2 (Eigenvalues and eigenvectors)**.** Let $T\colon V \to V$ be linear, λ be a scalar and $v \in V$ be nonzero. Then v is called an *eigenvector* of T, and v a corresponding *eigenvector*, iff

$$Tv = \lambda v.$$

Similarly, we define eigenvalues and eigenvalues of square matrices (which are precisely the linear operators on F^n).

Remark. *We had a choice here: We could've defined this so that 0 would be an eigenvector of each scalar. But then we'd have had to specify nonzero-ness of eigenvectors each time (like we now do for "nonzero zero divisors").*

Proposition 1.3 (Eigenspaces are subspaces)**.** *Let $T\colon V \to V$ be linear with an eigenvalue λ. Then*

$$\{v \in V : Tv = \lambda v\} = \ker(T - \lambda \operatorname{id}_V).$$

Theorem 1.4. *Vectors corresponding to distinct eigenvalues of a linear operator are independent.*

Definition 1.5 (Diagonalizability)**.** A linear operator $T\colon V \to V$ is called diagonalizable iff there exists a basis of V, comprising only of eigenvectors of T.[1]

In the same way, we define diagonalizability of square matrices.

Corollary 1.6. *If $T\colon V \to V$ with V being finite-dimensional and T having* $\dim V$ *many distinct eigenvalues, then T is diagonalizable.*

Proposition 1.7 (Matrices suffice for finite-dimensional spaces)**.** *Let $T\colon V \to V$ be linear with V being finite-dimensional. Let B be an ordered basis of V. Then the following hold:*

- *(i) For any vector v and any scalar λ, the following are equivalent:*
  - *(a) v is an eigenvector of T with eigenvalue λ.*
  - *(b) $[v]_B$ is an eigenvector of $[T]_{B\leftarrow B}$ with eigenvalue λ.*
- *(ii) T is diagonalizable $\iff$ $[T]_{B\leftarrow B}$ is diagonalizable.*

Remark. *All of this above can be seen elegantly by formulating "morphisms between maps" and then we'll have that isomorphisms between maps preserve eigenvalues, eigenvectors and diagonalizability.*

2 Diagonalizability of matrices

October 25, 2022

Definition 2.1 (Similar matrices)**.** Two square matrices A and B (of same size) are called similar iff there exists an invertible P such that

$$A = P^{-1}BP.$$

Proposition 2.2. *Similarity is an equivalence relation.*

Proposition 2.3 (Similar matrices have same eigenvalues[2])**.** *Let $A, P \in F^{n\times n}$ with P invertible. Then for any $\lambda \in F$ and any $v \in F^{n\times 1}$, the following are equivalent:*

- *(i) v is an eigenvector of A with eigenvalue λ.*
- *(ii) $P^{-1}v$ is an eigenvector of $P^{-1}AP$ with eigenvalue λ.*

Corollary 2.4. *Similarity preserves diagonalizability.*

[1] See Corollary 2.6 for the motivation to call it "diagonalizability".

[2] Also see Proposition 4.10.

Theorem 2.5 (Diagonalizing matrix contains eigenvectors)**.** *Let $A, P \in F^{n\times n}$ with $P =: [v_1, \dots, v_n]$ invertible and let $\lambda_1, \dots, \lambda_n \in F$. Then the following are equivalent:*

(i) $P^{-1} A P = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$*.*

(ii) $Av_i = \lambda_i v_i$ *for* $i = 1, \dots, n$*.*

Corollary 2.6. *An $n \times n$ matrix is diagonalizable iff it is similar to a diagonal matrix.*

Result 2.7. Orthogonal matrices over $\mathbb{K}$ have eigenvalues with absolute value 1.

Proposition 2.8. *Eigenspaces corresponding to distinct eigenvalues are independent.*

Theorem 2.9 (Characterizing diagonalizability[3])**.** *Let $T\colon V \to V$ be linear with V finite-dimensional and $\lambda_1, \dots, \lambda_k$ be distinct eigenvalues of T for $k \geq 0$. Let E_{λ_i} be the corresponding eigenspaces. Then the following are equivalent:*

(i) *T is diagonalizable.*

(ii) $V = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_k}$*.*

(iii) $\dim V = \dim E_{\lambda_1} + \cdots + \dim E_{\lambda_k}$*.*

2.1 Orthogonal diagonalization

October 25, 2022

Definition 2.10 (Orthogonal diagonalizability)**.** Let V be an inner product space. Then a linear operator $T\colon V \to V$ is called orthogonally diagonalizable iff there exists an orthonormal basis of V comprising of eigenvectors of T.

