University of Houston – Fall 2012 – Dr. G. Heier
Advanced Linear Algebra I (Math 4377/6308)

Contents

These lecture notes are based on the textbook Linear Algebra, 4th edition, by Friedberg, Insel, and Spence, ISBN 0-13-008451-4. They are provided "as is" and as a courtesy only. They do not replace use of the textbook or attending class.

0 Foundational material: The appendices

0.1 Appendix A: Sets

Definition 0.1. A *set* is a collection of objects, called elements.

Example 0.2.

- $\{1,2,3\} = \{2,1,1,1,2,3\}$ (no notion of "multiplicity")
- $[1,2]$ = the interval of reals between 1 and 2, inclduing 1 and 2.
- $\mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$ (later), $\mathbb{Z}^n, \mathbb{Q}^n, \mathbb{R}^n, \mathbb{C}^n$
- $\left\{ \begin{pmatrix} 8 \\ 0 \\ -1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 2 \\ 2 \end{pmatrix} \right\}$ = set of two vectors
- $\emptyset$: the empty set

Given two sets A, B, there are several operations that yield new sets from these. Most important are the following:

- $A \cup B$ (union of A and B)
- $A \cap B$ (intersection of A and B)
- $A \setminus B = \{x \in A : x \notin B\}$ (complement of B in A)
- $A \times B = \{(a,b) : a \in A, b \in B\}$ (product of A and B)

Definition 0.3. Let A be a set. A *relation* on A is a subset S of $A \times A$. Write $x \sim y$ if and only if $(x,y) \in S$.

Example 0.4.

- $A = \{1,2,3\}$, $S = \{(1,2),(1,3),(2,3)\}$. This relation is "$<$".
- $A = \{1,2,3\}$, $S = \{(1,2),(1,3),(2,3),(2,1),(2,3),(3,1),(3,2)\}$. This relation is "$\neq$".
- $A = \{1,2,3\}$, $S = \{(1,1),(2,2),(3,3)\}$. This relation is "$=$".

Recall the following symbols. $\forall$: "for all", $\exists$: "there exists".

Definition 0.5. Let A be a set with a relation S. Then S is called an *equivalence relation* if and only if

i. $\forall x \in A : x \sim x$ (reflexive)

ii. $\forall x, y \in A : x \sim y \Leftrightarrow y \sim x$ (symmetric)

iii. $\forall x, y, z \in A : (x \sim y \text{ and } y \sim z) \Rightarrow x \sim z$ (transitive)

Example 0.6. Let $A = \mathbb{Z}$. Let $x \sim y \Leftrightarrow \exists k \in \mathbb{Z} : x - y = 5k$. This defines an equivalence relation. (Check it.)

0.2 Appendix B: Functions

Definition 0.7. Let A, B be sets. A *function* $f : A \to B$ is a rule that associates to each element $x \in A$ a unique element of B, denoted $f(x)$. The set A is called the *domain*, the set B is called the *codomain*.

Definition 0.8.

- For $S \subseteq A, f(S) = \{f(x) : x \in S\}$ (*image of S under f*). $f(A)$ is called the *range*.
- For $T \subseteq B, f^{-1}(T) = \{x \in A : f(x) \in T\}$ (*pre-image of T under f*)
- $f : A \to B = g : A \to B \Leftrightarrow \forall x \in A : f(x) = g(x)$

Definition 0.9.

- $f : A \to B$ is *one-to-one* (aka *injective*) if and only if $f(x) = f(y) \Rightarrow x = y$.
- $f : A \to B$ is *onto* (aka *surjective*) if and only if $\forall b \in B \exists a \in A : f(a) = b$.
- f is *bijective* if and only if f is injective and surjective.
- For $S \in A$, the restriction of f to S is $f|_S : S \to B, x \mapsto f(x)$.

0.3 Appendix C: Fields

Definition 0.10. Let A be a set. A binary operation is any map $A \times A \to A$. We are very familiar with $\mathbb{Q}$ and $\mathbb{R}$ and the properties that the two binary operations $+$ and $\cdot$ have.

Definition 0.11. A field F is a set with two binary operations labelled $+$ and $\cdot$ such that

i. $a + b = b + a, \quad a \cdot b = b \cdot a$ (commutativity)

ii. $(a + b) + c = a + (b + c)$ and $(a \cdot b) \cdot c = a \cdot (b \cdot c)$ (associativity)

iii. $\exists 0 \in F : a + 0 = a \ \forall a$
$\exists 1 \in F : 1 \cdot a = a \ \forall a$ (neutral elements)

iv. $\forall a \in A : \exists b \in A : a + c = 0$
$\forall a \in A \backslash \{0\} : \exists b \in A : a \cdot b = 1$ (inverse elements)

v. $a \cdot (b + c) = a \cdot b + a \cdot c$ (distributive law)

Theorem 0.12 (Cancellation Laws). *Let F be a field and $a, b, c \in F$.*

i. $a + b = c + b \Rightarrow a = c$

ii. $a \cdot b = c \cdot b$ *and* $b \neq 0 \Rightarrow a = c$

Proof. Part i. Let d be an additive inverse of b. Now, observe that $(a+b)+d = a$ and $(c+b)+d = c$. Done.

Part ii is done in detail in the textbook. □

Proposition 0.13. *The neutral element of addition is unique.*

Proof. Let 0 and $0'$ be two neutral elements of addition. Then

$$0 = 0 + 0' = 0'.$$

□

Example 0.14. Some examples of fields.

- $\mathbb{Q}, \mathbb{R}, \mathbb{C}$
- $\mathbb{Q}[\sqrt{2}] = \{a + b\sqrt{2} | a, b \in \mathbb{Q}\}$.
- $\mathbb{Z}/p\mathbb{Z}$ when p is a prime.
- For all primes p and $n \in \mathbb{N}$, there exists a unique finite field, often denoted $GF(p^n)$, with p^n elements.

0.4 Appendix D: Complex Numbers

Motivation: In $\mathbb{R}$, $x^2 - 1 = 0$ has two solutions, namely $-1, 1$. However, the almost identical equation $x^2 + 1 = 0$ has no solutions. This means that the reals "leave something to be desired." In response, we introduce the imaginary unit i, which has the property $i^2 = -1$.

Definition 0.15. A *complex number* is an expression of the form $z = a + bi$ with $a, b \in \mathbb{R}$. Sum and product are defined by

$$z + w = (a + bi) + (c + di) = a + c + (b + d)i$$

and

$$zw = (a + bi)(c + di) = (ac - bd) + (ad + bc)i.$$

Remark 0.16. Memorize the multiplication by multiplying out as one would do naively, and then use $i^2 = -1$. Do some examples!

Theorem 0.17. *The complex numbers with sum and multiplication as above form a field.*

Proof. This just involves tedious checking of all the properties–you should try a few yourself at home. □

Remark 0.18. The multiplicative inverse of $z = a + bi$ is

$$\frac{1}{a+bi} = \frac{a-bi}{(a+bi)(a-bi)} = \frac{a-bi}{a^2+b^2} = \frac{a}{a^2+b^2} + i\frac{-b}{a^2+b^2}.$$

Definition 0.19. The *complex conjugate* of $z = a + bi$ is $\bar{z} = a - bi$.

Proposition 0.20. *i.* $\bar{\bar{z}} = z$

ii. $\overline{z+w} = \bar{z} + \bar{w}$

iii. $\overline{zw} = \bar{z} \cdot \bar{w}$

iv. $\overline{\frac{z}{w}} = \frac{\bar{z}}{\bar{w}}$

Remark 0.21. It is now clear that there is a bijection $\mathbb{C} \to \mathbb{R}^2$ via $a + bi \mapsto (a, b)$. By Pythagoras' Theorem, the length of a straight line from the origin to the point (a, b) is $\sqrt{a^2 + b^2}$.

Definition 0.22. The *absolute value* (or *modulus*) of $z = a + bi$ is $|z| = \sqrt{a^2 + b^2}$.

Remark 0.23. We have

$$z\bar{z} = (a + bi)(a - bi) = a^2 + b^2.$$

Thus,

$$|z| = \sqrt{z\bar{z}}.$$

Properties 0.24. i. $|zw| = |z||w|$

ii. $\left|\frac{z}{w}\right| = \frac{|z|}{|w|}$

iii. $|z+w| \leq |z| + |w|$

iv. $|z| - |w| \leq |z+w|$

Theorem 0.25 (Fundamental Theorem of Algebra)**.** *Let* $p(z) = a_n z^n + a_{n-1} z^{n-1} + \ldots + a_1 z + a_0$ *be a complex polynomial (i.e.,* $a_i \in \mathbb{C}$*). Then* $\exists z_0 \in \mathbb{C} : p(z_0) = 0$*.*

Proof. No proof is given here. This is a comparatively hard theorem to prove. □

Corollary 0.26. *For* p *as above,* $\exists r_1, \ldots, r_n \in \mathbb{C}$ *such that*

$$p(z) = a_n (z - r_1) \ldots (z - r_n).$$

Proof. Long division! □

Remark 0.27. The formula often taught in high school to solve quadratic equations still works. For example, to solve $x^2 - 2x + 5 = 0$, write $x = -\frac{-2}{2} \pm \sqrt{1-5} = 1 \pm \sqrt{-4} = 1 \pm \sqrt{4(-1)} = 1 \pm 2\sqrt{-1} = 1 \pm 2i$.

1 Vector spaces

1.1 Introduction

Geometrically, a vector in, say, $\mathbb{R}^2$, is the datum of a direction and a magnitude. Thus, it can be represented by an arrow which points in the given direction and has the given length. Two vector can be added using the parallelogram rule (see the textbook for some nice pictures explaining this).

Physically, the vectors may, e.g., represent forces that are exerted on an object. The result of the addition is the resulting net force that the object experiences when the original two forces are applied.

Algebraically, when $v = (a_1, a_2)$ and $w = (b_1, b_2)$, then $v + w = (a_1 + b_1, a_2 + b_2)$. Scalar multiplication is defined via $t(a_1, a_2) = (ta_1, ta_2)$.

Definition 1.1. The vectors v and w are parallel if and only if $\exists t \in \mathbb{R} : tv = w$.

A vector can be interpreted as the *displacement vector* between its start and end point. If the start point is (x_1, x_2) and the end point is (y_1, y_2), then the displacement vector is $(y_1 - x_1, y_2 - x_2)$.

Definition 1.2. The *line* through the points $A = (x_1, x_2)$ and $B = (y_1, y_2)$ is

$$\{(x_1, x_2) + t(y_1 - x_1, y_2 - x_2) : t \in \mathbb{R}\}.$$

Definition 1.3. The *line* through the points $A = (x_1, x_2, x_3)$ and $B = (y_1, y_2, y_3)$ is

$$\{(x_1, x_2, x_3) + t(y_1 - x_1, y_2 - x_2, y_3 - x_3) : t \in \mathbb{R}\}.$$

Definition 1.4. The *plane* through the points $A = (x_1, x_2, x_3)$, $B = (y_1, y_2, y_3)$ and $C = (z_1, z_2, z_3)$ (not all three on a line) is

$$\{(x_1, x_2, x_3) + s(y_1 - x_1, y_2 - x_2, y_3 - x_3) + t(z_1 - x_1, z_2 - x_2, z_3 - x_3) : s, t \in \mathbb{R}\}.$$

Example 1.5.

i. The line through $(1, 1, 2)$ and $(0, 3, -1)$ is

$$\{(1, 1, 2) + t(-1, 2 - 3) : t \in \mathbb{R}\}.$$

ii. The plane through the points $A = (1, 0, -1)$, $B = (0, 1, 2)$ and $C = (1, 1, 0)$ is

$$\{(1, 0, -1) + s(-1, 1, 3) + t(0, 1, 1) : s, t \in \mathbb{R}\}.$$

Now, observe that vector addition and scalar multiplication satisfy certain laws, e.g., $v + w = w + v$, $1 \cdot v = v$, $(ab)v = a(bv)$. Next, we will distill these obvious properties into an abstract definition.

1.2 Vector Spaces

Definition 1.6. A *vector space* (or *linear space*) V over a field F (think $F = \mathbb{R}$, or $\mathbb{C}$) is a set with a binary operation denoted "+" and a second map $\cdot : F \times V \to V$ such that

i. $\forall x, y \in V : x + y = y + x$

ii. $\forall x, y, z \in V : (x + y) + z = x + (y + z)$

iii. $\exists 0 \in V : \forall x \in V : x + 0 = x$

iv. $\forall x \in V \exists y \in V : x + y = 0$

v. $\forall x \in V : 1x = x$, where 1 is the neutral element of multiplication in F

vi. $\forall a, b \in F \forall x \in V : (ab)x = a(bx)$

vii. $\forall a \in F \forall x, y \in V : a(x + y) = ax + ay$

viii. $\forall a, b \in F \forall x \in V : (a + b)x = ax + bx$

Definition 1.7. The elements of F are called *scalars*. The elements of V are called *vectors*. Because of item ii above, sums like $x + y + z + w$ are well-defined.

Remark 1.8. To simplify typing, we will usually not adorn vectors with an arrow, i.e., we will write x instead of $\vec{x}$ and 0 instead of $\vec{0}$. Note that the neutral element of addition in the field is also denoted with 0, but it should always be clear from the context what is meant.

Example 1.9. i. (THE example, see later section on isomorphisms) Take a field F. (We will mostly just take $\mathbb{R}$, or perhaps $\mathbb{C}$.) An n-tuple is $(a_1, \ldots, a_n)$, where $a_1, \ldots, a_n \in F$. Note that $\{n\text{-tuples}\} \cong F^n$ naturally. Define $(a_1, \ldots, a_n) + (b_1, \ldots, b_n) = (a_1 + b_1, \ldots, a_n + b_n)$. Also, $c(a_1, \ldots, a_n) = (ca_1, \ldots, ca_n)$ for $c \in F$.

ii. $\mathrm{Mat}_{m,n}(F)$ is a vector space with componentwise addition and scalar multiplication

The following examples of vector spaces are substantially different from the examples above. They are "infinite dimensional," more about that later.

Example 1.10. i. Let S be a real interval, e.g., $(1, 2)$. Let $\mathcal{F}$ be the set of all real-valued functions on S. Then $\mathcal{F}$ is a vector space (over $\mathbb{R}$) with the usual addition and scalar multiplication of real-valued functions.

ii. Consider in $\mathcal{F}$ only those functions that are continuous. This is also a vector space (think back to calculus!)

iii. Consider in $\mathcal{F}$ only those functions that are differentiable. This is also a vector space (think back to calculus!)

iv. The set of all sequences of real numbers is a vector space with

$$(a_1, a_2, a_3, \ldots) + (b_1, b_2, b_3, \ldots) = (a_1 + b_1, a_2 + b_2, a_3 + b_3, \ldots)$$

and

$$c(a_1, a_2, a_3, \ldots) = (ca_1, ca_2, ca_3, \ldots).$$

We are free to define "+" as we like, e.g., by

$$(a_1, a_2) + (b_1, b_2) = (a_1 + b_1, a_2 - b_2).$$

However, most of the time, not all vector space conditions are satisfied. In the above case, already the first condition is violated.

Theorem 1.11 (Cancellation law for vector spaces)**.** *Let V be a vector space and $x, y, z \in V$. If $x + z = y + z$, then $x = y$.*

Proof. Let v be such that $z + v = 0$ (condition iv). Then

$$\begin{aligned} x &= x + 0 = x + (z + v) = (x + z) + v \\ &= (y + z) + v = y + (z + v) = y + 0 = y \end{aligned}$$

due to conditions ii and iii. □

Corollary 1.12. *The vector* 0 *is unique.*

Corollary 1.13. *The additive inverse is unique.*

Theorem 1.14. *Let V be a vector space. Then the following statements are true.*

i. $\forall x \in V : 0x = \vec{0}$

ii. $\forall x \in V\ \forall a \in F : (-a)x = -(ax) = a(-x)$

iii. $\forall a \in F : a\vec{0} = \vec{0}$

Proof. The textbook has detailed proofs of i and ii. The item iii is left to the reader. □

1.3 Subspaces

Definition 1.15. A subset W of a vector space V over the field F is called a *subspace of* V if W is a vector space with + and scalar multiplication from V.

Example 1.16.

- $\{\vec{0}\}$, V
- $\mathbb{R}^2 \cong \{(a, b, 0) | a, b \in \mathbb{R}\} \subset \mathbb{R}^3$

- The above examples 1.10ii and 1.10iii in 1.10i.

In order to verify that W is a subspace of V, it is not necessary to check all the vector space axioms in the definition of a vector space. For example, the restricted addition is clearly commutative since it was already commutative before the restriction.

Theorem 1.17. *A nonempty subset W of the vector space V is a subspace of V if and only if*

i. $\forall x, y \in W : x + y \in W$ *(closedness under +)*

ii. $\forall c \in F \forall x \in W : c \cdot x \in W$ *(closedness under scalar multiplication)*

Proof. First, observe that the implication $\Rightarrow$ is trivial. The proof of the other direction consists of some easy verifications. For example, let's see why $\vec{0} \in W$: Take an arbitrary element x of W. Since W is nonempty, such an element exists. Now, simply observe that $0 \cdot x = \vec{0}$, which is an element of W by ii. The remaining details are left to the reader. □

More examples:

Example 1.18.

