

Mathematics Year I, Linear Algebra Term 1, 2

Most Important theorems, definitions and propositions.

Szymon Kubica

April 16, 2021

1 Introduction to matrices and vectors

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

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

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

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

is called linear combination of the vectors v_1, \dots, v_n .

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

$$\text{span}\{v_1, \dots, v_n\}$$

.

Note 1. \mathbb{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

$$\text{dist}(u, v) := \|u - v\|$$

.

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

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

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 $(1, 1), (2, 2) \dots$ 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 i th entry is b_i . Then (v_1, \dots, v_n) is a solution to the system

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

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 i th 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 case 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 form

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

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 i th 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 $r_1, r_2, \dots, r_n \in \mathbb{R}^m$ for the rows of A and $c_1, c_2, \dots, c_n \in \mathbb{R}^m$ 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 $A \in M_{n \times m}(\mathbb{R})$ defines a function from \mathbb{R}^m to \mathbb{R}^n , which we denote by

$$\begin{aligned} T_A : \mathbb{R}^m &\rightarrow \mathbb{R}^n \\ \mathbf{v} &\mapsto A\mathbf{v}. \end{aligned}$$

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.

$$1. A(BC) = (AB)C \text{ (associativity).}$$

$$2. A(B + B') = AB + AB'$$

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 $A \in M_{n \times m}(\mathbb{R})$.

$$1. \forall k \in \mathbb{N} \mathbf{0}_{k \times n}A = \mathbf{0}_{k \times m} \text{ and } A\mathbf{0}_{m \times k} = \mathbf{0}_{n \times k}.$$

$$2. I_n A = A = A I_m.$$

Lemma 3. Let $\text{diag}(d_1, \dots, d_n)$ and $\text{diag}(d'_1, \dots, d'_n)$ be two diagonal matrices $\in M_{n \times n}(\mathbb{R})$.

Then their product is $\text{diag}(d_1 d'_1, \dots, d_n d'_n)$.

5 Inverse of a matrix

Definition 16. Let $A \in M_{n \times n}(\mathbb{R})$ 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 $\mathbf{v} \neq \mathbf{0} \in \mathbb{R}^n$ 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 $+$: $V \times V \rightarrow V$, which we call addition.
- A function $\mathbb{R} \times V \rightarrow V$, 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. $\forall \lambda, \mu \in \mathbb{R} \forall \mathbf{u}, \mathbf{v} \in V$ 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

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

where there are n terms on the right-hand-side.

2. $0\mathbf{x} = (\mathbf{0})_V$.
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

$$\mathbf{x} = \lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2 + \dots + \lambda_n \mathbf{v}_n$$

where $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ are vectors in \mathcal{S} , and $\lambda_1, \lambda_2, \dots, \lambda_n$ 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 $\text{span } \mathcal{S}$ to be $\{\mathbf{0}_V\}$.

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

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 $\text{span } \mathcal{S} \subset U$. 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 $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n \in \mathcal{L}$ and non-zero scalars $\lambda_1 \neq 0, \lambda_2 \neq 0, \dots, \lambda_n \neq 0$, such that

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

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 $\mathcal{B} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\} \subset V$ be a finite basis of a vector space V . Then every vector in V can be written as a 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 and 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 version). 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 version). Let $\mathcal{S} \subset V$ be a spanning set, and let $\mathcal{L} = \{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ be a finite linearly independent subset of V . Then there exists a subset $\mathcal{T} = \{\mathbf{y}_1, \dots, \mathbf{y}_n\} \subset \mathcal{S}$, 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