Similarly, we define orthogonal diagonalizability for square matrices over $\mathbb{K}$ (under the Euclidean inner product on $\mathbb{K}^n$).

Definition 2.11 (Hermitian matrices)**.** An $n\times n$ A matrix over $\mathbb{K}$ is called Hermitian iff

$$A^* = A.$$

Proposition 2.12. *Eigenvalues of a Hermitian matrix are real and its eigenspaces are orthogonal.*

Theorem 2.13 (Spectral theorem for matrices)**.** *A square matrix over $\mathbb{K}$ is orthogonally diagonalizable with real eigenvalues $\iff$ it is Hermitian.*

[3] Also see

3 Reflections

October 25, 2022

Proposition 3.1 (Reflections and reflection matrices)**.** *Let V be an inner product space and $u \in V$ with $\|u\| = 1$. Then the function $V \to V$ given by*

$$v \mapsto v - 2\langle v, u\rangle u$$

is linear and has eigenvalues ± 1 with

$$E_{-1} = \operatorname{span}(\{u\}), \text{ and}$$
$$E_1 = E_{-1}^{\perp}.$$

If $V = \mathbb{K}^{n\times 1}$ with the Euclidean inner product, then this map is given by

$$v \mapsto (I - uu^*)v.$$

The matrix $R := I - uu^*$ is Hermitian and orthogonal, and hence $R^2 = I$.

Remark. *When there's no confusion, we'll denote eigenspaces by E_λ.*

4 Cardinal polynomials

Remark. *For this section, fix a general ring R. As usual, we'll assume V to be a vector space over some fixed general field F.*

4.1 Modules and algebras

October 27, 2022

Definition 4.1 (R-modules)**.** An R-module is an abelian additive group $(M, +)$ along with a scalar multiplication $R \times M \to M$ (denoted by juxtaposition) such that the following hold:

(i) $(r + s)m = rm + sm$.

(ii) $r(m + n) = rm + rn$.

(iii) $(rs)m = r(sm)$.

(iv) If R has an identity, then $1_R m = m$.

Definition 4.2 (R-algebras)**.** An R-module A along with a multiplication $\times\colon A \times A \to A$ is said to be an R-algebra iff $\times$ is bilinear in both slots, *i.e.*,

(i) $a \times (b + c) = a \times b + a \times c$,

(ii) $(a + b) \times c = a \times c + b \times c$, and

(iii) $(ra) \times (sb) = (rs)(a \times b)$.

We say that A is associative (respectively commutative; has an identity) iff $\times$ is associative (respectively commutative; has an identity).

Definition 4.3 (Nice homomorphisms)**.** A ring homomorphism $\phi\colon R \to S$ is called nice iff this holds: R has an identity $\implies$ $\phi(1_R)$ is the identity in S.

Definition 4.4 (Algebras via homomorphisms)**.** Let S be a ring. Then a nice ring homomorphism $\phi\colon R \to S$ is called an algebra iff $\phi(R)$ is central in S.

We say that ϕ is commutative (respectively, has an identity) iff S is commutative (respectively, has an identity).

Theorem 4.5 (Interplay of Definitions 4.2 and 4.4)**.** *Let R have identity and A be an associative R-algebra with identity. Define $\phi\colon R \to A$ as*

$$\phi(r) := r1_A.$$

Then A is a ring and ϕ is an algebra with identity.

Conversely, let S be a ring and $\phi\colon R \to S$ be a nice ring homomorphism. Define scalar multiplication $R \times S \to S$ as

$$(r, s) \mapsto \phi(r)s.$$

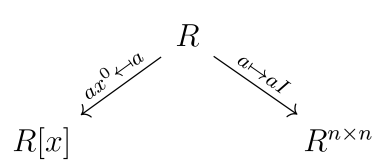
Then S forms an R-module. If $\phi(R)$ is further central in S (i.e., ϕ is an algebra), then S is an associative R-algebra.