- Let $W = \{(a, b) | a + b = 0\} \subset \mathbb{R}^2$. Closedness under $+$ is checked as follows. Let $(a, b), (c, d) \in W$. Then the result of the addition is $(a+c, b+d)$, which satisfies $(a+c)+(b+d) = (a+b)+(c+d) = 0+0 = 0$. Closedness under scalar multiplication is seen as follows. Let $(a, b) \in W$ and c a scalar. Then the result of the scalar multiplication is (ca, cb), which satisfies $ac + cb = c(a + b) = c0 = 0$.

- Let $W = \{(a, b, c) | 3a - b + 2c = 0\} \subset \mathbb{R}^3$. Check it as in i.

- What about $W = \{(a, b, c) | 3a - b + 2c = 1\} \subset \mathbb{R}^3$? Let $(a, b, c), (d, e, f) \in W$. Then the result of their addition is $(a+d, b+e, c+f)$, which satisfies $3(a+d) - b - e + 2(c+f) = 3a - b + 2c + 3d - e + 2f = 1 + 1 = 2 \neq 1$. Thus, W is not closed under addition and not a subspace.

- Recall that the trace of a square matrix $(a_{ij})_{i,j=1,\ldots,n}$ is $\operatorname{tr}((a_{ij})_{i,j=1,\ldots,n}) = \sum_{i=1}^{n} a_{ii}$. Let $W = \{(a_{ij})_{i,j=1,\ldots,n} | \operatorname{tr}((a_{ij})_{i,j=1,\ldots,n}) = 0\}$. It is easy to check that W is a subspace of the vector space of $n \times n$ square matrices.

- Upper triangular matrices also form subspaces.

- Any intersection of subspaces in a vector space is itself a subspace. (Checking this is on the HW.)

A major class of examples is given by sums and direct sums of subspaces.

Definition 1.19. Let V be a vector space. Let S, T be nonempty subsets of V. Then let $S + T = \{x + y | x \in S, y \in T\}$. We call $S + T$ the *sum* of S and T.

Definition 1.20. Let V be a vector space. Let W, U be subspaces of V. Then we call V the *direct sum* of W, U if $W + U = V$ and $W \cap U = \{0\}$. Write $V = W \oplus U$.

Proposition 1.21. *Let V be a vector space. Let W, U be subspaces of V. Then the sum $W + U$ is a subspace of V (containing both W and U).*

Proof. $(w_1 + u_1) + (w_2 + u_2) = (w_1 + w_2) + (u_1 + u_2)$, which is the sum of a vector in W, namely $w_1 + w_2$, and a vector in U, namely $u_1 + u_2$. Thus, $W + U$ is closed under addition. The closedness under scalar multiplication is completely analogous. □

Example 1.22.

- $\{(a, b, 0, c) | a, b, c \in \mathbb{R}\} + \{(d, 0, e, f) | d, e, f \in \mathbb{R}\} = \mathbb{R}^4$. But this is not a direct sum.
- $\{(a, 0, 0, b) | a, b \in \mathbb{R}\} + \{(0, c, d, 0) | c, d \in \mathbb{R}\} = \mathbb{R}^4$. This is a direct sum.
- $\{f : \mathbb{R} \to \mathbb{R} | f(5) = 0\} \oplus \{\text{constant functions } \mathbb{R} \to \mathbb{R}\} = \{\text{all functions } \mathbb{R} \to \mathbb{R}\}$.
- upper and lower triangular square matrices
- $\{(a, 0, 0) | a \in \mathbb{R}\} \oplus \{(0, b, 0) | b \in \mathbb{R}\} = \{(a, b, 0) | a, b \in \mathbb{R}\}$
- The vector space of all functions $f : \mathbb{R} \to \mathbb{R}$ can be written as the direct sum of the subspace $\{f : \mathbb{R} \to \mathbb{R} | f(5) = 0\}$ and some other subspace W. Try to find such a subspace W which works for this.

Quiz 1:

1. (3 points) Let $W = \{(a_1, a_2) \in \mathbb{R}^2 \mid a_1^2 = a_2^2\}$. Is W a subspace of $\mathbb{R}^2$? Prove your answer carefully.

2. (2 points) Let $W_1 = \{(a_1, a_2, a_3) \in \mathbb{R}^3 \mid a_1 + a_2 - a_3 = 0\}$. Prove that W_1 is a subspace of $\mathbb{R}^3$.

3. (5 points) Find a subspace W_2 of $\mathbb{R}^3$ such that $\mathbb{R}^3 = W_1 \oplus W_2$. Prove carefully that your W_2 is a subspace and that $\mathbb{R}^3$ is indeed the direct sum of W_1 and W_2.

Solution:

1. $(1,1),(1,-1) \in W$, but $(1,1)+(1,-1) = (2,0) \notin W$. Thus, W is not closed under addition and thus not a subspace.
2. Was treated in the above examples.
3. Let W_2 be $\{(0,0,s)|s \in \mathbb{R}\}$. Clearly, $W_1 \cap W_2 = \{\vec{0}\}$ and an arbitrary $(a,b,c) = (a,b,-a-b) + (0,0,c+a+b)$.

1.4 Linear Combinations and systems of linear equations

Definition 1.23. Let V be a vector space and S a nonempty subset of V. We call $v \in V$ a *linear combination* of vectors in S if there exist vectors $u_1, \ldots, u_n \in S$ and scalars $a_1, \ldots, a_n \in F$ such that $v = a_1u_1 + \ldots + a_nu_n$.

Example 1.24.

- $(3,4,1) = 3(1,0,0) + 4(0,1,0) + 1(0,0,1)$.
- If we want to write $(3,1,2)$ as a linear combination of $(1,0,1),(0,1,1),(1,2,1)$, how do we find the coefficients a_1, a_2, a_3? Answer: Make the Ansatz

$$(3,1,2) = a_1(1,0,1) + a_2(0,1,1) + a_3(1,2,1)$$

  and solve the system of linear equations

$$a_1 + a_3 = 3, \quad a_2 + 2a_3 = 1, \quad a_1 + a_2 + a_3 = 2.$$

  Solution: $a_1 = 2, a_2 = -1, a_3 = 1$.

- Find the a_i in a given situation may or may not be possible. E.g., writing

$$(3,1,2) = a_1(1,0,0) + a_2(0,1,0) + a_3(1,2,0)$$

  is clearly impossible.

- There may be many choices for the a_i:

$$(2,6,8) = a_1(1,2,1)+a_2(-2,-4,-2)+a_3(0,2,3)+a_4(2,0,-3)+a_5(-3,8,16)$$

  is equivalent to

$$(a_1, a_2, a_3, a_4, a_5) \in \{(-4+2s-t, s, 7-3t, 3+2t, t)|s,t \in \mathbb{R}\}.$$

  (There are two "free variables".)

There are three types of operations that we used to solve the above systems of linear equations:

i. Interchange the order of any two equations.

ii. Multiply an equation by a nonzero scalar.

iii. Add a constant multiple of one equation to another.

Key point: These operations do not change the set of solutions.

Definition 1.25. Let V be a vector space. Let S be a nonempty subset of V. We call $\text{span}(S)$ the set of all vectors in V that can be written as a linear combination of vectors in S.

Example 1.26. Let $S = \{(1,0,0),(0,1,0),(2,1,0)\}$. Then $\text{span}(S) = \{(s,t,0)|s,t \in \mathbb{R}\}$.

Theorem 1.27. *The span of any subset S of a vector space V is a subspace of V.*

Proof. Let $v = a_1u_1 + \ldots + a_nu_n \in \text{span}(S)$. Then $cv = (ca_1)u_1 + \ldots + (ca_n)u_n \in \text{span}(S)$. Thus, closedness under scalar multiplication is ok.

Let $v = a_1v_1 + \ldots + a_nv_n \in \text{span}(S)$ and let $w = b_1w_1 + \ldots + b_mw_m \in \text{span}(S)$. Then

$$v + w = a_1v_1 + \ldots + a_nv_n + b_1w_1 + \ldots + b_mw_m.$$

Thus, closedness under addition is ok. □

Example 1.28.

- $S = \{(1,0,0),(0,1,0),(0,0,1)\}$ spans $\mathbb{R}^3$.
- $S = \{(1,2),(2,1)\}$ spans $\mathbb{R}^2$.
- $S = \{(1,2)\}$ does not span $\mathbb{R}^2$.
- Which (a,b,c) are in $\text{span}(\{(1,1,2),(0,1,1),(2,1,3)\})$? Answer: Those that satisfy $a + b = c$.
- Does $S = \{(1,0,1),(1,1,1),(2,1,3)\}$ span $\mathbb{R}^3$? Answer: yes.

1.5 Linear dependence and linear independence

Motivation: Let W be a subspace of V. We are interested in a set $S \subset W$ such that $\operatorname{span}(S) = W$ and S is "as small as possible".

Definition 1.29. A subset S of a vector space V is called *linearly dependent* if there exist a finite number of vectors $u_1, \ldots, u_n \in S$ and scalars $a_1, \ldots, a_n$, not all equal to zero, such that

$$a_1 u_1 + \ldots + a_n u_n = 0.$$

We also say that the vectors in S are linearly dependent.

Example 1.30. Let $S = \{(1,3,-4,2),(2,2,-4,0),(1,-3,2-4),(-1,0,1,0)\}$. Then

$$4(1,3,-4,2) - 3(2,2,-4,0) + 2(1,-3,2-4) + 0(-1,0,1,0) = (0,0,0,0).$$

Thus, S is linearly dependent.

Definition 1.31. If S is not linearly dependent, we say S is *linearly independent.*

Remark 1.32. Linear independence is equivalent to: "$\sum a_i v_i = \vec{0} \Rightarrow$ all $a_i = 0$".

Remark 1.33. The empty set $\emptyset$ is linearly independent. The singleton set $\{v\}$ is linearly independent if and only if $v \neq \vec{0}$.

Theorem 1.34. *Let V be a vector space. If $S_1 \subseteq S_2$ and S_1 is linearly dependent, then S_2 is linearly dependent.*

Proof. This is immediate from the definition. □

Theorem 1.35. *Let S be a linearly independent subset of V. Let $v \in V \setminus S$. Then $S \cup \{v\}$ is linearly dependent if and only if $v \in \operatorname{span}(S)$.*

Proof. "$\Rightarrow$". Write $a_1 u_1 + \ldots + a_n u_n + a_{n+1} v = 0$ with not all a_i equal to zero and $u_i \in S$.

Claim: $a_{n+1} \neq 0$.

Proof of claim: If $a_{n+1} = 0$, then at least one of $a_1, \ldots, a_n$ is not equal to zero and $a_1 u_1 + \ldots + a_n u_n = 0$. Contradiction to linear independence of S.

So, $a_{n+1} \neq 0$, and we can write $v = \frac{-a_1}{a_{n+1}} u_1 + \ldots + \frac{-a_n}{a_{n+1}} u_n$, qed.

"$\Leftarrow$". Write $v = a_1 u_1 + \ldots + a_n u_n$. Then $a_1 u_1 + \ldots + a_n u_n + (-1)v = 0$. Thus, $S \cup \{v\}$ is linearly dependent, qed. $\square$

1.6 Bases and dimension

Definition 1.36. Let V be a vector space. A *basis* β is a linearly independent subset of V which satisfies $\operatorname{span}(\beta) = V$.

Quiz 2:

Find the condition(s) for (a, b, c, d) to be in $\operatorname{span}\{(1, 0, -1, 1), (-1, 1, 0, -1), (0, 1, -1, 0), (1, 1, -2, 1)\}$.

Solution:

Make the Ansatz

$$a_1(1, 0, -1, 1) + a_2(-1, 1, 0, -1) + a_3(0, 1, -1, 0) + a_4(1, 1, -2, 1) = (a, b, c, d).$$

We have to determine for which values of a, b, c, d the above system is consistent, i.e., solvable. Reducing this system to echelon form yields:

$$\begin{aligned} a_1 - a_2 \qquad + a_4 &= a \\ a_2 + a_3 + a_4 &= b \\ 0 &= a + b + c \\ 0 &= d - a \end{aligned}$$

It is now clear that the system is consistent if and only if $a + b + c = 0$ and $a = d$. Done.

Now, let's discuss some problems from HW3:

1. a) subspace b) not a subspace c) subspace d) subspace

2. a) subspace b) not a subspace

4. Prove: non-empty, and closed under addition and scalar multiplication.

$\vec{0} \in W_1 \cap W_2$, so non-empty is ok.

Let $v \in W_1 \cap W_2$, so $v \in W_1$ and $v \in W_2$. Thus, $cv \in W_1$ and $cv \in W_2$. Consequently, $cv \in W_1 \cap W_2$.

Let $v, w \in W_1 \cap W_2$, so $v, w \in W_1$ and $v, w \in W_2$. Thus, $v + w \in W_1$ and $v + w \in W_2$. Consequently, $v + w \in W_1 \cap W_2$. Done.

6. Let $W_2 = \{\text{constant functions}\}$. Then clearly $W_1 \cap W_2 = \{\vec{0}\}$ and for an arbitrary function $g(x)$, we have $g(x) = (g(x) - g(5)) + g(5)$. If we write $f_1(x) = g(x) - g(5)$ and $f_2(x) = g(5)$ for all $x \in \mathbb{R}$, we have $g(x) = f_1(x) + f_2(x)$ with $f_1 \in W_1$ and $f_2 \in W_2$. Done.

8. "$\Leftarrow$". By symmetry, it suffices to treat the case $W_2 \subseteq W_1$. Then $W_1 \cup W_2 = W_1$, which is a subspace by assumption.

"$\Rightarrow$" Assume that ($W_2 \subseteq W_1$ or $W_1 \subseteq W_2$) is false. Then $\exists x \in W_1 : x \notin W_2$ and $\exists y \in W_2 : y \notin W_1$. Clearly, $x, y \in W_1 \cup W_2$ and by assumption $x + y \in W_1 \cup W_2$. By symmetry, we may assume $x + y \in W_1$, i.e., $\exists z \in W_1 : x + y = z$, which is equivalent to $\exists z \in W_1 : y = z - x$. Since $x, z \in W_1$, we also have $y \in W_1$. Contradiction.

Let's now continue with new material.

Theorem 1.37. *Let V be a vector space. Let $\beta = \{u_1, \ldots, u_n\}$ be a subset of V. Then*

$$\beta \textit{ is a basis } \Leftrightarrow \forall v \in V : \exists!\, a_1, \ldots, a_n \in F : v = a_1 u_1 + \ldots + a_n u_n.$$

(Recall that $\exists!$ means **unique** *existence.)*

Proof. "$\Rightarrow$". Spanning property is already known. We just have to prove uniqueness.

Let

$$a_1 u_1 + \ldots + a_n u_n = v = b_1 u_1 + \ldots + b_n u_n.$$

This implies

$$(a_1 - b_1) u_1 + \ldots + (a_n - b_n) u_n = 0.$$

Linear independence implies $a_1 - b_1 = 0, \ldots, a_n - b_n = 0$. Done.

"$\Leftarrow$". Spanning property is already known. To show lin. indep., just observe that

$$a_1 u_1 + \ldots + a_n u_n = 0$$

is solved by the trivial solution $a_1 = \ldots = a_n = 0$. However, by assumption, this is the only solution. Thus, we have established linear independence. Done. $\square$

Theorem 1.38. *Let V be a vector space. Let S be a finite subset of V with* $\text{span}(S) = V$*. Then there exists a subset of S which is a basis for V. In particular, V has a finite basis.*

Proof. We conduct this proof by induction over the cardinality of S.

If $\#S = 1$, then $S = \{v\}$, and S is clearly linearly independent (unless we are in a trivial cases).

Now, assume that we know the theorem for $\#S = n$. We have to prove it for $\#S = n + 1$.

If S is not a basis, then S is lin. dep. Claim: $\exists v \in S : V = \text{span}(S) = \text{span}(S \setminus \{v\})$. Proof of Claim: lin. dep. means that there is a linear combination

$$a_1u_1 + \ldots + a_nu_n = 0$$

with some $a_{i_0} \neq 0$. We can solve the above equation for u_{i_0}. It is now clear that a linear combination of the vectors $u_1, \ldots, u_n$ can be expressed as a linear combination of the vectors $u_1, \ldots, u_{i_0-1}, u_{i_0+1}, \ldots u_n$. Thus, letting $v = u_{i_0}$ establishes the Claim.

If we let $\tilde{S} = S \setminus \{u_{i_0}\}$, then we can apply the induction hypothesis to obtain that $\tilde{S}$ contains a basis. Since $\tilde{S} \subset S$, we can conclude that S contains a basis. Done. □

Example 1.39.

- Let $S = \{(1, 0), (1, 1), (2, 3)\}$. Observe that S spans $\mathbb{R}^2$, but S is lin. dep. After removing any one of the three vectors from S, we obtain a basis.

- Let $S = \{(1, 0), (0, 1), (0, 2)\}$. Observe that S spans $\mathbb{R}^2$, but S is lin. dep. After removing the second or third vector, we obtain a basis. However, removing the first vector does not yield a basis.

- Let $S = \{(2, -3, 5), (8, -12, 20), (1, 0, -2), (0, 2, -1), (7, 2, 0)\}$. Observe that S spans $\mathbb{R}^3$, but S is lin. dep.. Consider the span of the first vector. Obviously, the span remains unchanged after adding the second vector, so the second vector should be removed. The third vector is not a multiple of the first, so we keep it. A direct computation shows that the first, third and fourth vector are lin. indep. and span $\mathbb{R}^3$. The fifth can be disregarded.

