Mathematics Year I, Linear Algebra Term 1, 2

Most Important theorems, definitions and propositions.

Szymon Kubica

April 17, 2021

1 Introduction to matrices and vectors

Definition 1. The standard basis vectors for $\mathbb{R}^n$ are the vectors

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \quad \dots \quad e_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}, \tag{1}$$

Definition 2. Let $v_1, ...v_n$ be vectors in $\mathbb{R}^n$. Any expression of the form

$$\lambda_1 v_1 + \lambda_2 v_2 + ... + \lambda_n v_n$$

is called linear combination of the vectors $v_1, ...v_n$.

Definition 3. The set of all linear combinations of a collection of vectors $v_1, ...v_n$ is called the span of the vectors $v_1, ...v_n$. Notation:

$$span\{v_1, ...v_n\}$$

.

Note 1. R^n is equal to the span of the standard basis vectors.

Definition 4. The norm of v is the non negative real number defined by

$$||v|| = \sqrt{v \cdot v}$$

.

Definition 5. A vector $v \in \mathbb{R}^n$ is called a unit vector if $||v|| = 1$.

Definition 6. Let u and v be vectors in $\mathbb{R}^n$. The distance between u and v is defined by

$$dist(u, v) := ||u - v||$$

.

Definition 7. The (i, j) entry of a matrix is the entry in row i and column j.

$$\begin{pmatrix} a_{11} & a_{12} & ... & a_{1m} \\ \vdots & & \ddots & \vdots \\ a_{n1} & a_{n2} & ... & anm \end{pmatrix}$$

Most often we use the condensed notation $M = (a_{ij})$.

Definition 8. The transpose of an $n \times m$ matrix $A = (a_{ij})$ is the $m \times n$ matrix whose (i, j) entry is a_{ji}. We denote it as A^T.

The leading diagonal of a matrix is the $(1, 1), (2, 2)...$ entries. So the transpose is obtained by doing a reflection in the leading diagonal.

Definition 9. The identity matrix $I_n = (a_{ij})$ is the square matrix such that $a_{ij} = 0 \; \forall \, i \neq j$, and $a_{ii} = 1$, where $0 < i, j \leq n$.

Definition 10. Let $A = (a_{ij})$ be a $n \times m$ matrix and $\mathbf{b}$ be the column vector of height n and whose ith entry is b_i. Then $(v_1, ..., v_n)$ is a solution to the system

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + ... + a_{1m}x_m & = b_1 \\ a_{21}x_1 + a_{22}x_2 + ... + a_{2m}x_m & = b_2 \\ \vdots & = \vdots \\ a_{n1}x_1 + a_{n2}x_2 + ... + a_{nm}x_m & = b_n \end{cases}$$

if and only if the vector $\mathbf{v} \in \mathbb{R}^n$ with entries v_i is a solution of the equation

$$Av = b$$

The matrix A is called the coefficient matrix of the system above. The augmented matrix associated to the system is the matrix obtained by adding b as an extra column to A. It is denoted as $(A|b)$.

2 Row operations

Definition 11. A **row operation** is one of the following procedures we can apply to a matrix:

1. $r_i(\lambda)$: Multiply each entry in the ith row by a real number $\lambda \neq 0$.
2. r_{ij} : Swap row i and row j.
3. $r_{ij}(\lambda)$: Add λ times row i to row j.

Proposition 1. Let $Ax = \mathbf{b}$ be a system of linear equations in matrix form. Let r be one of the row operations from Definition 11, and let $(A'|\mathbf{b}')$ be the result of applying r to the augmented matrix $(A|\mathbf{b})$. Then the vector $\mathbf{v}$ is a solution of $Ax = \mathbf{b}$ if and only if it is a solution of $A'x = \mathbf{b}'$.

3 A systematical way of solving linear systems.

Definition 12. The left-most non-zero entry in a non-zero row is called the **leading entry** of that row.

Definition 13. A matrix is in **echelon form** if

1. the leading entry in each non-zero row is 1,
2. the leading 1 in each non-zero row is to the right of the leading 1 in any row above it,
3. the zero rows are below any non-zero rows.

Definition 14. A matrix is in **row reduced echelon form (RRE)** if

1. it is in echelon form,
2. the leading entry in each non zero row is the only non-zero entry in its column.

Solution algorithm

If we have a system of equations

$$Ax = b$$

and A is in RRE form, then we can easily read off the solutions (if any exist). There are four cases to consider.

- **Case 1.** Every column of A contains a leading 1, and there are no zero rows. In this cas the only possibility is that $A = I_n$. Then the equations are simply

$$\begin{aligned} x_1 &= b_1 \\ x_2 &= b_2 \\ &\vdots \\ x_n &= b_n \end{aligned} \tag{2}$$

  and they have a unique solution, i.e. the entries of b.