Lemma 4.6 ("Transitivity" of modules and algebras)**.** *Let $\phi\colon R \to S$ and $\psi\colon S \to T$ be ring homomorphisms. Then the following hold:*

(i) ϕ, ψ are nice $\implies$ $\psi \circ \phi$ is nice.

(ii) $\psi(S)$ is central $\implies$ $(\psi \circ \phi)(R)$ is central.

Proposition 4.7. *We have the following nice ring homomorphisms:*



If R is commutative, then $a \mapsto a\,x^0$ is a commutative algebra as well.

Remark. *To distinguish elements of R from the constant polynomials in $R[x]$, we'll use x^0.*

Lemma 4.8 (When is $\phi\colon R[x] \to S$ a homomorphism?)**.** *For a ring S, a function $\phi\colon R[x] \to S$ is a ring homomorphism $\iff$ the following hold:*

(i) $\phi(0_{R[x]}) = 0_S$.

(ii) $\phi(p + ax^i) = \phi(p) + \phi(ax^i)$.

(iii) $\phi(ax^i\, bx^j) = \phi(ax^i)\,\phi(bx^j)$.

Proposition 4.9 (Substitution homomorphisms)**.** *Let $\phi\colon R \to S$ be an algebra and $s \in S$. Then the function $R[x] \to S$ given by*

$$a_0x^0 + \cdots + a_nx^n \mapsto \phi(a_0) + \phi(a_1)s^1 + \cdots + \phi(a_n)s^n$$

is a nice homomorphism.

Remark. *Strictly speaking, the well-defined-ness of this function needs to be shown.*

Notation. *For such homomorphisms, we'll use the notation $p \mapsto p(s)$, and also denote the image of $R[x]$ as $\phi(R)[s]$.*

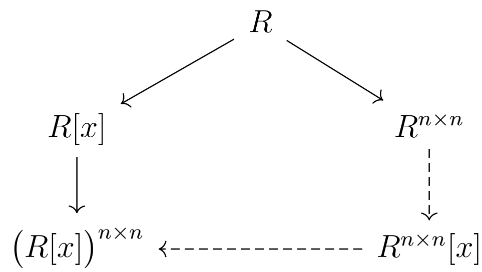
Remark. *If R is a subring of S, we'll take ϕ to be the inclusion $R \hookrightarrow S$ if not explicitly mentioned.*

Proposition 4.10 (T and $p(T)$ have the same eigenvectors)**.** *Let $T\colon V \to V$ be linear and $p \in F[x]$. Let $v \in V$ and $\lambda \in F$. Then*

$$Tv = \lambda v \implies p(T)v = p(\lambda)v.$$

4.2 Characteristic polynomial

Proposition 4.11 (Matrices of polynomials and vice-versa)**.** *We have the following commutative diagram with the canonical nice homomorphisms:*



Here, the homomorphism $R^{n\times n}[x] \to \big(R[x]\big)^{n\times n}$ *is given by*

$$\left[a_{ij}^{(0)}\right]x^0 + \cdots + \left[a_{ij}^{(n)}\right]x^n \mapsto \left[a_{ij}^{(0)}x^0 + \cdots + a_{ij}^{(n)}x^n\right].$$

The solid arrows become algebras if R is commutative.

Definition 4.12 (Characteristic polynomial of matrices)**.** Let R be commutative and $A \in R^{n\times n}$. Then we define the characteristic polynomial to be the polynomial in $R[x]$ given by

$$\det\big(f(-Ax^0 + Ix)\big)$$

where $f\colon R^{n\times n}[x] \to \big(R[x]\big)^{n\times n}$ is the homomorphism as given in Proposition 4.11.

Remark. *We have the nice properties of determinants holding only when the matrix entries come from a commutative rings. Hence we care to define* det *only here.*

Proposition 4.13. *Let R be commutative and $A \in R^{n\times n}$. Then the characteristic polynomial of A is monic and of degree n.*

Proposition 4.14. *The eigenvalues of a square matrix over a field are precisely the zeroes of its characteristic polynomial.*

Proposition 4.15. *Similar matrices have the same characteristic polynomial.*

Corollary 4.16 (Characteristic polynomial of operators)**.** *Let $T\colon V \to V$ be linear with V finite-dimensional*[4] *and B, C ordered bases of V. Then the characteristic polynomials of $[T]_{B\leftarrow B}$ and $[T]_{C\leftarrow C}$ are the same.*

[4]Finite-dimensionality is needed to talk of any matrix of T.