Theorem 1.40 (Replacement Theorem)**.** *Let V be a vector space. Let $V = \operatorname{span}(G)$, where G is a subset of V of cardinality n. Let L be a linearly independent subset of V of cardinality m. Then the following holds.*

- $m \leq n$
- *there exists a subset $H \subseteq G$ of cardinality $n - m$ such that* $\operatorname{span}(L \cup H) = V$

Remark 1.41. A typical situation is for example $m = 2$ and $n = 5$, i.e., $L = \{v_1, v_2\}$ and $G = \{w_1, w_2, w_3, w_4, w_5\}$. The replacement theorem now says that there are two vectors in G that can be **replaced** with the two vectors from L such that the set obtained by the replacement still spans V. In other words, L can be injected into G and the result still spans V.

Corollary 1.42. *Let V be a vector space with a finite basis. Then all bases contain the same number of elements.*

Proof. Let β basis of cardinality m and γ basis of cardinality n. Since β lin. indep. and γ spans, we have $m \leq n$. By symmetry, we have $m = n$. □

Definition 1.43. A vector space is called *finite dimensional* if there exists a basis consisting of finitely many vectors. The unique cardinality of a basis of a finite dimensional vector space is called the *dimension* of V, denoted $\dim(V)$.

Example 1.44. $\dim(\mathbb{R}^n) = n, \dim(Mat_{m \times n}) = mn$. (Consider the standard bases.)

Here are two more important corollaries. The second one is based on the Replacement Theorem.

Corollary 1.45. *Let $S \subset V$. If $V = \operatorname{span}(S)$ and $\#S = \dim(V)$, then S is a basis.*

Proof. By Theorem 1.38, S contains a basis. This basis must have $\dim V = \#S$ elements. Thus, this basis is S itself. □

Corollary 1.46. *Let $S \subset V$. If S is lin. indep. and $\#S = \dim(V)$, then S is a basis.*

Proof. Take any basis G. Apply the Replacement Theorem with $L = S$. Since $\#G = \#S = \dim V$, we have $H = \emptyset$ and $V = \operatorname{span} L = \operatorname{span} S$. $\square$

Finally, let us prove the Replacement Theorem.

Proof of Replacement Theorem. For a fixed $n = \#G$, we do induction over $\#L = m$.

For $m = 0$, we have $L = \emptyset$. Take $H = G$. Done.

Induction step: "$m \to m + 1$".

Let $L = \{v_1, \dots, v_{m+1}\}$, let $\tilde{L} = \{v_1, \dots, v_m\}$.

Induction hypothesis $\Rightarrow \exists \tilde{H} = \{u_1, \dots, u_{n-m}\}$ such that $V = \operatorname{span}(\tilde{L} \cup \tilde{H})$.

Write

$$v_{m+1} = a_1 v_1 + \ldots + a_m v_m + b_1 u_1 + \ldots + b_{n-m} u_{n-m}. \tag{1}$$

Since L is lin. indep. we know that there exists i such that $b_i \neq 0$. Thus $n - m > 0$, i.e., $n \geq m + 1$. This proves the first part of the claim for $m + 1$.

It remains to show that if $b_1 \neq 0$, then $H = \{u_2, \dots, u_{n-m}\}$ works, i.e., $V = \operatorname{span}(L \cup H)$. Let $v \in V$ be arbitrary. Write

$$v = \alpha_1 v_1 + \ldots + \alpha_m v_m + \gamma_1 u_1 + \ldots + \gamma_{n-m} u_{n-m}. \tag{2}$$

If we solve (1) for u_1 and substitute into (2), we see that v can be written as a linear combination of $v_1, \dots, v_{m+1}, u_2, \dots, u_{n-m}$. Done. $\square$

Now, let us discuss the dimension of subspaces.

Theorem 1.47. *Let V be a vector space. Let W be a subspace of V. Assume $\dim V$ is finite. Then $\dim W \leq \dim V$ and equality holds if and only if $V = W$.*

Proof. This is immediate from the Replacement Theorem. $\square$

In the following examples, the task is to find a basis for (and the dimension of) the subspace W.

Example 1.48.

- Let $V = \mathbb{R}^3$. Let $W = \{(a_1, a_2, a_3) \mid a_1 + a_3 = 0 \text{ and } a_1 + a_2 - a_3 = 0\}$. Solving the system

$$a_1 + a_3 = 0 \text{ and } a_1 + a_2 - a_3 = 0$$

yields $W = \{(-t, 2t, t) \mid t \in \mathbb{R}\}$. Thus $\{(-1, 2, 1)\}$ is a basis, and the dimension of W is one.

- Let $V = \mathbb{R}^5$. Let $W = \{(a_1, a_2, a_3, a_4, a_5) \mid a_1 + a_3 + a_5 = 0 \text{ and } a_2 = a_4\}$. Solving the system

$$a_1 + a_3 + a_5 = 0 \text{ and } a_2 = a_4$$

yields

$$\begin{aligned} W &= \{(-a_3 - a_5, a_4, a_3, a_4, a_5) \mid a_3, a_4, a_5 \in \mathbb{R}\} \\ &= \{a_3(-1, 0, 1, 0, 0) + a_4(0, 1, 0, 1, 0) + a_5(-1, 0, 0, 0, 1) |\ a_3, a_4, a_5 \in \mathbb{R}\} \end{aligned}$$

Thus $\{(-1, 0, 1, 0, 0), (0, 1, 0, 1, 0), (-1, 0, 0, 0, 1)\}$ is a basis for W, and the dimension of W is three.

2 Linear transformations and matrices

2.1 Linear transformations, null spaces, and ranges

Definition 2.1. Let V, W be vector spaces over the same field F. We call a function $T : V \to W$ a *linear transformation* from V to W if

i. $\forall x, y \in V : T(x + y) = T(x) + T(y)$

ii. $\forall c \in F \forall x \in V : T(cx) = cT(x)$

Remark 2.2. We say T is *linear* for short.

Properties 2.3.

- $T(\vec{0}) = \vec{0}$
- $T(x - y) = T(x) - T(y)$
- $T(a_1 v_1 + \ldots + a_n v_n) = a_1 T(v_1) + \ldots + a_n T(v_n)$

Example 2.4.

- $T(a_1, a_2) = (2a_1 + a_2, a_1)$ (Check it!)
- $T : \mathbb{R}^5 \to \mathbb{R}^7, T(a_1, \ldots, a_5) = (a_1, a_2, 0, a_7, 0, 0, a_1)$

- $T : Mat_{m\times n} \to Mat_{n\times m}, T(A) = A^T$
- $T : C^\infty(\mathbb{R}) \to C^\infty(\mathbb{R}), T(f) = \frac{df}{dx}$
- $T : C^\infty(\mathbb{R}) \to C^\infty(\mathbb{R}), T(f) = (x \mapsto \int_0^x f(t)dt)$

Conducted **Quiz 3**. See the end of this file for a pdf with solution.

Definition 2.5. LetV, W be vector spaces. Let $T : V \to W$ linear. We define the *null space* (aka *kernel*) of T to be

$$N(T) = \{x \in V \ : \ T(x) = \vec{0}\}.$$

Remark 2.6. Recall that the range of T is

$$R(T) = \{T(x) \ : \ x \in V\}.$$

Example 2.7. Let $T : \mathbb{R}^3 \to \mathbb{R}^2$, $T(a_1, a_2, a_3) = (a_1 - a_2, 2a_3)$. To find the null space, set

$$T(a_1, a_2, a_3) = (0, 0) \Leftrightarrow a_1 - a_2 = 0 \text{ and } 2a_3 = 0.$$

The solution of the above system of two equations in three variables is

$$N(T) = \{(t, t, 0) | t \in \mathbb{R}\}.$$

Moreover, it is clear that T is onto, so $R(T) = \mathbb{R}^2$.

Theorem 2.8. *Let V, W be vector spaces and $T : V \to W$ linear. Then*

i. $N(T)$ is a subspace of V

ii. $R(T)$ is a subspace of W

Proof. i. We just have to check closedness. If $T(x) = 0$ and $T(y) = 0$, then $T(ax + by) = aT(x) + bT(y) = 0 + 0 = 0$. Done.

ii. Again, just closedness. If $T(x) = v_1$ and $T(y) = v_2$, then $av_1 + bv_2 = aT(x) + bT(y) = T(ax + by)$. Done. □

Theorem 2.9. *Let V, W be vector spaces and $T : V \to W$ linear. Let $\{v_1, \ldots, v_n\}$ be a basis for V. Then $R(T) = \text{span}\{T(v_1), \ldots, T(v_n)\}$.*

Proof. Let $v = a_1v_1 + \ldots + a_nv_n$. Then $T(v) = T(a_1v_1 + \ldots + a_nv_n) = a_1T(v_1) + \ldots + a_nT(v_n) \in \operatorname{span}\{T(v_1), \ldots, T(v_n)\}$. Done. □

Example 2.10. Problem: Find (a basis for) $R(T)$ when $T : \mathbb{R}^3 \to \mathbb{R}^3, T(a_1, a_2, a_3) = (a_1 - 2a_2, a_2 + a_3, 2a_1 + a_2 + 5a_3)$.

First, we note $T(1,0,0) = (1,0,2)$, $T(0,1,0) = (-2,1,1)$, $T(0,0,1) = (0,1,5)$. Thus, according to Theorem 2.9, $R(T) = \operatorname{span}\{(1,0,2),(-2,1,1),(0,1,5)\}$. Now, note that twice the first vector plus the second equals the third, so $R(T) = \operatorname{span}\{(1,0,2),(-2,1,1)\}$. The set $\{(1,0,2),(-2,1,1)\}$ is clearly a basis for $R(T)$, and $\dim R(T) = 2$.

Note that $N(T)$ is easily computed to be one-dimensional, and $\dim N(T) + \dim R(T) = 3$.

Definition 2.11. Let V, W be vector spaces and $T : V \to W$ linear. If $N(T), R(T)$ are finite dimensional, then let

$$\operatorname{nullity}(T) = \dim N(T), \operatorname{rank}(T) = \dim R(T).$$

Theorem 2.12 (Dimension Theorem)**.** *Let V, W be vector spaces and $T : V \to W$ linear. If V is finite-dimensional, then*

$$\operatorname{nullity}(T) + \operatorname{rank}(T) = \dim V.$$

Proof. Let $\{v_1, \ldots, v_k\}$ be a basis for $N(T)$. In particular, $k = \operatorname{nullity}(T)$. Let $n = \dim V$. The Replacement Theorem implies that there are vectors $v_{k+1}, \ldots, v_n \in V$ such that $\{v_1, \ldots, v_n\}$ is a basis for V.

Claim: $\{T(v_{k+1}), \ldots, T(v_n)\}$ is a basis for $R(T)$.

Spanning: Let $v = a_1v_1 + \ldots + a_nv_n \in V$ arbitrary. Then

$$T(v) = a_1T(v_1) + \ldots + a_kT(v_k) + a_{k+1}T(v_{k+1}) + \ldots + a_nT(v_n).$$

However, the first k summands are zero due to $v_1, \ldots, v_k \in N(T)$.

Lin. indep.: Write

$$b_{k+1}T(v_{k+1}) + \ldots + b_nT(v_n) = 0.$$

We have to conclude $b_{k+1} = \ldots = b_n = 0$. To do this, note that $b_{k+1}T(v_{k+1}) + \ldots + b_nT(v_n) = T(b_{k+1}v_{k+1} + \ldots + b_nv_n)$, i.e., $b_{k+1}v_{k+1} + \ldots + b_nv_n \in N(T)$. Thus, there exist $a_1, \ldots, a_k$:

$$a_1v_1 + \ldots + a_kv_k = b_{k+1}v_{k+1} + \ldots + b_nv_n.$$

Since $\{v_1, \ldots, v_n\}$ is a basis for V, this is only possible if

$$a_1 = \ldots = a_k = b_{k+1} = \ldots = b_n = 0.$$

□

Theorem 2.13. *Let V, W vector spaces. Let $T : V \to W$ linear. Then T is one-to-one if and only if $N(T) = \{\vec{0}\}$.*

Proof. $\Rightarrow$. Saw: $T(\vec{0}) = \vec{0}$. Since T is one-to-one, this implies $N(T) = \{\vec{0}\}$.

$\Leftarrow$. Assume $T(x) = T(y)$. Then $T(x) - T(y) = 0$. By linerity of T, $T(x - y) = \vec{0}$. By assumption, $x - y = \vec{0}$. Done. □

Theorem 2.14. *Let V, W vector spaces with $\dim V = \dim W$. Let $T : V \to W$linear. Then the following are equivalent.*

i. *T is one-to-one*

ii. *T is onto*

iii. $\operatorname{rank} T = \dim V$

Proof. Immediate from the Dimension Theorem. □

Theorem 2.15. *Let V, W vector spaces. Let $\{v_1, \ldots, v_n\}$ be a basis for V. Let $w_1, \ldots, w_n$ be a list of arbitrary vectors in W. Then there exists a unique $T : V \to W$ linear such that $T(v_i) = w_i$ for all $i = 1, \ldots, n$.*

Proof. Recall that an arbitrary $v \in V$ can be written as $v = \sum_i a_i v_i$ with unique coefficients a_i. Then set $T(v) = \sum_i a_i w_i$. It is easy to check that this defines a well-defined linear map as required in the Theorem. Uniqueness is also clear. □

Corollary 2.16. *Let V, W vector spaces. Let $U, T : V \to W$ linear with $U(v_i) = T(v_i)$ on a basis $\{v_1, \ldots, v_n\}$ for V. Then $U = T$.*

Remark 2.17. If $T : V \to W$ linear and $\dim V < \dim W$, then T cannot be onto. If $T : V \to W$ linear and $\dim V > \dim W$, then T cannot be one-to-one.

Remark 2.18. On the (infinite-dimensional) vector space of sequences $(a_i)_{i \in \mathbb{N}}$ of real numbers, there exist maps which are one-to-one, but not onto. For example, $T(a_1, a_2, a_3, \ldots) = (0, a_1, a_2, a_3, \ldots)$. On a finite dimensional vector space, this is impossible due to Theorem 2.14.

2.2 The matrix representation of a linear transformation

Definition 2.19. Let V be a finite dimensional vector space. An *ordered basis* for V is a basis endowed with a specific order.

Example 2.20. As ordered bases,

$$\{(1,0,0),(0,1,0),(0,0,1)\} \neq \{(0,1,0),(1,0,0),(0,0,1)\}.$$

Definition 2.21. Let $\beta = \{u_1, \ldots, u_n\}$ ordered basis for V. We saw earlier:

$$\forall x \in V \ \exists! a_1, \ldots, a_n : x = a_1 u_1 + \ldots + a_n u_n.$$

Write

$$[x]_\beta = (a_1, \ldots, a_n)$$

for the *coordinate vector of* x relative to β. In particular, $[u_i]_\beta = e_i$.

Definition 2.22. Take V with $\beta = \{v_1, \ldots, v_n\}$, W with $\gamma = \{w_1, \ldots, w_m\}$. Let $T : V \to W$ linear. Write

$$T(v_j) = \sum_{i=1}^{m} a_{ij} w_i$$

for $j = 1, \ldots, n$. Call the matrix (a_{ij}) the *matrix representation of* T with respect to β and γ. When $V = W$ and $\beta = \gamma$, write $A = [T]_\beta$.

Remark 2.23. The key fact to remember is that the j-th column of the matrix representation is $[T(v_j)]_\gamma$.