- **Case 2.** Every column of A contains a leading 1, and there are some zero rows. Then A must have more rows than column, and it must be a matrix of the from

$$A = \begin{pmatrix} I_n \\ \mathbf{0}_{k\times n} \end{pmatrix} \tag{3}$$

  Which looks like an identity matrix with a block of zero rows underneath. In this case the respective equations are

$$\begin{aligned} x_1 &= b_1 \\ &\vdots \\ x_n &= b_n \end{aligned} \tag{4}$$

And the last k of them are

$$\begin{aligned} 0 &= b_{n+1} \\ &\vdots \\ x_n &= b_{n+k}. \end{aligned} \tag{5}$$

Now there are two possibilities:

1. If any of the last k entries of $\mathbf{b}$ are non-zero then this system has no solutions, i.e. it is inconsistent.
2. If the last k entries of $\mathbf{b}$ are all zero then the system has a unique solution, given by setting $x_i = b_i \; \forall \; i \in [1, n]$.

- **Case 3.** Some columns of A do not contain a leading 1, but there are no zero rows. For example

$$A = \begin{pmatrix} 1 & a_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{6}$$

  If the ith column of A does not contain a leading 1, then the corresponding x_i is called a **free variable**. The remaining variables are called basic variables. This kind of system has infinitely many solutions, we call such system underdetermined.

See the full version of the notes for a step by step proof-algorithm how to turn any matrix into RRE form.

Now we have a systematic procedure for solving a system of simultaneous linear equations $Ax = \mathbf{b}$.

1. Form the augmented matrix $(A|\mathbf{b})$.
2. Apply the algorithm to put the augmented matrix into RRE form $(A'|\mathbf{b}')$.
3. Read off the solutions to $A'x = \mathbf{b}'$.

Note 2. In fact we only need to put the left block A' to be able to read off the solutions.

Proposition 2. The number of solutions to a system $Ax = \mathbf{b}$ is always either 0, 1 or ∞.

4 Matrix Multiplication

Remark 1. The easiest way to remember how to multiply matrices is as follows
In order to multiply AB where $A \in M_{n\times m}(\mathbb{R})$ and $B \in M_{m\times p}(\mathbb{R})$ we write $r_1, r_2, ..., r_n \in \mathbb{R}^m$ for the rows of A and $c_1, c_2, ..., c_n \in \mathbb{R}^m$ for the columns of B. Then the definitions states that the (i, j) entry of $AB \in M_{n\times p}(\mathbb{R})$ is the dot product

$$r_i^T \cdot c_j$$

We can view this as in order to determine one column of AB we take the respective column of B and scan A with it from top to bottom multiplying each row of A by the current column.

Definition 15. A matrix $A \in M_{n\times m}(R)$ defines a function from $\mathbb{R}^m$ to $\mathbb{R}^n$, which we denote by

$$T_A : \mathbb{R}^m \to \mathbb{R}^n$$
$$\mathbf{v} \mapsto A\mathbf{v}.$$

Provided that two matrices have correct dimensions, we can compose the functions defined by them.

Lemma 1. *$T_A \circ T_B = T_{AB}$, i.e for any $\mathbf{v} \in \mathbb{R}^p$ we have*

$$A(B\mathbf{v}) = (AB)\mathbf{v}.$$

Where $A \in M_{n\times m}(\mathbb{R})$ and $B \in M_{m\times p}(\mathbb{R})$

4.1 Matrix multiplication properties.

Proposition 3. Let $A, A' \in M_{m\times n}(\mathbb{R})$, let $B, B' \in M_{n\times p}(\mathbb{R})$ let $C \in M_{p\times q}(\mathbb{R})$ Then the following holds.

1. $A(BC) = (AB)C$ (associativity).
2. $A(B + B') = AB + AB'$

   and

   $(A + A')B = AB + A'B$

   (left and right distributivity of multiplication over addition.)
3. $\forall\, \lambda \in \mathbb{R}(\lambda A)B = \lambda(AB) = A(\lambda B)$.

The usual rules about multiplication by zero and one translate onto matrix multiplication with a certain degree of caution.

Lemma 2. *Let $A \in M_{n\times m}(\mathbb{R})$.*

1. *$\forall\, k \in \mathbb{N}\ \mathbf{0}_{k\times n}A = \mathbf{0}_{k\times m}$ and $A\mathbf{0}_{m\times k} = \mathbf{0}_{n\times k}$.*
2. *$I_nA = A = AI_m$.*

Lemma 3. *Let $diag(d_1, ..., d_n)$ and $diag(d'_1, ..., d'_n)$ be two diagonal matrices $\in M_{n\times n}(\mathbb{R})$. Then their product is $diag(d_1d'_1, ..., d_nd'_n)$.*

5 Inverse of a matrix

Definition 16. Let $A \in M_{n\times m}(\mathbb{R})$ An $n \times n$ matrix A^{-1} such that

$$AA^{-1} = A^{-1}A = I_n$$

is called an **inverse** of A. A matrix A is called **invertible** if the inverse exists, **singular** if not.