Remark. *This allows us to talk of "the characteristic polynomial of T".*

Corollary 4.17. *Let $T\colon V \to V$ be linear with V being finite-dimensional. Then the eigenvalues of T are precisely the zeroes of its characteristic polynomial.*

Theorem 4.18 (Characterizing diagonalizability)**.** *Let $T\colon V \to V$ be linear with V finite-dimensional. Let $\lambda_1, \ldots, \lambda_k$ be eigenvalues of T and $E_{\lambda_1}, \ldots, E_{\lambda_k}$ be the corresponding eigenspaces. Then T is diagonalizable $\iff$ the characteristic polynomial of T is given by*

$$(x - \lambda_1)^{\dim E_{\lambda_1}} \cdots (x - \lambda_k)^{\dim E_{\lambda_k}}.$$

Remark. *In writing statements on vector spaces, we'll simply write x instead of $1_F x$, etc.*

Proposition 4.19. *Let R be commutative. Then for a square block matrix over R, we have*

$$\det \begin{bmatrix} A & B \\ 0 & D \end{bmatrix} = \det A \, \det D,$$

where A and B are square of possibly different sizes.

Similarly, we have that the characteristic polynomial also factorizes as above.

Definition 4.20 (Invariant subspaces)**.** Let $T\colon V \to V$ be linear. Then a subspace W of V is called T-invariant iff

$$T(W) \subseteq W.$$

Notation. *We'll denote the restriction of T on $W \to W$ by T_W.*

Corollary 4.21. *Let $T\colon V \to V$ be linear with V being finite-dimensional. Let W be a T-invariant subspace. Then the characteristic polynomial of T_W divides that of T.*

Result 4.22. Let $T\colon V \to V$ be linear. Then the following hold:

(i) $\{0\}$, V, $\ker T$, $\operatorname{im} T$ as well as eigenspaces are all T-invariant.

(ii) Let W be a subspace and $p \in F[x]$. Then W is T-invariant $\implies$ W is $p(T)$-invariant.

4.3 Division

Proposition 4.23 (Associates)**.** *Let R have an identity. Then the relation on R defined by*

$$a \sim b \text{ iff } a = ub \text{ for some invertible } u \in R$$

is an equivalence relation.

Remark. *Similarly, "$a = bu$" will define another equivalence relation. For commutative rings, both the relations coincide.*

Theorem 4.24 (Division is a "partial order" in integral domains)**.** *Let R be an integral domain. Then the following hold:*

(i) $a \mid a$.

(ii) $a \mid b$ *and* $b \mid a \implies a, b$ *are associates.*

(iii) $a \mid b$ *and* $b \mid c \implies a \mid c$.

Remark. *This allows to define* gcd *and* lcm *with these being "greatest" and "least" in some sense.*

Definition 4.25 (gcd and lcm in integral domains)**.** Let R be an integral domain and $a, b \in R$. Then an $x \in R$ is called

(i) a gcd of a, b iff
   - (a) x is a common divisor of a, b, and
   - (b) if d is any common divisor of a, b, then $d \mid x$;

(ii) an lcm of a, b iff
   - (a) x is a common multiple of a, b, and
   - (b) if m is any common multiple of a, b, then $x \mid m$.

Proposition 4.26. *In an integral domain,* gcd*'s (respectively* lcm*'s) of a pair of elements are unique up to associativity.*

Definition 4.27 (Coprimes)**.** Let R be an integral domain. Then $a, b \in R$ are said to be coprime iff they have 1_R as a gcd.