Example 2.24. (a) Let $T : \mathbb{R}^2 \to \mathbb{R}^3$ be given by $T(a_1, a_2) = (a_1 + 3a_2, 0, 2a_1 - 4a_2)$. Let β and γ be the respective standard bases. Then

$$T(1,0) = (1,0,2), \quad T(0,1) = (3,0,-4).$$

Thus,

$$[T]_\beta^\gamma = \begin{pmatrix} 1 & 3 \\ 0 & 0 \\ 2 & -4 \end{pmatrix}.$$

(b) Same map, but with $\gamma' = \{e_2, e_1, e_3\}$:

$$[T]_\beta^{\gamma'} = \begin{pmatrix} 0 & 0 \\ 1 & 3 \\ 2 & -4 \end{pmatrix}.$$

(c) Same map as in (a), but with $\beta' = \{e_2, e_1\}$:

$$[T]_{\beta'}^{\gamma} = \begin{pmatrix} 0 & 0 \\ 3 & 1 \\ -4 & 2 \end{pmatrix}.$$

(d) Let $T : \mathbb{R}^2 \to \mathbb{R}^3$ be given by $T(a_1, a_2) = (a_1 - a_2, a_1, 2a_1 + a_2)$. Let β be the standard basis and $\gamma = \{(1, 1, 0), (0, 1, 1), (2, 2, 3)\}$. By solving a system of linear equations, we find

$$T(1, 0) = (1, 1, 2) = -\frac{1}{3}(1, 1, 0) + \frac{2}{3}(2, 2, 3),$$

and

$$T(0, 1) = (-1, 0, 1) = -(1, 1, 0) + (0, 1, 1).$$

Thus,

$$[T]_{\beta}^{\gamma} = \begin{pmatrix} -\frac{1}{3} & -1 \\ 0 & 1 \\ \frac{2}{3} & 0 \end{pmatrix}.$$

Definition 2.25. Let $U, T : V \to W$ be linear. Then

$$(U + T)(x) = U(x) + T(x)$$

and

$$(cT)(x) = cT(x).$$

Theorem 2.26. *Let V, W be given vector spaces. The set of all linear transformations $V \to W$ is a vector space with $+$ and $\cdot$ defined as above. Write $\mathcal{L}(V, W)$ for this vector space.*

Proof. Check the axioms! □

Definition 2.27. Let $U, T : V \to W$ linear. Then

i. $[U + T]_{\beta}^{\gamma} = [U]_{\beta}^{\gamma} + [T]_{\beta}^{\gamma}$

ii. $[aT]_{\beta}^{\gamma} = a[T]_{\beta}^{\gamma}$

Proof. Write $U(v_j) = \sum a_{ij} w_i$, $T(v_j) = \sum b_{ij} w_i$. Then

$$(U + T)(v_j) = U(v_j) + T(v_j) = \sum a_{ij} w_i + \sum b_{ij} w_i = \sum (a_{ij} + b_{ij}) w_i.$$

Thus, the ij entry of $[U + T]_{\beta}^{\gamma}$ is $a_{ij} + b_{ij}$. □

2.3 Composition of linear transformations and matrix multiplication

Theorem 2.28. *Let V, W, Z be vector spaces over the same field. Let $T : V \to W$ and $U : W \to Z$ be linear. Then $U \circ T : V \to Z$ is linear.*

Proof.

$$(U \circ T)(ax+by) = U(T(ax+by)) = U(aT(x)+bT(y)) = aU(T(x))+bU(T(y))$$

$$= a(U \circ T)(x) + b(U \circ T)(y).$$

□

Theorem 2.29. *Let $U, S, T : V \to V$ linear. Then*

- $U \circ (S + T) = U \circ S + U \circ T$
- $(U + S) \circ T = U \circ T + S \circ T$
- $U \circ (S \circ T) = (U \circ S) \circ T$
- $\mathrm{id} \circ U = U \circ \mathrm{id} = U$
- $a(U \circ S) = (aU) \circ S = U \circ (aS)$

Now, let us investigate the matrix of a composition of linear transformations. Let $T : V \to W$, $U : W \to Z$. Let $\alpha = \{v_j\}, \beta = \{w_k\}, \gamma = \{z_i\}$ be the corresponding ordered basis, in alphabetical order. Let $[T]_\alpha^\beta = B$, $[U]_\beta^\gamma = A$. Then

$$(U \circ T)(v_j) = U(T(v_j)) = U(\sum_k b_{kj} w_k) = \sum_k b_{kj} U(w_k) =$$

$$\sum_k b_{kj} (\sum_i a_{ik} z_i) = \sum_i (\sum_k a_{ik} b_{kj}) z_i.$$

Consequently, if $C = [U \circ T]_\alpha^\gamma$, $c_{ij} = \sum_k a_{ik} b_{kj}$.

Definition 2.30. Let A be an $m \times n$ matrix and B an $n \times p$ matrix. Define the *matrix product* of A and B to be the $m \times p$ matrix given by

$$(AB)_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}.$$

We have just established the following theorem.

Theorem 2.31. *Let $T : V \to W$, $U : W \to Z$. Then* $[U \circ T]_\alpha^\gamma = [U]_\beta^\gamma [T]_\alpha^\beta$.

Quiz 4:

1. (7 points) Let $T_1 : \mathbb{R}^2 \to \mathbb{R}^2$, $T_1(a_1, a_2) = (a_1 + 2a_2, 2a_1 + 4a_2)$. Let $\beta = \{(1, -1), (0, 1)\}$ and $\gamma = \{(1, 2), (1, 1)\}$. Compute $[T_1]_\beta^\gamma$.

2. (3 points) Let $T_2 : \mathrm{Mat}_{2\times 2}(\mathbb{R}) \to \mathrm{Mat}_{2\times 2}(\mathbb{R}), T_2\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix}\right) = \begin{pmatrix} b & c \\ d & a \end{pmatrix}$. Let $\beta = \left\{\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},\right.$
$\left.\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\right\}$. Compute $[T_2]_\beta^\beta$.

Solution:

1. $T_1(1, -1) = (-1, -2) = (-1)(1, 2) + 0(1, 1)$, $T_1(0, 1) = (2, 4) = 2(1, 2) + 0(1, 1)$. Thus,

$$[T_1]_\beta^\gamma = \begin{pmatrix} -1 & 2 \\ 0 & 0 \end{pmatrix}.$$

2. Observe that

$$T_2 \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix},$$

$$T_2 \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix},$$

$$T_2 \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},$$

$$T_2 \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

Thus,

$$[T_2]_\beta^\beta = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

Theorem 2.32. *Let $A \in \mathrm{Mat}_{m\times n}$, $B, C \in \mathrm{Mat}_{n\times p}$, $D, E \in \mathrm{Mat}_{q\times m}$. Then*

i. $A(B + C) = AB + AC$

ii. $(D + E)A = DA + EA$

iii. $a(AB) = A(aB) = (aA)B$

iv. $I_m A = A I_n$

Theorem 2.33. *Let $T : V \to W$ linear. Then $\forall v \in V$:*

$$[T(v)]_\gamma = [T]_\beta^\gamma [v]_\beta.$$

Proof. I find the proof in the book somewhat uninformative. Here is an alternative. By linearity, it is clear that we have to check the identity only on the elements of the basis $\beta = \{v_1, \ldots, v_n\}$.

First, note that

$$[T(v_j)]_\gamma = [a_{1j}w_1 + \ldots + a_{mj}w_m]_\gamma = (a_{1j}, \ldots, a_{mj}).$$

On the other hand,

$$[T]_\beta^\gamma [v_j]_\beta = [T]_\beta^\gamma e_j = (a_{1j}, \ldots, a_{mj}).$$

Done. □

Definition 2.34. Let $A \in \mathrm{Mat}_{m\times n}(F)$. Define $L_A : F^n \to F^m$ by $L_A(x) = Ax$. L_A is the left-multiplication transformation given by the matrix A.

Theorem 2.35. *Let β, γ be the standard ordered bases.*

i. $L_A : F^n \to F^m$ *is linear.*

ii. $[L_A]_\beta^\gamma = A$

iii. $L_A = L_B \Leftrightarrow A = B$

iv. $L_{A+B} = L_A + L_B$, $L_{aA} = aL_A$

v. *For* $T : F^n \to F^m$, $T = L_{[T]_\beta^\gamma}$

2.4 Invertibility and Isomorphisms

Definition 2.36. Let V, W be vector spaces. Let $T : V \to W$ linear. We define $U : W \to V$ to be the *inverse* of T if $T \circ U = \mathrm{id}_W$ and $U \circ T = \mathrm{id}_V$. If T has an inverse, T is called *invertible*.

Remark 2.37. If T has an inverse, the inverse is certainly unique. We write T^{-1} for it. Recall that in general, for sets A, B, a function $f : A \to B$ is invertible if and only if f is one-to-one and onto.

Theorem 2.38. *If $T : V \to W$ is linear and invertible, then the inverse T^{-1} is linear also.*

Proof. Let $w_1, w_2 \in W$, $a \in F$. Since T is in particular onto, for some v_1, v_2 the following holds:

$$T^{-1}(w_1 + w_2) = T^{-1}(T(v_1) + T(v_2)) = T^{-1}(T(v_1 + v_2)) =$$
$$v_1 + v_2 = T^{-1}(w_1) + T^{-1}(w_2).$$

Moreover,

$$T^{-1}(aw_1) = T^{-1}(aT(v_1)) = T^{-1}(T(av_1)) = av_1 = aT^{-1}(w_1).$$

□

Lemma 2.39. *Let $T : V \to W$ linear and invertible. Let V, W be finite dimensional. Then*

$$\dim V = \dim W.$$

Proof. The dimension formula says $\mathrm{rank}(T) = \dim V$ since $\mathrm{nullity}(T) = 0$ due to T being one-to-one. But $\mathrm{rank}(T) \leq \dim W$ always. Thus, $\dim V \leq \dim W$. By symmetry, we are done. □

Definition 2.40. Let A be an $n \times n$ matrix. Say A is *invertible* if there exists an $n \times n$ matrix B such that

$$AB = I_n = BA.$$

Remark 2.41. Such a B is unique, if it exists. Reason: Let B, C be two matrices with the above property. Then

$$C = CI_n = C(AB) = (CA)B = I_nB = B.$$

Thus, we can write A^{-1} for B.

Theorem 2.42. *Let V, W be finite dimensional vector spaces with ordered bases β, γ respectively. Let $T : V \to W$ linear. Then T is invertible if and only if $[T]_\beta^\gamma$ is an invertible matrix. Moreover, in this case, $[T^{-1}]_\gamma^\beta = ([T]_\beta^\gamma)^{-1}$.*

Proof. Based on:

$$I_n = [\mathrm{id}_V]_\beta^\beta = [T^{-1} \circ T]_\beta^\beta = [T^{-1}]_\gamma^\beta [T]_\beta^\gamma.$$

See the textbook for complete details. □

Definition 2.43. Let V, W be vector spaces. Say V is *isomorphic* to W if and only if there exists $T : V \to W$ linear and invertible. Such a T is called an *isomorphism* from V onto W.

Remark 2.44. Being isomorphic is an equivalence relation on the set of vector spaces (over a given field).

Theorem 2.45. *Let V, W be finite dimensional vector spaces (over the same field). Then V is isomorphic to W if and only if $\dim V = \dim W$.*

Proof. We have already seen $\Rightarrow$. So it remains to prove $\Leftarrow$. Let $\beta = \{v_1, \ldots, v_n\}$ and $\gamma = \{w_1, \ldots, w_n\}$. Saw earlier that $T(v_j) = w_j$ for all j defines a linear transformation $V \to W$. It is onto because $\mathrm{Range}(T) = \mathrm{span}\{T(v_1), \ldots, T(v_n)\} = \mathrm{span}\{w_1, \ldots, w_n\} = W$. It is one-to-one because of the dimension formula. □

Corollary 2.46. *Let V be a vector space over F. Then V is isomorphic to F^n if and only $\dim V = n$.*

Let us recall the method you learned in your introductory linear algebra course to find the inverse of a matrix. Take your matrix, say,

$$A = \begin{pmatrix} 1 & 4 \\ 2 & 9 \end{pmatrix}.$$

Apply elementary row operations to A to obtain the identity matrix.

$$\begin{pmatrix} 1 & 4 \\ 2 & 9 \end{pmatrix} \to \begin{pmatrix} 1 & 4 \\ 0 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Then use the same elementary row operations to convert identity matrix to A^{-1}:

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix} \to \begin{pmatrix} 9 & -4 \\ -2 & 1 \end{pmatrix} = A^{-1}.$$

Also, recall the formula

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - cb} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

in case $ad - bc \neq 0$.

Theorem 2.47. *Let V, W be finite dimensional vector spaces. Let $\dim V = n$, $\dim W = m$. Let β, γ be ordered bases for V and W. Then $\Phi : \mathcal{L}(V, W) \to Mat_{m \times n}(F), T \mapsto [T]_\beta^\gamma$ is an isomorphism.*

Proof. We saw earlier that Φ is linear. It is onto, because for an arbitrary matrix (a_{ij}), define $T(v_j) = \sum_i a_{ij} w_i$, where $\beta = \{v_j\}$ and $\gamma = \{w_i\}$. Then $\Phi(T) = (a_{ij})$. It is also one-to-one, because obviously only the zero transformation maps to the zero matrix. □

2.5 The change of coordinate matrix

Theorem 2.48. *Let β, β' be ordered bases for the finite-dimensional vector space V. Let $Q = [\mathrm{id}_V]_{\beta'}^{\beta}$. Then*

i. Q is invertible

ii. $\forall v \in V : [v]_\beta = Q[v]_{\beta'}$.

Proof. (i) Saw earlier: id_V invertible implies Q is invertible. Concerning (ii), for all $v \in V$:

$$[v]_\beta = [\mathrm{id}_V(v)]_\beta = [\mathrm{id}_V]_{\beta'}^{\beta} [v]_{\beta'} = Q[v]_{\beta'}.$$

□

Remark 2.49. Q is called the *change of coordinates matrix.* It changes β'-coordinates to β-coordinates.

Example 2.50. Let $\beta = \{(1,1),(1,-1)\}$, $\beta' = \{(2,4),(3,1)\}$. Then

$$\mathrm{id}(2,4) = (2,4) = 3(1,1) - (1,-1)$$

and

$$\mathrm{id}(3,1) = (3,1) = 2(1,1) + (1,-1).$$

Thus,

$$Q = \begin{pmatrix} 3 & 2 \\ -1 & 1 \end{pmatrix}.$$

Now, we answer the following question: Let $T : V \to V$ linear. What is the relationship between $[T]_\beta$ and $[T]_{\beta'}$?

Theorem 2.51. *Let $T : V \to V$ linear with V finite dimensional vector space. Let β, β' be ordered bases for V. Let $Q = [\mathrm{id}_V]_{\beta'}^{\beta}$. Then*

$$[T]_{\beta'} = Q^{-1}[T]_\beta Q.$$

Proof.

$$\begin{aligned} Q[T]_{\beta'} &= [\mathrm{id}_V]_{\beta'}^{\beta}[T]_{\beta'}^{\beta'} = [\mathrm{id}_V \circ T]_{\beta'}^{\beta} = [T \circ \mathrm{id}_V]_{\beta'}^{\beta} \\ &= [T]_{\beta}^{\beta}[\mathrm{id}_V]_{\beta'}^{\beta} = [T]_{\beta}^{\beta} Q \end{aligned}$$

□

Example 2.52. Let $T(a,b) = (3a-b, a+3b)$. Let again $\beta = \{(1,1),(1,-1)\}$, $\beta' = \{(2,4),(3,1)\}$. Then

$$[T]_\beta = \begin{pmatrix} 3 & 1 \\ -1 & 3 \end{pmatrix}.$$

We saw earlier:

$$Q = \begin{pmatrix} 3 & 2 \\ -1 & 1 \end{pmatrix}.$$

Thus,

$$[T]_{\beta'} = Q^{-1}[T]_\beta Q = \begin{pmatrix} 4 & 1 \\ -2 & 2 \end{pmatrix}.$$

An interesting special case of Theorem 2.51 is when $V = \mathbb{R}^n$, β is the standard basis, $\beta' = \{v_1, \ldots, v_n\}$ a general ordered basis, and $T = L_A$ with $A \in Mat_{n\times n}(\mathbb{R})$. Specializing Theorem 2.51 to this case yields $[L_A]_{\beta'} = Q^{-1}AQ$ where Q is the matrix whose columns are the vectors $v_1, \ldots, v_n$.

Definition 2.53. Let $A, B \in Mat_{n\times n}(F)$. We say B is *similar* to A if and only if there exists an invertible $Q \in Mat_{n\times}(F)$: $B = Q^{-1}AQ$.

Remark 2.54. Being similar is an equivalence relation on $Mat_{n\times n}(F)$.

Now, in preparation for the midterm exam on Oct. 22, here are a few solutions for problems from the first sample exam.

2. (a) (10 points) Let $W_1 = \{(a_1, a_2, a_1 - a_2)|a_1, a_2 \in \mathbb{R}\} \subset \mathbb{R}^3$. Let $W_2 = \{(b, -b, 0)|b \in \mathbb{R}\} \subset \mathbb{R}^3$. Is $W_1 \oplus W_2 = \mathbb{R}^3$? Justify your answer.

First, note that W_1, W_2 are indeed subspaces of $\mathbb{R}^3$. Then, note that $(x, y, z) \in W_1 \cap W_2$ is equivalent to

$$z = x - y,\ z = 0,\ x = -y.$$

The set of solutions for this system of three equations in three variables is $\{(0, 0, 0)\}$. Thus, it remains to show that $W_1 + W_2 \in \mathbb{R}^3$. Since $\dim W_1 = 2$ and $\dim W_2 = 1$ it is clear that we can pick a total of three linearly independent vectors from W_1 and W_2. These three vectors span $\mathbb{R}^3$, so we are done.

4. (a) (15 points) Find bases for the kernel and range of $T : \mathbb{R}^5 \to \mathbb{R}^4$, $(a_1, a_2, a_3, a_4, a_5) \mapsto (a_1 + a_3 - a_4 + a_5, -a_1 + a_2 + a_4, -a_1 + 2a_4, -a_1 + a_2 + a_3 + 2a_4 + a_5)$.

First, let's find a basis for the kernel:

$$T(a_1, a_2, a_3, a_4, a_5) = (0, 0, 0, 0)$$

is equivalent to

$$a_1+a_3-a_4+a_5 = 0,\ -a_1+a_2+a_4 = 0,\ -a_1+2a_4 = 0,\ -a_1+a_2+a_3+2a_4+a_5 = 0.$$

The solution of this system is

$$N(T) = \{(2a_4, a_4, -a_4 - a_5, a_4, a_5)|a_4, a_5 \in \mathbb{R}\}.$$

Thus, a basis for $N(T)$ is $\{(2, 1, -1, 1, 0), (0, 0, -1, 0, 1)\}$. The dimension formula now says $\operatorname{rank}(T) = \dim \mathbb{R}^5 - \operatorname{nullity}(T) = 5 - 2 = 3$. The three vectors

$$\begin{aligned} T(1, 0, 0, 0, 0) &= (1, -1, -1, -1) \\ T(0, 0, 0, 0, 1) &= (1, 0, 0, 1) \end{aligned}$$

$$T(0,0,0,1,0) \;=\; (-1,1,2,2)$$

are obviously linearly independent, so they form a basis for the range of T.
(b) (10 points) Give a complete and explicit list of all linear transformations $T : \mathbb{R}^2 \to \mathbb{R}^2$ satisfying both $T(1,1) = (1,2)$ and $T(1,0) = (3,1)$. For every T on your list, compute $T(0,1)$.