Lemma 4. *1. If A is invertible, then the inverse is unique.*

2. If A is invertible and either $AB = I_n$ or $BA = I_n$ for some $B \in M_{n\times n}(\mathbb{R})$, then $B = A^{-1}$.

Lemma 5. *Suppose $A, B \in M_{n\times n}(\mathbb{R})$ are invertible. Then AB is invertible and*

$$(AB)^{-1} = B^{-1}A^{-1}.$$

Lemma 6. *Let $A \in M_{n\times n}(\mathbb{R})$.*

1. *If there exists $\mathbf{v} \neq 0 \in \mathbb{R}^n$ such that $A\mathbf{v} = \mathbf{0}$ then A is not invertible.*

2. *If there exists a non zero matrix $B \in M_{n\times n}(\mathbb{R})$ such that $AB = \mathbf{0}_{n\times n}$ or $BA = \mathbf{0}_{n\times n}$ then A is not invertible.*

Corollary 1. If $A \in M_{n\times n}(\mathbb{R})$ has a column of zeros, then A is not invertible.

Corollary 2. If $A \in M_{n\times n}(\mathbb{R})$ has a row of zeros, then A is not invertible.

Lemma 7. *If $A \in M_{n\times n}(\mathbb{R})$ and A' is obtained from A by a row operation, then A' is invertible if and only if A is.*

Lemma 8. *If $A \in M_{n\times n}(\mathbb{R})$ and A' is the RRE form of A , then A' is invertible if and only if it has no zero rows.*

Lemma 9. *A matrix $A \in M_{n\times n}(\mathbb{R})$ is invertible if and only if its RRE form is the identity matrix.*

Proposition 4. A matrix $A \in M_{n\times n}(\mathbb{R})$ is invertible if and only if there is no non-zero vector $\mathbf{v} \in \mathbb{R}^n$ such that $A\mathbf{v} = \mathbf{0}$.

A following algorithm can be used for computing the inverse of a matrix.

1. Write the augmented matrix $(A|I_n)$.

2. Row reduce $(A|I_n)$ and bring it into RRE form.

3. If A is invertible this process will transform the A to its RRE form and I_n to A^{-1}.

Remark 2. The algorithm described above gives us another method for finding solutions of linear systems of equations. Assuming we have such a system in matrix form $Ax = b$, and A is an invertible matrix, we can compute A^{-1} and rewrite our system as $x = A^{-1}b$. Note that it works if and only if A is invertible and therefore the system has a unique solution.

6 Vector spaces

Definition 17. A **vector space** is the following data:

- A set V. We will refer to the elements of V as vectors.
- A binary operation $+ : V \times V \to V$, which we call addition.
- A function $\mathbb{R} \times V \to V$, which we call scalar multiplication. We usually omit the symbol of it.

We require that the following axioms hold:

1. The set V with the binary operation $+$ forms an Abelian group. We denote the identity element of it as $\mathbf{0}_V$
2. $\forall\, \mathbf{v} \in V$ we have $1\mathbf{v} = \mathbf{v}$.
3. $\forall\, \lambda, \mu \in \mathbb{R}\, \forall\, \mathbf{v} \in V$ we have $\lambda(\mu\mathbf{v}) = (\lambda\mu)\mathbf{v}$.
4. $\forall\, \lambda, \in \mathbb{R}\, \forall\, \mathbf{u}, \mathbf{v} \in V$ we have $\lambda(\mathbf{u} + \mathbf{v}) = \lambda\mathbf{u} + \lambda\mathbf{v}$.

   (distributivity of scalar multiplication over addition)
5. $\forall\, \lambda, \mu \in \mathbb{R}\, \forall\, \mathbf{v} \in V$ we have $(\lambda + \mu)\mathbf{v} = \lambda\mathbf{v} + \mu\mathbf{v}$.

   (distributivity of scalar multiplication over scalar addition)

Lemma 10. *Let V be a vector space and let $\mathbf{x} \in V$.*

1. *For any positive integer $n \in \mathbb{R}$, we have*

$$n\mathbf{x} = \mathbf{x} + \mathbf{x} + ... + \mathbf{x}$$

   where there are n terms on the right-hand-side.
2. $0\mathbf{x} = (0)_V$.
3. $(-1)\mathbf{x}$ *is the additive inverse of* $\mathbf{x}$.

7 Subspaces

Definition 18. Let V be a vector space. A subset $U \subset V$ is called a subspace of V if:

1. If $\mathbf{x} + \mathbf{y} \in U$ then $\mathbf{x} + \mathbf{y} \in U$ (Closure under vector addition).
2. $\mathbf{0}_V \in U$.
3. If $\mathbf{x} \in U$ then $\forall\, \lambda \in \mathbb{R}.\ \lambda\mathbf{x} \in U$.

8 Spanning sets

Definition 19. Let $\mathcal{S} \subset V$ be any subset. A linear combination of elements in $\mathcal{S}$ is a vector $\mathbf{x} \in V$ which can be written as

$$\mathbf{x} = \lambda_1 \mathbf{v_1} + \lambda_2 \mathbf{v_2} + ... + \lambda_n \mathbf{v_n}$$

where $\mathbf{v_1}, \mathbf{v_2}, ..., \mathbf{v_n}$ are vectors in $\mathcal{S}$, and $\lambda_\mathbf{1}, \lambda_\mathbf{2}, ..., \lambda_\mathbf{n}$ are real numbers.