Definition 4.28 (Primes)**.** Let R be commutative. Then $p \in R \setminus \{0_R\}$ ($p \neq 1_R$ too if R has identity) is called prime iff

$$p \mid ab \implies p \mid a \text{ or } p \mid b.$$

Proposition 4.29 (Divisors of prime products)**.** *Let R be an integral domain $p_1, \dots, p_n$ be primes for $n \geq 0$. Let u be invertible and $d_1, \dots, d_n \geq 0$. Then divisors of $up_1^{d_1} \cdots p_n^{d_n}$ are precisely of the form*

$$vp_1^{e_1} \cdots p_n^{e_n}$$

where v is invertible and $0 \leq e_i \leq d_i$.

Lemma 4.30. *Primes are irreducible in an integral domain.*

Proposition 4.31. *Let R be an integral domain and $\lambda, \mu \in R$ be distinct. Then $1_R x^1 - \lambda x^0$ and $1_R x^1 - \mu x^0$ are non-associate primes.*

Proposition 4.32. *R is an integral domain $\iff$ $R[x]$ is an integral domain.*

Theorem 4.33 (Division in $R[x]$ and the factor theorem)**.** *Let R have identity and $f, g \in R[x]$ with g being monic.*[5] *Then the following hold:*

(i) There exist $q, r \in R[x]$ such that

$$f = qg + r \text{ with } r = 0, \text{ or else, } \deg r < \deg q.$$

(ii) For $\alpha \in R$, we have that

$$p(\alpha) = 0_R \iff (1_R x^1 - \alpha x^0) \text{ divides } p \text{ from both sides.}$$

(iii) If R further has no nonzero zero divisor, then these q, r are unique.

Remark. *Exactly similar proposition will hold for the quotient q appearing on the right of g.*

4.4 Annihilators

Definition 4.34 (Annihilating and minimal polynomials[6])**.** Let $\phi \colon R \to S$ be an R-algebra.[7] Let I be an ideal of S and $s \in S$. Then a polynomial $p \in R[x]$ is called an I-annihilator of s iff

$$p(s) \in I.$$

If p is such a nonzero polynomial with least degree, it's called a minimal I-annihilator of s.

When $I = \{0_S\}$, we'll simply call these annihilators.

[5] We can weaken this by demanding the leading coefficient of q to be invertible.

[6] In this case, S fails in general to be an $R[x]$-algebra (we could define a product $(p, s) \mapsto p(s)$ on $R[x] \times S \to S$), hence the "annihilators of a module" will not suffice for us.

[7] ϕ is required to be an R-algebra to talk of $p(s)$.

Remark. *I am allowing minimal annihilators to be non-monic.*

In the case of a vector space V over F, we'll have $R = F$, $S = \mathcal{L}(V, V)$ or $F^{n\times n}$ (all of which have identities), and $I = \{0\}$ with $\phi\colon \alpha \mapsto \alpha\, \mathrm{id}_V$ or $\alpha\, I_n$.

Proposition 4.35 (Monic minimals divide all the annihilators, and are unique)**.** *Let $\phi\colon R \to S$ be an algebra. Let I be an ideal of S and $s \in S$. Then the following hold:*

(i) The set

$$\mathcal{A} := \{I\text{-}annihilators\ of\ s\}$$

forms an ideal of $R[x]$.

(ii) If R further has an identity[8] and $m \in R[x]$ is any monic minimal I-annihilator for s, then

$$\mathcal{A} = m\, R[x] = R[x]\, m = (\!| m |\!).$$

(iii) In addition, R further has no nonzero zero divisors, then the monic minimal I-annihilator of s, if existent, is unique.

Proposition 4.36 (Polynomials over matrices suffice for finite-dimensional)**.** *Let $T\colon V \to V$ be linear with V being finite-dimensional. Let B be an ordered basis for V and $p \in F[x]$. Then*

$$\big[p(T)\big]_{B\leftarrow B} = p\big([T]_{B\leftarrow B}\big).$$

Hence, the (minimal) annihilators of T are precisely the (minimal) annihilators of $[T]_{B\leftarrow B}$.

Corollary 4.37. *Each $T \in \mathcal{L}(V, V)$ for finite-dimensional V has a unique minimal annihilator.*

Proposition 4.38 (Minimals also give eigenvalues)**.** *For any $A \in F^{n\times n}$, the characteristic and minimal polynomials have the same zeros.*

Proposition 4.39. *Similarity preserves (minimal) annihilators.*

[8] This is required to talk of monic m which is in turn required for division. See Theorem 4.33.