T is given with respect to the standard basis by a matrix

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

The information $T(1,1) = (1,2)$ and $T(1,0) = (3,1)$ is equivalent to

$$a + b = 1,\ c + d = 2,\ a = 3,\ c = 1,$$

which in turn is equivalent to $a = 3,\ c = 1,\ b = -2, d = 1$. Thus, there is a unique such T, which is given by the matrix we determined. $T(0,1) = (-2,1)$.

6. (a) (10 points) Let $T : V \to W$ be a linear transformation. Assume that T is one-to-one. Let S be a subset of V. Prove that S is linearly independent if and only if $T(S)$ is linearly independent.

"$\Rightarrow$". Let $T(v_1), \ldots, T(v_n) \in T(S)$. We have to show that, for $v_i \in S$

$$a_1 T(v_1) + \ldots + a_n T(v_n) = \vec{0}$$

only has the trivial solution $a_1 = \ldots = a_n = 0$. To this end, note that the above equation is equivalent to

$$T(a_1 v_1 + \ldots, a_n v_n) = \vec{0}.$$

Since T is one-to-one, this is equivalent to

$$a_1 v_1 + \ldots, a_n v_n = \vec{0}.$$

Since S is linearly independent, this gives $a_1 = \ldots = a_n = 0$.

"$\Leftarrow$" is completely analogous.

(b) (10 points) Let $T : V \to W$ be a linear transformation. Suppose $\beta = \{v_1, \ldots, v_n\}$ is a basis for V and T is one-to-one and onto. Prove that $T(\beta) = \{T(v_1), \ldots, T(v_n)\}$ is a basis for W.

Since T is one-to-one and onto, we have $\dim V = \dim W$. Thus $T(\beta)$ has the correct cardinality, and we only have to prove that it is linearly independent. This follows from part (a).

2.6 Dual Spaces

Definition 2.55. Let V be a vector space over the field F. A linear transformation $f : V \to F$, where F is considered as a vector space over itself, is called a *linear functional.*

Example 2.56.

i. $f : \mathbb{R}^n \to \mathbb{R}$, $(x_1, \ldots, x_n) \mapsto x_1$

ii. $f : \mathbb{R}^2 \to \mathbb{R}$, $(x_1, x_2) \mapsto 2x_1 - 3x_2$

iii. $f : \mathbb{R}^n \to \mathbb{R}$, $(x_1, \ldots, x_n) \mapsto x_1 + \ldots + x_n$

iv. $\mathrm{tr} : Mat_{n\times n} \to \mathbb{R}$, $A \mapsto \mathrm{tr}(A)$

v. $\mathrm{eval}_5 : \{g : \mathbb{R} \to \mathbb{R}\} \to \mathbb{R}$, $g \mapsto g(5)$.

vi. Let V be a vector space over F, $\beta = \{v_1, \ldots, v_n\}$ ordered basis. Let $f_i : V \to F$ be defined by $f(v) =$ the i-th coordinate of v with respect to β. In other words, if $[v]_\beta = (a_1, \ldots, a_n)$, then $f_i(v) = a_i$. In particular, $f_i(v_j) = \delta_{ij}$.

Definition 2.57. For a vector space V over F, let the *dual space* V^* be $\mathcal{L}(V, F)$.

Remark 2.58. $\dim V^* = \dim \mathcal{L}(V, F) = \dim V \cdot \dim F = \dim V$. Consequently, V and V^* are isomorphic as vector spaces over F. (But not canonically so.)

Theorem 2.59. *Let V be a vector space and $\beta = \{v_1, \ldots, v_n\}$ ordered basis. Let f_i be the i-th coordinate function with respect to β as in Example 2.56. Then $\beta^* = \{f_1, \ldots, f_n\}$ is an ordered basis for V^* and for all $f \in V^*$,*

$$f = \sum_{i=1}^{n} f(v_i) f_i.$$

Proof. We only have to prove the displayed equality. Let $v = \sum_i a_i v_i$. Then the left hand side, applied to v, can be expanded to

$$\sum_i a_i f(v_i).$$

The right hand side becomes

$$\sum_i f(v_i) f_i(\sum a_i v_i) = \sum_i f(v_i) a_i.$$

□

Definition 2.60. We call β^* the *dual basis* of β.

Example 2.61. Let $V = \mathbb{R}^2$. Let $\beta = \{(2,1),(3,1)\}$. Task: Find f_1, f_2.

Answer: Write $f_1(x,y) = ax + by$. Then f_1 is determined by the system:

$$f_1(2,1) = 1,\ f_1(3,1) = 0 \Leftrightarrow 2a + b = 1,\ 3a + b = 0.$$

Solving for a, b yields $a = -1,\ b = 3$, i.e., $f_1(x,y) = -x + 3y$.

For f_2:

$$f_2(2,1) = 0,\ f_2(3,1) = 1 \Leftrightarrow 2a + b = 0,\ 3a + b = 1,$$

which yields $f_2(x,y) = x - 2y$.

Example 2.62. In the other direction, assume we are given $f_1(x,y) = -x + 3y$ and $f_2(x,y) = x - 2y$. How do we find $\beta = \{v_1, v_2\}$?

Answer: Write $v_1 = (a,b)$. It is determined by the system

$$f_1(a,b) = 1,\ f_2(a,b) = 0 \Leftrightarrow -a + 3b = 1,\ a - 2b = 0,$$

which yields $a = 2, b = 1$.

Write $v_2 = (a,b)$. It is determined by the system

$$f_1(a,b) = 0,\ f_2(a,b) = 1 \Leftrightarrow -a + 3b = 0,\ a - 2b = 1,$$

which yields $a = 3, b = 1$.

Next, let us define the *transpose* T^t of a linear transformation T.

Theorem 2.63. *Let V, W be vector spaces over the same field F with ordered bases β, γ respectively. Let $T : V \to W$ be linear. Then $T^t : W^* \to V^*$ defined by*

$$T^t(g) = g \circ T,$$

is a linear transformation with $[T^t]_{\gamma^*}^{\beta^*} = ([T]_\beta^\gamma)^t$.

Proof. It is easy to see that T^t is well-defined and linear. The identity is verified simply by careful checking. □

Let us finish this section with a discussion of the *canonical* isomorphism between a vector space and its double dual. Here is a basic observation:

Let V be a finite dimensional vector space. Any vector $v \in V$ defines a functional $\hat{v} : V^* \to F$ by letting $\hat{v}(f) = f(v)$. This linear functional is also called the *evaluation functional with respect to* V. Note $\hat{v} \in (V^*)^*$.

Theorem 2.64. *Let V be a finite dimensional vector space. Then $\psi : V \to (V^*)^*, v \mapsto \hat{v}$ is a (canonical) isomorphism.*

Proof. The following shows that ψ is linear:

$$\psi(av + bw)(f) = f(av + bw) = af(v) + bf(w) = a\psi(v)(f) + b\psi(w)(f).$$

To show that ψ is one-to-one, we have to show its null space is $\{\vec{0}\}$. Let $v \in V$ be an arbitrary non-zero vector. Let $\beta = \{v, v_2, \ldots, v_n\}$ be a completion of $\{v\}$ to an ordered basis. Let $\beta^* = \{f_1, \ldots, f_n\}$ be its dual basis. Then $f_1(v) = 1 \neq 0$. Thus, $(\psi(v))(f_1) = 1$, and $\psi(v)$ is not the zero functional.

Finally, ψ is an isomorphism due to the dimension formula, since

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

Note that ψ is canonical because it does not depend on the choice of a basis. □

3 Elementary Matrix Operations and Systems of Linear Equations

3.1 Elementary matrix operations and elementary matrices

Definition 3.1. Let A be an $m \times n$ matrix. An *elementary row operation* is any one of the following.

i. interchanging any two rows of A

ii. multiplying any row of A with a non-zero scalar

iii. adding any scalar multiple of a row of A to another row.

Definition 3.2. An $n \times n$ *elementary matrix* E is obtained by performing an elementary operation on I_n. We say E is of type 1,2, or 3 if the elementary operation was of that type according to Definition 3.1.

Theorem 3.3. *Let $A \in Mat_{m\times n}$. Let B be obtained from A by an elementary row operation corresponding to the elementary matrix E (of size $m \times m$). Then*

$$B = EA.$$

Proof. Direct verification in the three cases. □

Theorem 3.4. *Elementary matrices are invertible. The inverse of an elementary matrix is an elementary matrix of the same type.*

Proof. Explicit checking. □

3.2 The rank of a matrix and matrix inverses

Definition 3.5. Let $A \in Mat_{m\times n}(F)$. We define the *rank* of A, denoted by $\operatorname{rank} A$, to be the rank of the linear transformation $L_A : F^n \to F^m$ given by A.

Theorem 3.6. *Let $A \in Mat_{m\times n}(F)$. If $P \in Mat_{m\times m}(F)$ and $Q \in Mat_{n\times n}(F)$ are invertible, then*

i. $\operatorname{rank}(AQ) = \operatorname{rank}(A)$

ii. $\operatorname{rank}(PA) = \operatorname{rank}(A)$

iii. $\operatorname{rank}(PAQ) = \operatorname{rank}(A)$

Proof.

$$\operatorname{range}(L_{AQ}) = \operatorname{range}(L_A \circ L_Q) = (L_A \circ L_Q)(F^n) = L_A(F^n)$$

where the last equality holds because L_Q is onto (it is actually an isomorphism). Thus,

$$\operatorname{range}(L_{AQ}) = \operatorname{range}(L_A).$$

In particular, the ranks agree.

The two remaining statements are left as exercises. □

Corollary 3.7. *Elementary row operations do not change the rank of a matrix.*

Theorem 3.8. *The rank of any matrix A equals the maximum number of linearly independent* **columns***, i.e., the dimension of the subspace generated by the columns.*

Proof. Let β be the standard ordered basis for F^n. Then

$$\mathrm{range}(L_A) = \mathrm{span}(L_A(\beta)) = \mathrm{span}\{L_A(e_1), \ldots, L_A(e_n)\}.$$

Note that $L_A(e_i)$ simply is the i-th column of A. □

The following sums up the discussion of Theorem 3.6, Corollary 1 and Corollary 2 in the textbook.

By elementary row and column operations, any $m \times n$ matrix can be transformed to an $m \times n$ matrix of the form

$$D = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix},$$

where I_r is the identity matrix of size $r \times r$. More precisely, there exist elementary matrices $E_1, \ldots, E_p$ and $G_1, \ldots, G_q$ such that

$$D = E_p \ldots E_1 A G_1 \ldots G_q.$$

Then

$$\operatorname{rank} A = \operatorname{rank} D = \operatorname{rank} D^t = \operatorname{rank} A^t,$$

where the last inequality is due to $D^t = G_q^t \ldots G_1^t A E_1^t \ldots E_p^t$ and the fact that the transpose of an elementary matrix is an elementary matrix.

We have just proven that

$$\operatorname{rank} A = \operatorname{rank} A^t.$$

Definition 3.9. Let the *row rank* of a matrix denote the maximum number of linearly independent rows.

Since the row rank of a matrix is clearly equal to the rank of its transpose, we have proven that the row rank and the rank of any matrix agree.

Quiz 5:

(10 points) Let f_1, f_2, f_3 be linear functionals on $\mathbb{R}^3$ defined by

$$f_1(x, y, z) = x - y, \quad f_2(x, y, z) = x + y + z, \quad f_3(x, y, z) = y - z.$$

Find an ordered basis for $\mathbb{R}^3$ whose dual is the ordered basis $\{f_1, f_2, f_3\}$.

Solution:

We have to find the three vectors in the basis $\{v_1, v_2, v_3\}$.

Assume $v_1 = (a, b, c)$. Then v_1 is determined by

$$f_1(v_1) = 1,\ f_2(v_1) = 0,\ f_3(v_1) = 0.$$

This is equivalent to

$$a - b = 1, a + b + c = 0, b - c = 0,$$

which is equivalent to $a = \frac{2}{3}, b = -\frac{1}{3}, c = -\frac{1}{3}$, i.e., $v_1 = (\frac{2}{3}, -\frac{1}{3}, -\frac{1}{3})$.

Now we assume $v_2 = (a, b, c)$, and v_2 is determined by

$$f_1(v_2) = 0,\ f_2(v_2) = 1,\ f_3(v_2) = 0.$$

This is equivalent to

$$a - b = 0, a + b + c = 1, b - c = 0,$$

which is equivalent to $a = \frac{1}{3}, b = \frac{1}{3}, c = \frac{1}{3}$, i.e., $v_2 = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$.

Finally, we assume $v_3 = (a, b, c)$, and v_3 is determined by

$$f_1(v_3) = 0,\ f_2(v_3) = 0,\ f_3(v_3) = 1.$$

This is equivalent to

$$a - b = 0, a + b + c = 0, b - c = 1,$$

which is equivalent to $a = \frac{1}{3}, b = \frac{1}{3}, c = -\frac{2}{3}$, i.e., $v_3 = (\frac{1}{3}, \frac{1}{3}, -\frac{2}{3})$.

We continue with the discussion of the rank of a matrix. What we have discussed yields yields the following corollary.

Corollary 3.10. *The rank of a matrix A can be found by reducing A to echelon form via row operations and counting the number of non-zero rows.*

Here is an example. To find the rank of $A =$

$$\begin{pmatrix} 0 & 2 & 4 & 2 & 2 \\ 4 & 4 & 4 & 8 & 0 \\ 8 & 2 & 0 & 10 & 2 \\ 6 & 3 & 2 & 9 & 1 \end{pmatrix},$$

we simply reduce it, via elementary row operations to echelon form:

$$\begin{pmatrix} 1 & 1 & 1 & 2 & 0 \\ 0 & 1 & 2 & 1 & 1 \\ 4 & 1 & 0 & 5 & 1 \\ 6 & 3 & 2 & 9 & 1 \end{pmatrix},$$

$$\begin{pmatrix} 1 & 1 & 1 & 2 & 0 \\ 0 & 1 & 2 & 1 & 1 \\ 0 & -3 & -4 & -3 & 1 \\ 0 & -3 & -4 & -3 & 1 \end{pmatrix},$$

$$\begin{pmatrix} 1 & 1 & 1 & 2 & 0 \\ 0 & 1 & 2 & 1 & 1 \\ 0 & 0 & 2 & 0 & 4 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

The number of non-zero rows is 3, which means that the rank of A is 3.

Finally, a remark on the method of computing the inverse of a matrix.

Let A be an $n \times n$ invertible matrix. If we reduce A by elementary row operations to the identity, then we are effectively executing the following multiplication of A with elementary matrices:

$$E_p \dots E_1 A = I_n.$$

Observe that obviously $E_p \dots E_1 = A^{-1}$. Writing trivially $A^{-1} = E_p \dots E_1 = E_p \dots E_1 I_n$, we conclude that the inverse of A can be obtained by applying the same sequence of elementary row operations to the identity. This method is commonly taught in an introductory course to Linear Algebra, but we now have a very nice justification for it.

3.3 Systems of linear equations – theoretical aspects

A general system of m linear equations in n variables is

$$a_{11}x_1 + \dots + a_{1n}x_n = b_1, \dots, a_{m1}x_1 + \dots + a_{mn}x_n = b_m.$$

This can be rewritten as

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$$

or

$$Ax = b.$$

Definition 3.11. A solution of the above system is $s = (s_1, \dots, s_n) \in F^n$ such that $As = b$. The system is called *consistent* if there exists a solution. The system is called *homogeneous* if and only if $b = \vec{0}$.

Theorem 3.12. *The set of solutions of $Ax = \vec{0}$ is the null space of L_A.*

Proof. Clear! □

Corollary 3.13. *If $m < n$, the system $Ax = \vec{0}$ has a non-zero solution.*

Proof. The dimension formula yields

$$\text{nullity}(L_A) = n - \text{rank}(L_A) \geq n - m > 0.$$

□

Theorem 3.14. *Let K be the set of solutions of $Ax = b$ and K_H be the set of solutions of $Ax = 0$. Then for any s with $As = b$,*

$$K = \{s\} + K_H = \{s + k : k \in K_H\}.$$

Proof. "$\supseteq$". Let $t \in K_H$. Then

$$A(s+t) = As + At = As + \vec{0} = As = b.$$

"$\subseteq$". Let $t \in K$. Let $w = t - s$. Then

$$Aw = A(t - s) = At - As = b - b = 0.$$

Thus, $w \in K_H$ and $t = s + w \in \{s\} + K_H$. □

Example 3.15. Let us consider the single inhomogeneous equation in three variables $x_1 - 2x_2 + x_3 = 4$. An obvious solution is $s = (4, 0, 0)$. Moreover, the solution set of the homogeneous equation $x_1 - 2x_2 + x_3 = 0$ is

$$K_H = \{(2x_2 - x_3, x_2, x_3) : x_2, x_3 \in \mathbb{R}\},$$

which can be rewritten as

$$\{x_2(2, 1, 0) + x_3(-1, 0, 1) : x_2, x_3 \in \mathbb{R}\}.$$

Theorem 3.14 now says that

$$K = \{(4, 0, 0) + x_2(2, 1, 0) + x_3(-1, 0, 1) : x_2, x_3 \in \mathbb{R}\}$$

(which we could have found directly, of course).