Lemma 11. *A non-empty subset U of a vector space V is a subspace if and only if every linear combination of elements of U is again in U.*

Definition 20. Let V be a vector space and let $\mathcal{S} \subset V$ be a non-empty subset.
Then **span** of $\mathcal{S}$ written

$$\text{span } \mathcal{S} \subset V$$

is the set of all linear combinations of elements of $\mathcal{S}$. If $\mathcal{S} = \emptyset$ then we define span $\mathcal{S}$ to be $\{\mathbf{0}_V\}$.

Lemma 12. *Let V be a vector space and let $\mathcal{S} \subset V$ be any subset. Then* span $\mathcal{S}$ *is a subspace of V.*

Definition 21. Let V be a vector space. A subset $\mathcal{S} \subset V$ is called a **spanning set** if span $\mathcal{S} = V$.

Lemma 13. *Let V be a vector space, and let $\mathcal{S} \subset V$ be any subset. Suppose $\mathcal{S} \subset U$ for some subspace $U \subset V$. Then* span $\mathcal{S} \subset U$*. So if $\mathcal{S}$ is a spanning set for V, then $\mathcal{S}$ cannot be contained in any proper subspace $U \subsetneq V$*

Definition 22. A vector space is called **finite-dimensional** if it has a finite spanning set.

Definition 23. Let V be a finite-dimensional vector space. The dimension of V, denoted as dim V, is the smallest $n \in \mathbb{N}$ such that V has a spanning set of size n.

9 Linear independence

Definition 24. A subset $\mathcal{L}$ of a vector space V is called linearly dependent if we can find **distinct** vectors $\mathbf{v_1}, \mathbf{v_2}, ..., \mathbf{v_n} \in \mathcal{L}$ and non-zero scalars $\lambda_1 \neq 0,\ \lambda_2 \neq 0,\ \dots, \lambda_n \neq 0$, such that

$$\lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2 + ... + \lambda_n \mathbf{v}_n = \mathbf{0}_V.$$

Note 3. Any subset of a linearly independent set is also linearly-independent.

Definition 25. A **basis** of a vector space is a linearly independent spanning set.

Proposition 5. Let $\mathcal{B} = \{\mathbf{v_1}, ..., \mathbf{v_n}\} \subset V$ be a finite basis of a vector space V. Then every vector in V can be written as as linear combination of elements in $\mathcal{B}$, in a unique way. Conversely, any finite subset $\mathcal{B}$ with this property is a basis.

10 Bases an dimension

Lemma 14. *Let $\mathcal{S} \subset V$ be a spanning set, and suppose that $\mathcal{S}$ is not linearly independent. Then there exists a vector $\mathbf{x} \in \mathcal{S}$ such that $\mathcal{S}' = \mathcal{S} \setminus \{\mathbf{x}\}$ is still a spanning set.*

Corollary 3. Any finite spanning set contains a basis.

Corollary 4. Any finite-dimensional vector space V has a basis.

Proposition 6 (Steinitz exchange lemma - easy verion)**.** Let $\mathcal{S} \subset V$ be a spanning set, and let $\mathbf{x} \in V$ be any non-zero vector. Then there exists a vector $\mathbf{y} \in \mathcal{S}$ such that the set

$$\mathcal{S}' = (\mathcal{S} \setminus \{\mathbf{y}\}) \cup \{\mathbf{x}\}$$

is still a spanning set.

Proposition 7 (Steinitz exchange lemma - full verion)**.** Let $\mathcal{S} \subset V$ be a spanning set, and let $\mathcal{L}\{\mathbf{x}_1, ..., \mathbf{x}_n\}$ be a finite linearly independent subset of V. Then there exists a subset $\mathcal{T}\{\mathbf{y}_1, ..., \mathbf{y}_n\} \subset \mathcal{S}$, with the same size as $\mathcal{L}$, such that

$$\mathcal{S}' = (\mathcal{S} \setminus \mathcal{T}) \cup \mathcal{L}$$

is still a spanning set.

Corollary 5. Let V be a finite-dimensional vector space, let $\mathcal{S} \subset V$ be a finite spanning set and let $\mathcal{L} \subset V$ be a linearly independent subset. Then $\mathcal{L}$ is finite and $\#\mathcal{L} \leq \#\mathcal{S}$.

Theorem 1. *Let V be a finite-dimensional vector space with* $\dim V = n$*. Then any basis of V is finite and has size n.*

11 Dimensions of subspaces

Lemma 15. *Suppose $\mathcal{L} \subset V$ is a linearly independent subset. Let $\mathbf{v} \in V$ be a vector which does not lie in* span $\mathcal{L}$*. Then $\mathcal{L} \cup \{\mathbf{v}\}$ is linearly independent.*

Lemma 16. *If V is not finite-dimensional, then for any $n \in \mathbb{N}$ we can find linearly independent subset $\mathcal{L} \subset V$ of size n.*

Lemma 17. *Let V be a finite-dimensional vector space with* $\dim V = n$*. Then any linearly independent subset $\mathcal{L} \subset V$ of size n must be a basis.*

Lemma 18. *If V is finite-dimensional then any linearly independent subset is contained in a basis.*

Proposition 8. Let V be a finite-dimensional vector space and let $U \subset V$ be a subspace.