Theorem 3.16. *Let $Ax = b$ be a system of n equations in n variables. Then A is invertible if and only if the system has exactly one solution.*

Proof. $\Rightarrow$. Let us multiply the equation from the left with A^{-1}. Then it becomes

$$A^{-1}Ax = x = A^{-1}b.$$

Thus, $x = A^{-1}b$ is the unique solution.

$\Leftarrow$. We saw: $K = \{s\} + K_H$, where S is a solution. Since the solution is unique, we can infer $K_H = \{\vec{0}\}$. Thus, the null space of L_A is $\{\vec{0}\}$, and A is invertible. $\square$

Example 3.17. Let

$$A = \begin{pmatrix} 0 & -1 \\ 2 & 4 \end{pmatrix}.$$

Then

$$A \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \end{pmatrix}$$

is equivalent to

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = A^{-1} \begin{pmatrix} 2 \\ 2 \end{pmatrix} = \begin{pmatrix} 2 & \frac{1}{2} \\ -1 & 0 \end{pmatrix} \begin{pmatrix} 2 \\ 2 \end{pmatrix} = \begin{pmatrix} 5 \\ -2 \end{pmatrix}.$$

Theorem 3.18. *The system $Ax = b$ is consistent if and only if* $\operatorname{rank} A = rank(A|b)$*, where $(A|b)$ is the augmented matrix*

$$\begin{pmatrix} a_{11} & \dots & a_{1n} & b_1 \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & \dots & a_{mn} & b_m \end{pmatrix}.$$

Proof. The system is consistent if and only if $b \in R(L_A)$, which means $b \in \operatorname{span}\{L_A(e_1), \dots, L_A(e_n)\}$. This in turn is equivalent to

$$\operatorname{span}\{L_A(e_1), \dots, L_A(e_n)\} = \operatorname{span}\{L_A(e_1), \dots, L_A(e_n), b\},$$

which equivalent to $\operatorname{rank} A = rank(A|b)$. □

Example 3.19. Let's consider the system $x_1 + 2x_2 = 1,\ 2x_1 + 4x_2 = 0$. It is not consistent because

$$\operatorname{rank} \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} = 1,$$

but

$$\operatorname{rank} \begin{pmatrix} 1 & 2 & 1 \\ 2 & 4 & 0 \end{pmatrix} = 2.$$

4 Determinants

4.1 Determinants of 2×2 matrices

Definition 4.1. Let $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ be a 2×2 matrix over some field F. Then we define the *determinant* of A (also $\det A$ or $|A|$) to be the scalar $ad - bc$.

Example 4.2. $\det \begin{pmatrix} 5 & 0 \\ 1 & 2 \end{pmatrix} = 5 \cdot 2 - 1 \cdot 0 = 10,\ \det \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} = 1 \cdot 4 - 2 \cdot 2 = 0.$

The following is an important observation:

$$\det \left(\begin{pmatrix} 5 & 0 \\ 1 & 2 \end{pmatrix} + \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \right) = \det \begin{pmatrix} 6 & 2 \\ 4 & 6 \end{pmatrix} = 36 - 8 = 28.$$

On the other hand,

$$\det\begin{pmatrix}5 & 0\\ 1 & 2\end{pmatrix} + \det\begin{pmatrix}1 & 2\\ 3 & 4\end{pmatrix} = 10 - 2 = 8.$$

The above clearly shows that the determinant is not linear, which we state as the following remark.

Remark 4.3. $\det : M_{n\times n}(F) \to F$ is not a linear functional.

Instead, the right kind of linearity is the following.

Theorem 4.4. $\det : M_{n\times n}(F) \to F$ *is a linear function of each row when the other row is held fixed, i.e., for the first row, we have*

$$\det\begin{pmatrix}a_{11}+b_{11} & a_{12}+b_{12}\\ a_{21} & a_{22}\end{pmatrix} = \det\begin{pmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{pmatrix} + \det\begin{pmatrix}b_{11} & b_{12}\\ a_{21} & a_{22}\end{pmatrix}$$

and

$$\det\begin{pmatrix}ca_{11} & ca_{12}\\ a_{21} & a_{22}\end{pmatrix}.$$

The statements for the second row are analogous.

Proof. Expand all expressions according to the definition and compare both sides of the equations. □

Theorem 4.5. *Let* $A \in M_{2\times 2}(F)$. *Then*

$$\det A \neq 0 \Leftrightarrow A \text{ is invertible.}$$

Proof. "⇒" Since $\det A \neq 0$, the matrix $B = \frac{1}{\det A}\begin{pmatrix}a_{22} & -a_{12}\\ -a_{21} & a_{11}\end{pmatrix}$ is well-defined. Since $AB = I_2 = BA$, it is clear that A is invertible and its inverse is B.

"⇐" Since A is invertible, its rank is two. Thus, at least one of a_{11}, a_{21} must be not equal to zero. Without loss of generality, let's assume it is a_{11}. If we perform on A the elementary row operation of adding $\frac{-a_{21}}{a_{11}}$ times the first row to the second, we obtain the matrix

$$\begin{pmatrix}a_{11} & a_{12}\\ 0 & a_{22} - \frac{a_{12}a_{21}}{a_{11}}\end{pmatrix}.$$

We know that elementary row operations do not change the rank, so this matrix still has rank two, which implies

$$a_{22} - \frac{a_{12}a_{21}}{a_{11}} \neq 0.$$

Multiplying this relation by a_{11} yields

$$\det A = a_{11}a_{22} - a_{12}a_{21} \neq 0.$$

□

Quiz 6:

1. (8 points) Determine the rank of the following matrix:

$$A = \begin{pmatrix} 1 & -1 & 1 & 2 \\ 2 & 0 & 3 & 3 \\ -1 & 2 & 3 & 2 \\ 1 & 1 & 2 & 1 \\ 0 & 3 & 5 & 3 \end{pmatrix}$$

2. (2 points) Let $A, B \in M_{n\times n}$. Let $\operatorname{rank} A = s$, $\operatorname{rank} B = t$. Make the best possible statement about $\operatorname{rank} A + B$ based on the information provided. You do not have to prove your answer.

Solution:

1. One uses elementary row operations to transform the matrix into echelon form. This does not change the rank. The matrix in echelon form has three nonzero rows, and thus the rank is three. (We omit the explicit computation.)

2. $|t - s| \leq \operatorname{rank}(A + B) \leq \min\{n, s + t\}$.

We conclude this subsection with a geometric argument concerning the area of the parallelogram spanned by two vectors in $\mathbb{R}^2$.

Theorem 4.6. *The area of the parallelogram spanned by* (a_1, a_2) *and* (b_1, b_2) *is*

$$|\det \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix}|.$$

Proof. We rotate the parallelogram so that the first vector point in the direction of the positive x-axis. After the rotation, the new parallelogram is spanned by vectors of the form $(\tilde{a}_1, 0)$ and $(\tilde{b}_1, \tilde{b}_2)$. We state two facts.

- The rotation does not change the area of the parallelogram.
- $|\det \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix}| = |\det \begin{pmatrix} \tilde{a}_1 & \tilde{a}_2 \\ 0 & \tilde{b}_2 \end{pmatrix}|$.

Recall that the area of a parallelogram is given by the formula base · height. After the rotation, the base is $\tilde{a}_1$ and the height is $|\tilde{b}_2|$. Thus, the area of the parallelogram is

$$\tilde{a}_1 \cdot |\tilde{b}_2| = |\det \begin{pmatrix} \tilde{a}_1 & \tilde{a}_2 \\ 0 & \tilde{b}_2 \end{pmatrix}| = |\det \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix}|.$$

□

4.2 Determinants of order n

We now define the determinant of a matrix $A \in M_{n\times n}$ for $n \geq 3$.

Remark 4.7. In class and in these notes so far, our convention was to refer to the entries of matrices with lower case letters, e.g., as a_{ij}. The textbook however uses upper case letters, e.g., A_{ij}. We follow the convention of the textbook in this section to make going back and forth between these notes and the textbook easier. Be sure not to confuse A_{ij} with $\tilde{A}_{ij}$ as defined below.

Definition 4.8. Let $A \in M_{n\times n}$. For fixed $i, j \in \{1, \dots, n\}$ denote by $\tilde{A}_{ij}$ the $(n-1)\times(n-1)$ matrix obtained by deleting the i-th row and j-th column of A.

Example 4.9. Let

$$A = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 3 & \pi \\ \sqrt{2} & 1 & -1 \end{pmatrix}.$$

Then

$$\tilde{A}_{11} = \begin{pmatrix} 3 & \pi \\ 1 & -1 \end{pmatrix}, \ \tilde{A}_{12} = \begin{pmatrix} 1 & \pi \\ \sqrt{2} & -1 \end{pmatrix}, \ \tilde{A}_{32} = \begin{pmatrix} 2 & 0 \\ 1 & \pi \end{pmatrix}.$$

Definition 4.10. Let $A \in M_{n\times n}$. If $n = 1$, then let $\det A = A_{11}$.

If $n \geq 2$, define $\det A$ recursively as

$$\det A = \sum_{j=1}^{n} (-1)^{1+j} A_{1j} \det(\tilde{A}_{1j})$$

Definition 4.11. The determinant $\det \tilde{A}_{ij}$ is called the (i,j)-*minor*. The scalar $c_{ij} = (-1)^{i+j} \det(\tilde{A}_{ij})$ is called the (i,j)-*cofactor*.

With the above definitions the definition of the determinant can be restated as

$$\det A = \sum_{j=1}^{n} c_{1j} A_{1j}.$$

For obvious reasons, this is also referred to as the cofactor expansion of the determinant along the first row. Here are some examples:

Let $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Then $\det A = (-1)^2 a \det(d) + (-1)^3 b \det(c) = ad - bc$, as defined in the previous subsection.

Let $A = \begin{pmatrix} 1 & 3 & -3 \\ -3 & -5 & 2 \\ -4 & 4 & -6 \end{pmatrix}$. Then $\det A =$

$$(-1)^2 \cdot 1 \cdot \det \begin{pmatrix} -5 & 2 \\ 4 & -6 \end{pmatrix} + (-1)^3 \cdot 3 \cdot \det \begin{pmatrix} -3 & 2 \\ -4 & -6 \end{pmatrix} + (-1)^4 \cdot (-3) \cdot \det \begin{pmatrix} -3 & -5 \\ -4 & 4 \end{pmatrix} = 40.$$

Let $A = \begin{pmatrix} 0 & 1 & 3 \\ -2 & -3 & -5 \\ 4 & -4 & 4 \end{pmatrix}$. Then $\det A =$

$$0 + (-1)^3 \cdot 1 \cdot \det \begin{pmatrix} -2 & -5 \\ 4 & 4 \end{pmatrix} + (-1)^4 \cdot 3 \cdot \det \begin{pmatrix} -2 & -3 \\ 4 & -4 \end{pmatrix} = 48.$$

Let $A = \begin{pmatrix} 2 & 0 & 0 & 1 \\ 0 & 1 & 3 & -3 \\ -2 & -3 & -5 & 2 \\ 4 & -4 & 4 & -6 \end{pmatrix}$. Then $\det A =$

$$(-1)^2 \cdot 2 \cdot \det \begin{pmatrix} 1 & 3 & -3 \\ -3 & -5 & 2 \\ -4 & 4 & -6 \end{pmatrix} + (-1)^5 \cdot 1 \cdot \det \begin{pmatrix} 0 & 1 & 3 \\ -2 & -3 & -5 \\ 4 & -4 & 4 \end{pmatrix} = 32.$$

Theorem 4.12. *The determinant of an $n \times n$ matrix is a linear function of each row when the remaining rows are held fixed.*

Proof. Write the matrix in questions as

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{r1} + kb_{r1} & \dots & a_{rn} + kb_{rn} \\ \vdots & \vdots & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix}.$$

Also, let

$$B = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{r1} & \dots & a_{rn} \\ \vdots & \vdots & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix}$$

and

$$C = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \vdots & \vdots \\ b_{r1} & \dots & b_{rn} \\ \vdots & \vdots & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix}.$$

The theorem is immediate from the definition of the determinant if the row in question is the first row, i.e., $r = 1$. So we may assume that the row in question is not the first row, i.e., $r \geq 2$. By induction, we know that

$$\det \tilde{A}_{1j} = \det \tilde{B}_{1j} + k \det \tilde{C}_{1j},$$

because $\tilde{A}_{1j}, \tilde{B}_{1j}, \tilde{C}_{1j}$ are matrices of size $(n-1) \times (n-1)$ and have the exact form necessary to apply the theorem. Thus,

$$\det A = \sum_{j=1}^{n} (-1)^{1+j} A_{1j} \det(\tilde{A}_{1j}) = \sum_{j=1}^{n} (-1)^{1+j} A_{1j} (\det(\tilde{B}_{1j}) + k \det(\tilde{C}_{1j})).$$

Since clearly

$$A_{1j} = B_{1j} = C_{1j},$$

the above equals

$$\sum_{j=1}^{n} (-1)^{1+j} B_{1j} \det(\tilde{B}_{1j}) + \sum_{j=1}^{n} (-1)^{1+j} C_{1j} k \det(\tilde{C}_{1j}) = \det B + k \det C.$$

□

Corollary 4.13. *If A has a row of zeroes, then its determinant is zero.*

Proof. Let it be the r-th row of A which is zero. Let B be a matrix which agrees with A outside of the r-th row. Then by the preceding theorem,

$$\det B = \det B + \det A.$$

Thus, $\det A = 0$. $\square$

Next, we establish that it is not necessary to use the first row for the expansion to compute the determinant: any row will work.

Theorem 4.14. *The determinant of a square matrix $A \in Mat_{n\times n}$ can be found by cofactor expansion along any row, i.e., for all $i = 1, \ldots, n$,*

$$\det A = \sum_{j=1}^{n} (-1)^{i+j} A_{ij} \det(\tilde{A}_{ij}).$$

Proof. The case $i = 1$ is the definition, so we may assume $i \geq 2$. By linearity with respect to the i-th row, we obtain

$$\det A = a_{i1} \det \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & \vdots & \vdots \\ 1 & 0 & 0 \\ \vdots & \vdots & \vdots \\ a_{n1} & \ldots & a_{nn} \end{pmatrix} + \ldots + a_{in} \det \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & \vdots & \vdots \\ 0 & 0 & 1 \\ \vdots & \vdots & \vdots \\ a_{n1} & \ldots & a_{nn} \end{pmatrix}.$$

The next step in the proof uses the Lemma on pages 213-214 in the textbook. We do not give its proof here, because a detailed rigorous proof (filling the entire page 214) is given in the textbook. In essence, the lemma says that the above is equal to

$$a_{i1}(-1)^{i+1} \det \tilde{A}_{i1} + \ldots + a_{1n}(-1)^{i+n} \det \tilde{A}_{in},$$

which is what we intended to prove. $\square$

Corollary 4.15. *If $A \in Mat_{n\times n}$ has two identical rows, then* $\det A = 0$.

Proof. We proceed by induction. Let $n = 2$. Then clearly

$$\det A = \det \begin{pmatrix} a & b \\ a & b \end{pmatrix} = 0.$$

In the case of general $n \geq 3$, let the two identical rows be rows r and s. Let $i \neq r,\ i \neq s$. Then we compute $\det A$ by an expansion along the i-th row:

$$\det A = \sum_{j=1}^{n} (-1)^{i+j} A_{ij} \det(\tilde{A}_{ij}).$$

However, the matrices $\tilde{A}_{ij}$ also have two identical rows (since we are deleting a third row), so by induction their determinants are 0. Thus, $\det A$ is also zero. □

Next, we investigate the influence of elementary row operations on the determinant.

Theorem 4.16. *Let $A \in Mat_{n\times n}$. If B is obtained from A by interchanging two rows, then* $\det B = -\det A$.

Proof. Write

$$A = \begin{pmatrix} a_1 \\ \vdots \\ a_r \\ \vdots \\ a_s \\ \vdots \\ a_n \end{pmatrix}, \quad B = \begin{pmatrix} a_1 \\ \vdots \\ a_s \\ \vdots \\ a_r \\ \vdots \\ a_n \end{pmatrix},$$

with the a_i representing rows. Next, observe that

$$0 = \det \begin{pmatrix} a_1 \\ \vdots \\ a_r + a_s \\ \vdots \\ a_r + a_s \\ \vdots \\ a_n \end{pmatrix} = \det \begin{pmatrix} a_1 \\ \vdots \\ a_r \\ \vdots \\ a_r + a_s \\ \vdots \\ a_n \end{pmatrix} + \det \begin{pmatrix} a_1 \\ \vdots \\ a_s \\ \vdots \\ a_r + a_s \\ \vdots \\ a_n \end{pmatrix}$$

due to Corollary 4.15 and linearity. We can expand this further to

$$\det \begin{pmatrix} a_1 \\ \vdots \\ a_r \\ \vdots \\ a_r \\ \vdots \\ a_n \end{pmatrix} + \det \begin{pmatrix} a_1 \\ \vdots \\ a_r \\ \vdots \\ a_s \\ \vdots \\ a_n \end{pmatrix} + \det \begin{pmatrix} a_1 \\ \vdots \\ a_s \\ \vdots \\ a_r \\ \vdots \\ a_n \end{pmatrix} + \det \begin{pmatrix} a_1 \\ \vdots \\ a_s \\ \vdots \\ a_s \\ \vdots \\ a_n \end{pmatrix}.$$

Due to Corollary 4.15, the two outer determinants are zero, which proves our claim. □

Theorem 4.17. *Let $A \in Mat_{n\times n}$ and let B be obtained from A by adding a multiple of one row to another row of A. Then $\det B = \det A$.*

Proof.

$$\det B = \det \begin{pmatrix} a_1 \\ \vdots \\ a_r \\ \vdots \\ ka_r + a_s \\ \vdots \\ a_n \end{pmatrix} = k \det \begin{pmatrix} a_1 \\ \vdots \\ a_r \\ \vdots \\ a_r \\ \vdots \\ a_n \end{pmatrix} + \det \begin{pmatrix} a_1 \\ \vdots \\ a_r \\ \vdots \\ a_s \\ \vdots \\ a_n \end{pmatrix}.$$

Due to Corollary 4.15, the determinant of the matrix with the two identical rows is zero, which proves our claim. □

Corollary 4.18. *Let $A \in Mat_{n\times n}$. Then A has rank less than n if and only if $\det A = 0$.*

Proof. "⇒". We have seen that performing an elementary row operation on a matrix may change the value of its determinant. However, it will not change whether the determinant is zero or not. Now, if we transform A into echelon form, the fact that A has rank less than n means that the echelon form will contain a row of zeroes. Thus, the determinant of the matrix in echelon form is equal to zero due to Corollary 4.13. This means that the determinant of A was zero to begin with.

"$\Leftarrow$". If $\det A = 0$, then a matrix in echelon form obtained from A via elementary row operations also has determinant zero, because this is unaffected by elementary row operations. However, it is a consequence of homework problem 4.2, # 23 and the definition of *echelon form* that a square matrix in echelon form which has determinant zero actually has a row of zeroes at the bottom. Thus, the rank of A is less than n. □

Remark 4.19. In the above Corollary, the textbook chooses to state only the "$\Rightarrow$" direction, because the "$\Leftarrow$" direction strictly speaking requires the use of the homework problem. Moreover, a very simple argument for the "$\Leftarrow$" direction will be available in the next subsection.

Quiz 7:

1. (5 points) Over the real numbers, find the value of k that satisfies the following equation:

$$\det \begin{pmatrix} 2a_1 & 2a_2 & 2a_3 \\ 3b_1 + 5c_1 & 3b_2 + 5c_2 & 3b_3 + 5c_3 \\ 4c_1 & 4c_2 & 4c_3 \end{pmatrix} = k \det \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix}.$$

Justify your answer carefully.

2. (5 points) Let $A \in \mathrm{Mat}_{n\times n}(\mathbb{R})$. What is the condition for $\det(A) = \det(-A)$? Justify your answer carefully. Hint: Be sure to consider all possible cases.

Solution:

1.

$$\det \begin{pmatrix} 2a_1 & 2a_2 & 2a_3 \\ 3b_1 + 5c_1 & 3b_2 + 5c_2 & 3b_3 + 5c_3 \\ 4c_1 & 4c_2 & 4c_3 \end{pmatrix} = \det \begin{pmatrix} 2a_1 & 2a_2 & 2a_3 \\ 3b_1 & 3b_2 & 3b_3 \\ 4c_1 & 4c_2 & 4c_3 \end{pmatrix}$$

because adding $-\frac{5}{4}$ times the third row to the second one does not change the determinant. Moreover,

$$\det \begin{pmatrix} 2a_1 & 2a_2 & 2a_3 \\ 3b_1 & 3b_2 & 3b_3 \\ 4c_1 & 4c_2 & 4c_3 \end{pmatrix} = 2 \cdot 3 \cdot 4 \cdot \det \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix}.$$

Thus, $k = 24$.

2. If n is even, then

$$\det(-A) = \det((-1)A) = (-1)^n \det A = \det(A),$$

i.e., there is no condition.

If n is odd, then

$$\det(-A) = \det((-1)A) = (-1)^n \det A = -\det(A),$$

which is true if and only if $\det A = 0$, i.e., A has rank less than n.

4.3 Properties of determinants

It is clear from our previous discussion that

$$\det(\text{elementary matrix which interchanges two rows}) = -1,$$

$$\det(\text{elementary matrix which adds a multiple of one row to another}) = 1,$$

$$\det(\text{elementary matrix which multiplies a given row by } k) = k.$$

While we saw earlier that matrix addition is not compatible with taking the determinant, we now show that matrix multiplication is.

Theorem 4.20. *Let* $A, B \in Mat_{n\times n}$. *Then*

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

Proof. We saw earlier that $\det(EC) = \det(E)\det(C)$ for any $C \in Mat_{n\times n}$ when E is an elementary matrix. We may assume that $\operatorname{rank} A = \operatorname{rank} B = n$, because otherwise both sides are clearly equal to zero. As we did earlier, we write A as a product of elementary matrices $A = E_1 \dots E_p$. Now

$$\begin{aligned}
\det(AB) &= \det(E_1 \dots E_p B) \\
&= \det(E_1)\det(E_2 \dots E_p B) \\
&= \det(E_1)\dots\det(E_p)\det(B) \\
&= \det(E_1)\dots\det(E_{p-1}E_p)\det(B) \\
&= \det(E_1 \dots E_p)\det(B) \\
&= \det A \cdot \det B
\end{aligned}$$

□

Corollary 4.21. *Let $A \in Mat_{n\times n}$. Then A is invertible if and only if* $\det A \neq 0$*. Moreover, if* $\det A \neq 0$*, then* $\det A^{-1} = \frac{1}{\det A}$.

Proof. We already gave an argument for "$\Leftarrow$" in the previous section. We now prove "$\Rightarrow$", for which we just observe that

$$1 = \det I_n = \det(AA^{-1}) = \det A \cdot \det A^{-1},$$

which implies that $\det A$ cannot be zero. Moreover, $\det A^{-1} = \frac{1}{\det A}$ is also clear now. □

Theorem 4.22. *Let $A \in Mat_{n\times n}$. Then* $\det A = \det A^T$.

Proof.

$$\begin{aligned} \det(A^T) &= \det((E_1 \dots E_p)^T) \\ &= \det(E_p^T \dots E_1^T) \\ &= \det(E_p^T) \dots \det(E_1^T) \\ &= \det(E_1^T) \dots \det(E_p^T) \\ &= \det(E_1) \dots \det(E_p) \\ &= \det(E_1 \dots E_p) \\ &= \det(A). \end{aligned}$$

Note that the fourth equal sign is simply due to the commutativity of the reals. The equality $\det E_i^T = \det E_i$ is obvious for all three types of elementary matrices. □

Theorem 4.23 (Cramer's Rule)**.** *Let $A \in Mat_{n\times n}$ be an invertible matrix. Let $x = (x_1, \dots, x_n)$ be the unique solution of $Ax = b$. Let M_k be obtained from A by replacing the k-th column of A by b. Then*

$$x_k = \frac{\det M_k}{\det A}.$$

Proof. Let X_k be obtained from I_n by replacing the k-th column by the vector x. The expansion of X_k along the k-th row shows that $\det X_k = x_k$. Moreover, it is straightforward to check that

$$A \cdot X_k = M_k.$$

Thus,

$$\det M_k = \det(A \cdot X_k) = \det A \cdot \det X_k = (\det A) \cdot x_k,$$

which yields

$$x_k = \frac{\det M_k}{\det A}.$$

□

Let us apply Cramer's rule to the example of

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 0 & 1 \\ 1 & 1 & -1 \end{pmatrix}, \quad b = \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix}.$$

In this case,

$$M_1 = \begin{pmatrix} 2 & 2 & 3 \\ 3 & 0 & 1 \\ 1 & 1 & -1 \end{pmatrix}, \quad M_2 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 1 \\ 1 & 1 & -1 \end{pmatrix}, \quad M_3 = \begin{pmatrix} 1 & 2 & 2 \\ 1 & 0 & 3 \\ 1 & 1 & 1 \end{pmatrix}.$$

One easily computes

$$\det A = 6, \quad \det M_1 = 15, \quad \det M_2 = -6, \quad \det M_3 = 3.$$

Thus,

$$x_1 = \frac{15}{6} = \frac{5}{2}, \quad x_2 = \frac{-6}{6} = -1, \quad x_3 = \frac{3}{6} = \frac{1}{2}.$$

Remark 4.24. We end this section by remarking that the comments on determinants and area in the previous section can be extended to arbitrary dimension. Namely, n linearly independent vectors in $\mathbb{R}^n$ span a parallelepiped whose volume is the absolute value of the determinant of the $n \times n$ matrix consisting of the n vectors as columns (or rows).

5 Diagonalization

5.1 Eigenvalues and eigenvectors

Motivation: Let $T : V \to V$ be a linear transformation. It may happen that a certain nonzero vector v gets mapped to a scalar multiple of itself, i.e., $T(v) = \lambda v$ for some $\lambda \in F$. Such vectors are clearly special, and we now study these types of pairs of scalars and vectors (v, λ) and related questions.

Definition 5.1. A linear operator T on a finite dimensional vector space V is called diagonalizable if there is an ordered basis β for V such that $[T]_\beta$ is a diagonal matrix. Moreover, we call a square matrix diagonalizable if $L_A : F^n \to F^n, v \mapsto Av$ is diagonalizable.

Definition 5.2. Let T be a linear operator on the vector space V. A nonzero $v \in V$ is called *eigenvector* of T if

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

for some scalar λ in F. The scalar λ is called the corresponding *eigenvalue*. Moreover, for $A \in Mat_{n\times n}(F)$, the nonzero vector $v \in F^n$ is an *eigenvector of* A if and only if $L_A(v) = \lambda v$, i.e., $Av = \lambda v$, for some scalar $\lambda \in F$, which we again call the corresponding *eigenvalue*.

Theorem 5.3. *Let T be a linear operator on V. Then T is diagonalizable if and only if there exists an orered basis $\beta = \{v_1, \ldots, v_n\}$ of eigenvectors.*

Proof. "$\Rightarrow$" It is clear that

$$[T]_\beta = \begin{pmatrix} \lambda_1 & 0 & 0\ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \ldots & 0 & \lambda_n \end{pmatrix},$$

which is a diagonal matrix, so T is diagonalizable.

"$\Leftarrow$" Let $\beta = \{v_1, \ldots, v_n\}$ be the ordered basis for which $[T]_\beta$ is diagonal. Then

$$[T]_\beta[v_i]_\beta = \lambda_i e_i = \lambda_i[v_i]_\beta.$$

Thus, $T(v_i) = \lambda_i v_i$, which implies that v_i is an eigenvector. □

Example 5.4. Let

$$A = \begin{pmatrix} 1 & 3 \\ 4 & 2 \end{pmatrix}, \quad v_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 3 \\ 4 \end{pmatrix}.$$

Then $Av_1 = (-2, 2) = (-2)(1, -1)$ and $Av_2 = (15, 20) = 5(3, 4)$. Thus, $\beta = \{(1, -1), (3, 4)\}$ and

$$[L_A]_\beta = \begin{pmatrix} -2 & 0 \\ 0 & 5 \end{pmatrix}.$$

Remark 5.5. Linear operators and matrices do not necessarily have eigenvectors. For example, the rotation by 90 degrees in the plane clearly has none.

Example 5.6. Let $T : C^\infty(\mathbb{R}, \mathbb{R}) \to C^\infty(\mathbb{R}, \mathbb{R})$ be given by $f \mapsto \frac{df}{dx}$. Then $T(f) = \lambda f$ is equivalent to $\frac{df}{dx} = \lambda f$, which is solved by $f(x) = ce^{\lambda x}$.

Theorem 5.7. *Let $A \in Mat_{n\times n}(F)$. Then $\lambda \in F$ is an eigenvalue of A if and only if $\det(A - \lambda I_n) = 0$.*

Proof. For a nonzero vector $v \in V$, the following are equivalent:

$$\begin{aligned} & Av = \lambda v \\ \Leftrightarrow\ & Av - \lambda v = \vec{0} \\ \Leftrightarrow\ & Av - \lambda I_n v = \vec{0} \\ \Leftrightarrow\ & (A - \lambda I_n)v = \vec{0} \\ \Leftrightarrow\ & \vec{0} \neq v \in \ker(A - \lambda I_n) \end{aligned}$$

Thus, $\det(A - \lambda I_n) = 0$, and conversely, $\det(A - \lambda I_n) = 0$ implies the existence of a vector v with $Av = \lambda v$. □

Definition 5.8. Let $A \in Mat_{n\times n}(F)$. The polynomial $f(t) = \det(A - tI_n)$ is called the characteristic polynomial of A.

Remark 5.9. Theorem 5.7 says that the set of eigenvalues of A equals the set of roots of the characteristic polynomial.

Let us find the eigenvalues of $A = \begin{pmatrix} 1 & 1 \\ 4 & 1 \end{pmatrix}$.

$$\det(A - tI_2) = \det \begin{pmatrix} 1-t & 1 \\ 4 & 1-t \end{pmatrix} = (1-t)^2 - 4 = t^2 - 2t - 3.$$

The roots of this polynomial are $\lambda = 3$ or $\lambda = -1$.

Let us find the eigenvalues of $A = \begin{pmatrix} 5 & 8 & 16 \\ 4 & 1 & 8 \\ -4 & -4 & -11 \end{pmatrix}$.

$$\det(A - tI_3) = \det \begin{pmatrix} 5-t & 8 & 16 \\ 4 & 1-t & 8 \\ -4 & -4 & -11-t \end{pmatrix} = \det \begin{pmatrix} 5-t & 8 & 16 \\ 4 & 1-t & 8 \\ 0 & -3-t & -3-t \end{pmatrix}.$$

Extracting the common factor from the third row and then doing an expansion along the same row yields.

$$(-3-t)\left(-\det\begin{pmatrix}5-t & 16\\ 4 & 8\end{pmatrix}+\det\begin{pmatrix}5-t & 8\\ 4 & 1-t\end{pmatrix}\right)=-(t+3)(t^2+2t-3).$$

This is a third degree polynomial which can be factored into $-(t+3)^2(t-1)$, which means that the eigenvalues of the matrix A are $\lambda = 1$ and $\lambda = -3$, the latter with multiplicity two.

Finally, let us find the eigenvalues of $A = \begin{pmatrix}0 & -2 & -3\\ -1 & 1 & -1\\ 2 & 2 & 5\end{pmatrix}$.

$$\det(A-tI_3)=\det\begin{pmatrix}-t & -2 & -3\\ -1 & 1-t & -1\\ 2 & 2 & 5-t\end{pmatrix}=\det\begin{pmatrix}-1 & 1-t & -1\\ -t & -2 & -3\\ 2 & 2 & 5-t\end{pmatrix}.$$

In the above, as a first step, we have moved $-t$ out of the top left hand corner. Under obvious row operations, this becomes

$$\det\begin{pmatrix}1 & t-1 & 1\\ 0 & t^2-t-2 & t-3\\ 0 & -2t+4 & 3-t\end{pmatrix}.$$

The determinant can be further simplified to

$$(t-3)\det\begin{pmatrix}t^2-t-2 & 1\\ -2t+4 & -1\end{pmatrix}=-(t-3)(t^2-t+2)=-(t-3)(t-1)(t-2).$$

Thus, the eigenvalues of A are $1, 2, 3$.

Theorem 5.10. *The characteristic polynomial of an $n \times n$ square matrix if a polynomial of degree n with leading coefficient $(-1)^n$. Moreover, a $n \times n$ square matrix has at most n eigenvalues.*

Proof. Clear. □

To find the eigenvectors of a matrix, the essential observation is the following.

Theorem 5.11. *Let $T : V \to V$ be a linear operator. Let λ be an eigenvalue of T. Then a nonzero vector $v \in V$ is an eigenvector of T corresponding to λ if and only if $v \in N(T - \lambda\,\mathrm{id})$.*

Proof.

$$T(v) = \lambda v \quad \Leftrightarrow \quad T(v) - \lambda\,\mathrm{id}(v) = \vec{0} \quad \Leftrightarrow \quad v \in N(T - \lambda\,\mathrm{id}).$$

□

As an example, let us find the eigenvectors of $A = \begin{pmatrix} 0 & -2 & -3 \\ -1 & 1 & -1 \\ 2 & 2 & 5 \end{pmatrix}$, whose eigenvalues we found to be $\lambda = 1, 2, 3$.

To find the eigenvectors corresponding to $\lambda = 1$, we have to find the null space of the matrix

$$A - 1I_3 = \begin{pmatrix} -1 & -2 & -3 \\ -1 & 0 & -1 \\ 2 & 2 & 4 \end{pmatrix},$$

i.e., solve the linear system

$$\begin{pmatrix} -1 & -2 & -3 \\ -1 & 0 & -1 \\ 2 & 2 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

A simple computation shows that $N(A - I_3) = \mathrm{span}\{(-1, -1, 1)\}$.

To find the eigenvectors corresponding to $\lambda = 2$, we have to find the null space of the matrix

$$A - 2I_3 = \begin{pmatrix} -2 & -2 & -3 \\ -1 & -1 & -1 \\ 2 & 2 & 3 \end{pmatrix},$$

i.e., solve the linear system

$$\begin{pmatrix} -2 & -2 & -3 \\ -1 & -1 & -1 \\ 2 & 2 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

A simple computation shows that $N(A - 2I_3) = \mathrm{span}\{(-1, 1, 0)\}$.