1. U is finite-dimensional.
2. $\dim U \leq \dim V$.
3. If $\dim U = \dim V$ then $U = V$.

12 Linear maps

Definition 26. Let U and V be vector spaces. A function $f : U \to V$ is called a **linear map** if

- $f(\mathbf{x} + \mathbf{y}) = f(\mathbf{x}) + f(\mathbf{y})$ for all $\mathbf{x}, \mathbf{y} \in U$,
- $f(\lambda \mathbf{x}) = \lambda f(\mathbf{x})$ for all $\mathbf{x} \in U$ and all $\lambda \in \mathbb{R}$.

Lemma 19. *If $f : U \to V$ is linear then $f(\mathbf{0}_U) = \mathbf{0}_V$.*

Lemma 20. *A function composition of linear maps is also a linear map.*

Definition 27. Let $f : U \to V$ be a linear map.

- The **image** of f, denoted as $\operatorname{Im} f$ is defined to be the subset

$$\{f(\mathbf{x}) \,|\, \mathbf{x} \in U\} \subset V$$

- The **ker** of f, denoted as $\ker f$ is defined to be the subset

$$\{\mathbf{x} \in U \,|\, f(\mathbf{x}) = \mathbf{0}_V\} \subset U$$

Lemma 21. *Let $f : U \to V$ be a linear map between vector spaces. Then $\ker f$ is a subspace of U and $\operatorname{Im} f$ is a subspace of V.*

Lemma 22. *A linear map $f : U \to V$ is injective if and only if $\ker f = \{\mathbf{0}_U\}$.*

Lemma 23. *Let $f : U \to V$ be a linear map, and fix $\mathbf{y} \in V$. Suppose $\mathbf{x} \in U$ is such that $f(\mathbf{x}) = \mathbf{y}$. Then*

$$f^{-1}(\mathbf{y}) = \{\mathbf{x} + \mathbf{v} \,|\, \mathbf{v} \in \ker f\}$$

13 Linear maps and bases

Proposition 9. Let $f : \mathbb{R}^k \to \mathbb{R}^n$ be linear. Then $f = T_A$ for some matrix $A \in M_{n \times k}(\mathbb{R})$

Proposition 10. Let $f : U \to V$ and $g : U \to V$ be two linear maps, and let $\mathcal{B} = \{\mathbf{b}_1, ..., \mathbf{b}_k\}$ be a basis for U. Suppose $f(\mathbf{b}_i) = g(\mathbf{b}_i)$ for each $i = 1, ..., k$. Then $f = g$.

Proposition 11. Let U and V be vector spaces. Let $\mathcal{B} = \{\mathbf{b}_1, ..., \mathbf{b}_k\}$ be a basis for U, and let $\{\mathbf{v}_1, ..., \mathbf{v}_k\}$ be any set of k vectors in V. Then there is a unique linear map $f : U \to V$ such that $f(\mathbf{b}_i) = \mathbf{v}_i$ for each i.

14 Isomorphisms

Definition 28. A linear map $f : U \to V$ between two vector spaces is called an **isomorphism** if f is bijective. If there exists an isomorphism from U to V we say that U is isomorphic to V and write

$$U \cong V.$$

Proposition 12. Let V be a vector space with $\dim V = n$. Then V is isomorphic to $\mathbb{R}^n$.

Lemma 24. *Let $f : U \to V$ be a linear map, and let $\mathcal{B} = \{\mathbf{b}_1, ..., \mathbf{b}_k\}$ be a basis for U. Let $\mathcal{C} = \{f(\mathbf{b}_1), ..., f(\mathbf{b}_k)\} \subset V$. Then:*

1. *$\mathcal{C}$ is a spanning set if and only if f is surjective.*

2. *$\mathcal{C}$ is linearly independent if and only if f is injective.*

3. *$\mathcal{C}$ is a basis if and only if f is an isomorphism.*

Corollary 6. If $U \cong V$ then $\dim U = \dim V$.

Corollary 7. Let $f : U \to V$ be a linear map, and suppose $\dim U = \dim V$. Then the following are equivalent:

1. f is injective.

2. f is surjective.

3. f is an isomorphism.

Corollary 8. If $f : \mathbb{R}^n \to V$ is an isomorphism, then the set

$$\mathcal{C} = \{f(\mathbf{e}_1), ..., f(\mathbf{e}_n)\}$$

is a basis for V.

15 The Rank-Nullity Theorem

Definition 29. Let U and V be vector spaces and let $f : U \to V$ be a linear map. Then

- The **Rank** of f denoted as $\operatorname{rank} f$ is defined as $\dim \operatorname{Im} f$.

- The **Nullity** of f denoted as $\operatorname{nullity} f$ is defined as $\dim \ker f$.