To find the eigenvectors corresponding to $\lambda = 3$, we have to find the null space of the matrix

$$A - 3I_3 = \begin{pmatrix} -3 & -2 & -3 \\ -1 & -2 & -1 \\ 2 & 2 & 2 \end{pmatrix},$$

i.e., solve the linear system

$$\begin{pmatrix} -3 & -2 & -3 \\ -1 & -2 & -1 \\ 2 & 2 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

A simple computation shows that $N(A - 3I_3) = \text{span}\{(-1, 0, 1)\}$.

Altogether, $\beta = \{(-1, -1, 1), (-1, 1, 0), (-1, 0, 1)\}$ is an ordered basis of $\mathbb{R}^3$ consisting of eigenvectors of A. Thus, A is diagonalizable, and its diagonal form in this basis is

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}.$$

Quiz 8:

1. Let

$$A = \begin{pmatrix} 1 & 1 \\ -2 & 4 \end{pmatrix}.$$

(a) (3 points) Find the eigenvalues of A.
(b) (3 points) Find the eigenvectors of A.
(c) (2 points) Find a matrix Q such that $Q^{-1}AQ$ is diagonal.

2. (2 points) Give an example of a matrix B such that $(2, 1)$ is an eigenvector with eigenvalue 2 and $(1, 1)$ is an eigenvector with eigenvalue -1.

Solution:

1(a). The characteristic polynomial of A is

$$\det \begin{pmatrix} 1-t & 1 \\ -2 & 4-t \end{pmatrix} = t^2 - 5t + 6 = (t-3)(t-2).$$

Thus, the eigenvalues of A are 2 and 3.

1(b). Let us find the eigenvectors corresponding to $\lambda = 3$ first. To this end, we have to solve the system

$$\begin{pmatrix} 1-3 & 1 \\ -2 & 4-3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

which is equivalent to $x_2 = 2x_1$. Thus, the eigenvectors corresponding to $\lambda = 3$ are of the form $(a, 2a)$, with $a \in \mathbb{R}\backslash\{\vec{0}\}$.

Now, let us find the eigenvectors corresponding to $\lambda = 2$. To this end, we have to solve the system

$$\begin{pmatrix} 1-2 & 1 \\ -2 & 4-2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

which is equivalent to $x_2 = x_1$. Thus, the eigenvectors corresponding to $\lambda = 3$ are of the form (a, a), with $a \in \mathbb{R}\backslash\{\vec{0}\}$.

1(c). The matrix Q can be obtained by inserting the eigenvectors into Q as columns, i.e.,

$$\begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix}.$$

Then

$$Q^{-1} = \begin{pmatrix} -1 & 1 \\ 2 & -1 \end{pmatrix}$$

and a direct computation yields

$$Q^{-1}AQ = \begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix}.$$

2. We take the two eigenvectors and create a matrix Q from them:

$$Q = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}.$$

We know that

$$\begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix} = Q^{-1}BQ.$$

We can solve this equation for B to obtain

$$B = Q \begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix} Q^{-1} = \begin{pmatrix} 5 & -6 \\ 3 & -4 \end{pmatrix}.$$

5.2 Diagonalizability

Theorem 5.12. *Let $T : V \to V$ linear operator, let $\lambda_1, \ldots, \lambda_k$ be distinct eigenvalues of T. If $v_1, \ldots, v_k$ are eigenvectors corresponding to $\lambda_1, \ldots, \lambda_k$ respectively, then $v_1, \ldots, v_k$ are linearly independent.*

Proof. We proceed by induction on k. The case $k = 1$ is trivial. Now, we assume that the statement is true for $k - 1$ and show that it also holds for k. To this end, we write

$$a_1 v_1 + \ldots + a_k v_k = \vec{0}$$

and have to conclude that $a_1 = \ldots = a_k = 0$. We apply the linear operator $T - \lambda_k \operatorname{id}$ to this equation, which yields

$$a_1(\lambda_1 - \lambda_k)v_1 + \ldots + a_{k-1}(\lambda_{k-1} - \lambda_k)v_{k-1} + a_k(0)v_k = \vec{0},$$

i.e.,

$$a_1(\lambda_1 - \lambda_k)v_1 + \ldots + a_{k-1}(\lambda_{k-1} - \lambda_k)v_{k-1} = \vec{0}.$$

Due to the induction assumption, $v_1, \ldots, v_{k-1}$ are linearly independent, so

$$a_i(\lambda_i - \lambda_k) = 0$$

for $i = 1, \ldots, k - 1$. Since all eigenvalues are distinct, $\lambda_i - \lambda_k \neq 0$ when $i \neq k$. Thus, $a_i = 0$ for $i = 1, \ldots, k - 1$. Going back to the first displayed equation in this proof, we finally obtain $a_k = 0$. □

Corollary 5.13. *Let $T : V \to V$ linear. If T has $n = \dim V$ distinct eigenvalues, then T is diagonalizable.*

Proof. Let $v_1, \ldots, v_n$ be eigenvectors corresponding to the distinct eigenvalues, respectively. The previous theorem says that $v_1, \ldots, v_n$ are linearly independent. Since $n = \dim V$, this is actually a basis for V. Thus, we have found a basis of eigenvectors of T for V, which means T is diagonalizable. □

Definition 5.14. A polynomial $f(t)$ over the field F *splits* if there exist constants $c, a_1, \ldots, a_n \in F$:

$$f(t) = c(t - a_1) \cdot \ldots \cdot (t - a_n).$$

Definition 5.15. Let λ be an eigenvalue of $T : V \to V$. The *multiplicity* of λ is the largest positive integer k such that $(t - \lambda)^k$ is a factor in the characteristic polynomial of T.

As an example, let us consider the linear transformation $\mathbb{R}^3 \to \mathbb{R}^3$ given by the upper triangular matrix

$$\begin{pmatrix} 3 & 1 & 0 \\ 0 & 3 & 7 \\ 0 & 0 & 4 \end{pmatrix}.$$

Since this matrix is upper triangular, it is immediately clear that the characteristic polynomial is $-(t-3)^2(t-4)$. Thus, the multiplicity of the eigenvalue 3 is two, while the multiplicity of the eigenvalue 4 is one. We cannot decide the question whether A is diagonalizable based on our discussion so far.

Definition 5.16. Let $T : V \to V$ be a linear operator. Let λ be an eigenvalue of T. We define the *eigenspace* of T corresponding to λ to be

$$E_\lambda := \{x \in V | T(x) = \lambda x\} = N(T - \lambda \operatorname{id}).$$

In other words, E_λ is the set of eigenvectors corresponding to λ together with the zero vector. It is easy to check that E_λ is a subspace of V.

Theorem 5.17. *Let $T : V \to V$ linear. Let λ be an eigenvalue of multiplicity m. Then*

$$1 \leq \dim E_\lambda \leq m.$$

Proof. Choose an ordered basis $\{v_1, \ldots, v_p\}$ for E_λ. We extend it to an ordered basis $\beta = \{v_1, \ldots, v_n\}$ for V. With these definitions, $A = [T]_\beta$ is of the form

$$A = \begin{pmatrix} \lambda I_p & B \\ 0 & C \end{pmatrix}.$$

Its characteristic polynomial is

$$f(t) = \det(A - tI_n) = \det \begin{pmatrix} (\lambda - t)I_p & B \\ 0 & C - tI_{n-p} \end{pmatrix}.$$

By repeated expansion of the determinant along the first row, we see that this equals $(\lambda - t)^p \det(C - tI_{n-p})$. Thus, $(\lambda - t)^p$ is clearly a factor in $f(t)$, which implies $\dim E_\lambda = p \leq m$ by definition of m. □

We remark that it is easy to produce examples where the extreme cases of the inequality in the above theorem are achieved. For example, the matrix

$$A = \begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix}$$

has characteristic polynomial equal to $(t-2)^4$ (i.e., the multiplicity of the eigenvalue $\lambda = 2$ is four), but

$$E_2 = N(A - 2I_4) = \operatorname{span}\{(1,0,0,0)\},$$

and thus $\dim E_2 = 1$.

On the other hand, the diagonal matrix

$$A = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}$$

has characteristic polynomial equal to $(t-2)^4$ (i.e., the multiplicity of the eigenvalue $\lambda = 2$ is again four), and $E_2 = \mathbb{R}^4$, i.e., $\dim E_2 = 4$.

Theorem 5.18. *Let $T : V \to V$ linear operator. Let $\lambda_1, \ldots, \lambda_k$ be distinct eigenvalues. Let S_i be a linearly independent subset of E_{λ_i} for $i = 1, \ldots, k$. Then*

$$S_1 \cup \ldots \cup S_k$$

is linearly independent.

Proof. Write $S_i = \{v_1^{(i)}, \ldots, v_{n_i}^{(i)}\}$ for $i = 1, \ldots, k$. As always, we set

$$a_1^{(1)} v_1^{(1)} + \ldots + a_{n_1}^{(1)} v_{n_1}^{(1)} + \ldots + a_1^{(k)} v_1^{(k)} + \ldots + a_{n_k}^{(k)} v_{n_k}^{(k)} = \vec{0}.$$

We have to show that all the coefficients are equal to zero. By Theorem 5.12, we have

$$a_1^{(i)} v_1^{(i)} + \ldots + a_{n_1}^{(i)} v_{n_1}^{(i)} = \vec{0} \quad \text{for } i = 1, \ldots, k.$$

By the assumed linear independence of all S_i,

$$a_j^{(i)} = 0$$

for $i = 1, \ldots, k$ and $j = 1, \ldots, n_i$. $\square$

The following theorem is now evident.

Theorem 5.19. *Let $T : V \to V$ linear operator on a finite dimensional vector space V. Assume that the characteristic polynomial splits with distinct roots $\lambda_1, \ldots, \lambda_k$. Then* *i. T is diagonalizable* $\Leftrightarrow \operatorname{mult} \lambda_i = \dim E_{\lambda_i}$ *for* $i = 1, \ldots, k$.

ii. If T is diagonalizable with β_i being an ordered basis for $\dim E_{\lambda_i}$*, then $\beta = \beta_1 \cup \ldots \cup \beta_k$ is an ordered basis for V consisting of eigenvectors of T.*

Let us consider another example. Let

$$A = \begin{pmatrix} 3 & 1 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 4 \end{pmatrix}.$$

The characteristic polynomial is $-(t-3)^2(t-4)$. It is now clear that A is diagonalizable if and only if $\dim E_3 = 2$. To see if this holds, we compute $E_3 = N(A - 3I_3)$ by solving

$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

This is clearly equivalent to $x_2 = x_3 = 0$, i.e., $E_3 = \operatorname{span}\{(1, 0, 0)\}$, which is one-dimensional. Thus, A is not diagonalizable.

Remark 5.20. It is clear that a matrix cannot be diagonalizable if the characteristic polynomial is not split. In other words, the characteristic polynomial being split is a necessary (but not sufficient) condition for a matrix to be diagonalizable.

Quiz 9:

1. (7 points) Let

$$A = \begin{pmatrix} 1 & -3 & 3 \\ 3 & -5 & 3 \\ 6 & -6 & 4 \end{pmatrix}.$$

Find bases for the eigenspaces of A. Is A diagonalizable?

2. (3 points) Give an example of a matrix that is not diagonalizable. Prove that your example has this property. If you found the matrix in Problem 1 not to be diagonalizable, you must choose a different example.

Solution:

1. To find the characteristic polynomial, we compute:

$$\begin{aligned}
\det(A - tI_3) &= \det\begin{pmatrix} 1-t & -3 & 3 \\ 3 & -5-t & 3 \\ 6 & -6 & 4-t \end{pmatrix} \\
&= -\det\begin{pmatrix} 3 & -5-t & 3 \\ 1-t & -3 & 3 \\ 6 & -6 & 4-t \end{pmatrix} \\
&= -\frac{1}{3}\det\begin{pmatrix} 3 & -5-t & 3 \\ 3-3t & -9 & 9 \\ 6 & -6 & 4-t \end{pmatrix} \\
&= -\frac{1}{3}\det\begin{pmatrix} 3 & -5-t & 3 \\ 0 & (5+t)(1-t)-9 & 3t+6 \\ 0 & 2t+4 & -2-t \end{pmatrix} \\
&= -(t+2)\frac{1}{3}\det\begin{pmatrix} 3 & -5-t & 3 \\ 0 & (5+t)(1-t)-9 & 3t+6 \\ 0 & 2 & -1 \end{pmatrix} \\
&= -(t+2)\det\begin{pmatrix} (5+t)(1-t)-9 & 3t+6 \\ 2 & -1 \end{pmatrix} \\
&= -(t+2)(t^2-2t-8) = -(t+2)^2(t-4).
\end{aligned}$$

Thus, the eigenvalues are $\lambda = -2$ (with multiplicity two) and $\lambda = 4$ (with multiplicity one). A simple computation finds the following basis for the eigenspace $E_{-2} = N(A + 2I_3)$:

$$\{(1,1,0),(0,1,1)\}.$$

A simple computation finds the following basis for the eigenspace $E_4 = N(A - 4I_3)$:

$$\{(1,1,2)\}.$$

Since $\dim E_{-2} = 2 =$ the multiplicity of the eigenvalue $\lambda = -2$, we can use Theorem 5.19 conclude that A is diagonalizable.

2. The matrix

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

has characteristic polynomial $t^2 + 1$, which has no roots over the reals. Thus, it is not diagonalizable (over the reals).

To conclude, here are the solutions to a couple of homework problems.

Section 5.1, #14. Let $A \in Mat_{n\times n}(\mathbb{R})$. Prove that A and A^T have the same characteristic polynomial. Answer:

$$\det(A^T - tI) = \det(A^T - (tI)^T) = \det((A - tI)^T) = \det(A - tI).$$

Note that we used $(A + B)^T = A^T + B^T$ for all matrices A, B of the same size.

Section 4.2, #4. Let

$$\det \begin{pmatrix} b_1 + c_1 & b_2 + c_2 & b_3 + c_3 \\ a_1 + c_1 & a_2 + c_2 & a_3 + c_3 \\ a_1 + b_1 & a_2 + b_2 & a_3 + b_3 \end{pmatrix} \quad = \quad k \det \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix}.$$

Determine k. Answer: First, use the linearity of the determinant with respect to the first row. Then, use elementary row operations to see that

$$\det \begin{pmatrix} b_1 + c_1 & b_2 + c_2 & b_3 + c_3 \\ a_1 + c_1 & a_2 + c_2 & a_3 + c_3 \\ a_1 + b_1 & a_2 + b_2 & a_3 + b_3 \end{pmatrix} = \det \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix} + \det \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix}.$$

Thus, $k = 2$.

Good luck for the Final Exam!

UH - Math 4377 - Dr. Heier - Fall 2012
Quiz 3 – 10/01/2012

Name:

(10 points) Let $W = \{(a_1, a_2, a_3, a_4, a_5, a_6) \in \mathbb{R}^6 : a_1 + 2a_3 + a_5 + 4a_6 = 0,\ -a_1 + 2a_2 + 2a_4 + a_5 + 6a_6 = 0,\ a_3 = a_4 + a_5 + 2a_6\}$ be a subspace of $\mathbb{R}^6$. Find a basis for W. Determine $\dim W$.

$$\begin{aligned} a_1 + 2a_3 + a_5 + 4a_6 &= 0 \\ -a_1 + 2a_2 + 2a_4 + a_5 + 6a_6 &= 0 \\ a_3 &= a_4 + a_5 + 2a_6 \end{aligned}$$

$$\begin{aligned} (1)\quad a_1 + 2a_3 + a_5 + 4a_6 &= 0 \\ (2)\quad 2a_2 + 2a_3 + 2a_4 + 2a_5 + 10a_6 &= 0 \\ a_3 &= a_4 + a_5 + 2a_6 \end{aligned}$$

(2) becomes $a_2 = -a_3 - a_4 - a_5 - 5a_6 = -(a_4 + a_5 + 2a_6) - a_4 - a_5 - 5a_6 = -2a_4 - 2a_5 - 7a_6$

(1) becomes $a_1 = -2a_3 - a_5 - 4a_6 = -2(a_4 + a_5 + 2a_6) - a_5 - 4a_6 = -2a_4 - 3a_5 - 8a_6$ | $a_4, a_5, a_6 \in \mathbb{R}$

$$\Rightarrow W = \left\{ \begin{pmatrix} -2a_4 - 3a_5 - 8a_6 \\ -2a_4 - 2a_5 - 7a_6 \\ a_4 + a_5 + 2a_6 \\ a_4 \\ a_5 \\ a_6 \end{pmatrix} \middle| a_4, a_5, a_6 \in \mathbb{R} \right\} = \left\{ a_4 \begin{pmatrix} -2 \\ -2 \\ 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + a_5 \begin{pmatrix} -3 \\ -2 \\ 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} + a_6 \begin{pmatrix} -8 \\ -7 \\ 2 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\}$$

$$\Rightarrow B := \left\{ \begin{pmatrix} -2 \\ -2 \\ 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -3 \\ -2 \\ 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -8 \\ -7 \\ 2 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\}$$ is a basis for W.

$\#B = 3 \Rightarrow \dim W = 3$