Theorem 2 (Rank-Nullity theorem)**.** *Let U and V be vector spaces and let $f : U \to V$ be a linear map. Then*

$$\operatorname{rank} f + \operatorname{nullity} f = \dim U$$

.

16 Linear maps and matrices

Suppose U and V are vector spaces and $f : U \to V$ is a linear map. We want to associate a matrix with f as we did before for a linear map between $\mathbb{R}^n$ and $\mathbb{R}^k$ In order to do it let

$$\mathcal{B} = \{\mathbf{b}_1, ..., \mathbf{b}_k\} \subset U \text{ and } \mathcal{C} = \{\mathbf{c}_1, ..., \mathbf{c}_n\} \subset V$$

be bases. We have seen in Proposition 12 that this gives us the following isomorphisms

$$F_{\mathcal{B}} : \mathbb{R}^k \to U \text{ and } F_{\mathcal{C}} : \mathbb{R}^n \to V$$

Now by Lemma 20 the following composition of linear maps

$$F_{\mathcal{C}}^{-1} \circ f \circ F_{\mathcal{B}}$$

is also a linear map. Therefore it must be given by some matrix $A \in M_{n\times k}(\mathbb{R})$ such that

$$F_{\mathcal{C}}^{-1} \circ f \circ F_{\mathcal{B}}(\mathbf{v}) = A\mathbf{v}$$

This matrix A is called the **matrix representing f with respect to $\mathcal{B}$ and $\mathcal{C}$**, we denote it as:

$${}_{\mathcal{C}}[f]_{\mathcal{B}} \text{ or } [f]_{\mathcal{B}}^{\mathcal{C}}$$

To compute the matrix ${}_{\mathcal{C}}[f]_{\mathcal{B}}$ we can use the fact that the product $A\mathbf{e}_j$ is given by the jth column of A. Therefore the jth column of ${}_{\mathcal{C}}[f]_{\mathcal{B}}$ is the vector

$$F_{\mathcal{C}}^{-1} \circ f \circ F_{\mathcal{B}}(\mathbf{e}_j) = F_{\mathcal{C}}^{-1} \circ f(\mathbf{b}_j) \in \mathbb{R}^n.$$

The procedure of finding ${}_{\mathcal{C}}[f]_{\mathcal{B}}$ is as follows

- For each $j = 1, ..., k$, take the jth basis vector $\mathbf{b}_j \in \mathcal{B}$, and apply the map f to it to get a vector $f(\mathbf{b}_j) \in V$

- Express each $f(\mathbf{b}_j)$ as a linear combination of vectors in $\mathcal{C}$

  $$f(\mathbf{b}_j) = a_{1j}\mathbf{c}_1 + a_{2j}\mathbf{c}_2 + ... + a_{nj}\mathbf{c}_n$$

  for some scalars $a_{1j}, ..., a_{nj} \in \mathbb{R}$.

  Now $F_{\mathcal{C}}$ is a map that takes a vector in $\mathbb{R}^n$ and returns a linear combination of the elements $\mathcal{C}$ (the basis of V) where scalar of the ith basis vector of V is the ith entry in the vector. Therefore

  $$F_{\mathcal{C}}^{-1}(f(\mathbf{b}_j)) = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{nj} \end{pmatrix}, \tag{7}$$

  so ${}_{\mathcal{C}}[f]_{\mathcal{B}}$ is the matrix (aij).

Definition 30. Let V be a vector space, and let $\mathcal{B} = \{\mathbf{b}_1, ..., \mathbf{b}_k\}$ and $\mathcal{C} = \{\mathbf{c}_1, ..., \mathbf{c}_n\}$ be two bases for V. The **change-of-basis matrix** from $\mathcal{B}$ to $\mathcal{C}$ is the matrix

$$_{\mathcal{C}}[\mathrm{id}_V]_{\mathcal{B}}$$

That represents the identity map with respect to $\mathcal{B}$ and $\mathcal{C}$.

Lemma 25. *Let V be a vector space, and let $\mathcal{B} = \{\mathbf{b}_1, ..., \mathbf{b}_k\}$ and $\mathcal{C} = \{\mathbf{c}_1, ..., \mathbf{c}_n\}$ be two bases for V, and let $P =_{\mathcal{C}} [id_V]_{\mathcal{B}}$ be the change-of-basis matrix. Pick any $\mathbf{x} \in V$. If the coefficients of $\mathbf{x}$ with respect to $\mathcal{B}$ are the vector $\mathbf{v} \in \mathbb{R}$, then the coefficients of $\mathbf{x}$ with respect to $\mathcal{C}$ are given by the vector*

$$P\mathbf{v}.$$

Proposition 13 (Change-of-basis formula)**.** Let $f : U \to V$ be a linear map, Let $\mathcal{B}$ and $\mathcal{B}'$ be two bases for U, and let $_{\mathcal{B}}P_{\mathcal{B}'}$ be the change-of-basis matrix between them. Also let $\mathcal{C}$ and $\mathcal{C}'$ be two bases for V, and let $_{\mathcal{C}'}P_{\mathcal{C}}$ be the change-of-basis matrix between them. Then the matrices representing f with respect to $\mathcal{B}$ and $\mathcal{C}$ or with respect to $\mathcal{B}'$ and $\mathcal{C}'$ are related by

$$_{\mathcal{C}'}[f]_{\mathcal{B}'} = {_{\mathcal{C}'}P_{\mathcal{C}}}\ {_{\mathcal{C}}[f]_{\mathcal{B}}}\ {_{\mathcal{B}}P_{\mathcal{B}'}}.$$

17 Determinants

Definition 31. Let A be an $n \times n$-matrix, and A_{ji} be the submatrix obtained by deleting the i-th row and j-th column of A. A_{ij} is called the minor of A.

Definition 32. The determinant of an $n \times n$-matrix $A = (a_{ij})$ is given by

$$\det A = \sum_j (-1)^{i+j}\, a_{ij} \det A_{ij} = \sum_i (-1)^{i+j}\, a_{ij} \det A_{ij}$$

The first sum is called expansion along the i-th row, the second is called expansion along the j-th column.

Definition 33. Let $\sigma \in S_n$ be a permutation. An **inversion** (i, j) of σ is a pair of positive integers i, j such that $i < j$, but $\sigma(i) \geq \sigma(j)$. The **sign** of a permutation is defined to be

$$sgn(\sigma) := \begin{cases} 1, & \text{if the number of inversions in } \sigma \text{ is even.} \\ -1, & \text{if the number of inversions in } \sigma \text{ is odd.} \end{cases}$$

Definition 34. Let $A = (a_{ij}) \in M_{n\times n}(\mathbb{R})$. The **determinant** of A is the number:

$$\det(A) = \sum_{\sigma \in S_n} sgn(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} ... a_{n\sigma(n)}.$$

Proposition 14. Let A be a square matrix. Then

1. If 2 rows of A are swapped to produce a matrix B, then det $B = -$ det A.
2. If one row of A is multiplied by a scalar λ to produce a matrix B, then det $B = \lambda \det A$.
3. If a multiple of one row of A is added to another row to produce a matrix B, then det $B =$ det A.

Proposition 15. Let A be a square matrix. Then

1. $\det A = \det A^T$.
2. $\det AB = \det A \det B$.

Proposition 16. An $n \times n$-matrix A is invertible if and only if det $A \neq 0$.

18 Eigenvalues and eigenvectors

Definition 35. Let A be a $n \times n$- matrix, and let $\lambda \in \mathbb{R}$ be a number. We say that λ is an **eigenvalue** of A if there exists a non-zero vector $\mathbf{v} \in \mathbb{R}^n$ satisfying

$$A\mathbf{v} = \lambda \mathbf{v}$$

We call such a $\mathbf{v}$ an **eigenvector** of A (for the eigenvalue λ).

Remark 3. Let $A \in M_{n \times n}(\mathbb{R})$. Obviously, the non-zero vector $\mathbf{v}$ is an eigenvector of A with eigenvalue λ if and only if $(A - \lambda I_n)\mathbf{v} = \mathbf{0}$.

Proposition 17. Let $A \in M_{n \times n}(\mathbb{R})$. Then λ is an eigenvalue of A if and only if the matrix $A - \lambda I_n$ is not invertible.

Definition 36. Let $A \in M_{n \times n}(\mathbb{R})$. The determinant

$$p_A(\lambda) := \det(A - \lambda I_n)$$

is called the characteristic polynomial.

19 Diagonalization

Definition 37. Let A and B be $n \times n$-matrices. We call A similar to B if there exists an invertible matrix P such that $A = P^{-1}BP$.

Lemma 26. *Let A and B be $n \times n$-matrices. If A is similar to B, then B is similar to A.*

Proposition 18. Similar matrices have the same eigenvalues.

Definition 38. A matrix A is called diagonalizable if it is similar to a diagonal matrix.

Lemma 27. *Let $A \in M_{n\times n}(\mathbb{R})$. Suppose we can find n eigenvectors $\mathbf{v}_1, ..., \mathbf{v}_n$ of A, with corresponding eigenvalues $\lambda_1, ..., \lambda_n$. Let*

$$P := (\mathbf{v}_1|\mathbf{v}_2|...|\mathbf{v}_n),$$

i.e. the matrix whose columns are the vectors $\mathbf{v}_i$. If P is invertible, then A is diagonalizable and the corresponding diagonal matrix is given by

$$P^{-1}AP = diag(\lambda_1, ..., \lambda_n).$$

Definition 39. Let $A \in M_{n\times n}(\mathbb{R})$. The eigenspace E_λ of A is the set

$$E_\lambda := \{v \in \mathbb{R}^n | Av = \lambda v\}.$$

Note that this set includes all of the eigenvectors of A plus the zero vector.

Proposition 19. Any eigenspace E_λ of a matrix $A \in M_{n\times n}(\mathbb{R})$ is a subspace of $\mathbb{R}^n$.

Remark 4. The proposition above is not surprising as

$$E_\lambda = \ker(T_{A-\lambda I_n}) = span\{v_1, ..., v_k\},$$

where $v_1, ..., v_k$ are the eigenvectors to the eigenvalue λ.

20 Gram-Schmidt Process and orthogonality

Definition 40. A set of vectors $S = \{\mathbf{v}_1, ..., \mathbf{v}_n\} \subset \mathbb{R}^n$ is called an orthogonal set if each pair of distinct vectors in the set is orthogonal:

$$v_i \cdot v_j = 0 \; \forall i, j, i \neq j.$$

The set S is called orthonormal if it is orthonormal and all vectors in the set are unit vectors.

Proposition 20. Let $\{\mathbf{v}_1, ..., \mathbf{v}_n\}$ be an orthogonal set. Then $\{\mathbf{v}_1, ..., \mathbf{v}_n\}$ is linearly independent.

Definition 41. Let $\{\mathbf{v}_1, ..., \mathbf{v}_k\}$ be an orthogonal set of vectors in $\mathbb{R}^n$ and $U = span\{v_1, ..., v_k\}$. Let $y \in \mathbb{R}^n$. The orthogonal projection of y onto U is the vector given by

$$proj_U \; y := \frac{y \cdot v_1}{||v_1||^2} v_1 + \frac{y \cdot v_2}{||v_2||^2} v_2 + ... + \frac{y \cdot v_k}{||v_k||^2} v_k.$$

Theorem 3 (Gram-Schmidt Process)**.** *Let $\{\mathbf{v}_1, ..., \mathbf{v}_k\}$ be a linearly independent set of vectors in $\mathbb{R}^n$.*

Define inductively, $i \leq k$:

$$w_1 := v_1 \tag{8}$$

$$w_2 := v_2 - proj_{span\{w_1\}}\, v_2 = v_2 - \frac{v_2 \cdot w_1}{||w_1||^2} w_1 \tag{9}$$

$$w_3 := v_3 - proj_{span\{w_1, w_2\}}\, v_3 = v_3 - \left(\frac{v_3 \cdot w_1}{||w_1||^2} w_1 + \frac{v_3 \cdot w_2}{||w_2||^2} w_2 \right) \tag{10}$$

$$\ldots \tag{11}$$

$$w_i := v_i - proj_{span\{w_1, w_2, \ldots, w_{i-1}\}}\, v_i = v_i - \left(\frac{v_i \cdot w_1}{||w_1||^2} w_1 + \frac{v_i \cdot w_2}{||w_2||^2} w_2 + \ldots + \frac{v_i \cdot w_{i-1}}{||w_{i-1}||^2} w_{i-1} \right) \tag{12}$$

$$\tag{13}$$

Then

1. $\{w_1, ..., w_k\}$ is an orthogonal set.

2. Let $u_i = \frac{w_i}{||w_i||}$. Then $\{u_1, ..., u_k\}$ is an orthonormal set.

3. $span\{w_1, ..., w_k\} = span\{u_1, ..., u_k\} = span\{v_1, ..., v_k\}$.

Definition 42. A matrix is called orthogonal if its columns as an orthogonal set. It is called orthonormal if it is orthogonal and the columns are unit vectors.

Note 4. Since all columns of an orthogonal matrix are orthogonal, they are linearly independent and therefore an orthogonal matrix is always invertible.

Proposition 21. A matrix $A \in M_{n \times n}(\mathbb{R})$ is orthogonal if and only if $AA^T = I_n$.

21 Real symmetric matrices and orthogonality

Definition 43. A matrix $A \in M_{n \times n}(\mathbb{R})$ is called symmetric if $A = A^T$.

Proposition 22. Let $A \in M_{n \times n}(\mathbb{R})$, A symmetric. Then A has at least one real eigenvalue, and all its eigenvectors are real.

Lemma 28. *Let $A \in M_{n \times n}(\mathbb{R})$,$A$ symmetric. Let λ_1, λ_2 be distinct eigenvalues of A with corresponding eigenvectors v_1 and v_2. Then v_1 and v_2 are orthogonal.*

Theorem 4. *Let $A \in M_{n \times n}(\mathbb{R})$,$A$ symmetric. Then A is diagonalizable, i.e. there exists a diagonal matrix D and an invertible matrix P with $A = PDP^{-1}$ such that the matrix P of eigenvectors is orthogonal. We call such a matrix orthogonally diagonalizable.*

Remark 5.

1. From theorem 4 we can deduce that a symmetric matrix always has a basis of real eigenvectors. If all the eigenvalues are distinct, by Lemma 28, we get immediately that the eigenvectors are pairwise orthogonal.

2. Not every diagonalizable matrix is orthogonally diagonalizable. That is because the Gram-Schmidt process doesn't preserve eigenvectors in general. Actually a matrix is orthogonally diagonalizable if and only if it is symmetric.

Theorem 5 (Spectral Theorem). *A matrix $A \in M_{n\times n}(\mathbb{R})$ is orthogonally diagonalizable if and only is it is symmetric.*