Linear Algebra I & II Summary

Paul Kim

June 11, 2023

Preface

Note that this is not your actual lecture note, and I do not condone only sticking to this as your resource. On the other hand, if you are reading this for your intuition, you are at the right place.

This document is based on 2019 linear algebra I and 2018 linear algebra II lecture notes by Vicky Neale and Ulrike Tillmann. You can check them out at `https://courses.maths.ox.ac.uk/` in the Archive tab.

Most examples here will include the ones from these lecture notes.

Note that this summary note will gloss over quite many of the proofs, but will try to give geometric intuition behind them.

Part I
Linear Algebra I

1 Linear Equations and Matrices

You've probably already seen linear system of equations like the following:

$$\begin{cases} 3x + 5y = -1 \\ 4x - y = 10 \end{cases} \tag{1.1}$$

$$\begin{cases} 2x + y = 0 \\ 4x + 2y = 1 \end{cases} \tag{1.2}$$

$$\begin{cases} 8x - 7y + 6z = 59 \\ x + 2z = 9 \\ 3x + 2y - z = 11 \end{cases} \tag{1.3}$$

Note that (1.1) and (1.3) are uniquely solvable, but (1.2) does not have solution.

You've probably also seen examples of ones with infinite solutions (because for example, you have more variables than "effective" equations).

You can analyze them by hand, but is there a systematic approach to analyzing them? Say, you are telling a computer to solve these. What are your options?

Definition 1.1 (Matrix)**.** For $m, n \geq 1$, an $\underline{m \times n \text{ matrix}}$ is a rectangular array with m rows and n columns with entries from $\mathbb{R}$ (or $\mathbb{C}$ or other fields).

Remark 1.1. **Always count the number of rows, then number of columns**.

Remark 1.2. Notation: For matrix $A \in \mathbb{R}^{m \times n}$, $A_{i,j}$ represents the entry at row i and column j.

Example 1.1 (Matrices)**.** Here are examples of 3×2 matrices:

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

$$\begin{pmatrix} 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}$$

Definition 1.2 (Vector)**.** An $n \times 1$ matrix is called a **column vector**. A $1 \times n$ matrix is called a **row vector**.

Example 1.2 (Vector). Here are examples of a column vector and a row vector respectively:

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

Remark 1.3. **I highly recommend using variables for denoting column vectors by default**. For example $v = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ If you get a row vector, always denote it with a transpose: $v^T = \begin{pmatrix} 1 & 2 \end{pmatrix}$ This is because by convention, we will use left multiplication by a matrix far more often then right multiplication.

Definition 1.3 (Square Matrices). A matrix with the same number of rows and columns is called a **square matrix**.

1.1 Matrix Multiplication and Transpose

Addition and subtractions: $A + B$ are defined naturally by their entry-wise operation. (Note that A and B have to have the same dimensions!) Multiplication of matrices A and B are defined for $A \in \mathbb{R}^{m \times n}$, $B \in \mathbb{R}^{n \times \ell}$ as the following

Definition 1.4 (Matrix Multiplication).

$$(AB)_{i,j} = \sum_{k=1}^{n} (A_{i,k})(B_{k,j})$$

Note that the resulting AB has dimension $m \times \ell$.

Example 1.3 (Matrix Multiplication). Consider the following two matrices:

$$A = \underbrace{\begin{pmatrix} 3 & 1 & 2 \\ 4 & 5 & -1 \end{pmatrix}}_{2\times 3}$$
$$B = \underbrace{\begin{pmatrix} 10 \\ 15 \\ -5 \end{pmatrix}}_{3\times 1}$$

Then:

$$AB = \underbrace{\begin{pmatrix} 3 & 1 & 2 \\ 4 & 5 & -1 \end{pmatrix}}_{2\times 3} \underbrace{\begin{pmatrix} 10 \\ 15 \\ -5 \end{pmatrix}}_{3\times 1} = \underbrace{\begin{pmatrix} 3\times 10 + 1\times 15 + 2\times(-5) \\ 4\times 10 + 5\times 15 + (-1)\times(-5) \end{pmatrix}}_{2\times 1}$$

Remark 1.4. One way to remember if matrix multiplication is well-defined is to note that: $m\times \underline{n}$ and $\underline{n} \times\ell$ results in $m \times \ell$.

Remark 1.5. The way matrix multiplication is defined may not be intuitive. However, it is a natural way to capture *linear transforms.*

Exercise 1.1 (Associativity of Matrix Multiplication)**.** Show that for $A \in \mathbb{R}^{m\times n}$, $B \in \mathbb{R}^{n\times\ell}$, and $C \in \mathbb{R}^{\ell\times p}$, then $(AB)C = A(BC)$. Is it true that $AB = BA$ in general?

Exercise 1.2 (Diagonal matrices form a ring)**.** Show that for **diagonal matrices** $A \in \mathbb{R}^{n\times n}$ and $B \in \mathbb{R}^{n\times n}$ such that $A_{i,j} = B_{i,j} = 0$ for $i \neq j$, AB should also be a diagonal matrix.

Exercise 1.3 (Triangular matrices form a ring)**.** Show that for **upper triangular** $A \in \mathbb{R}^{n\times n}$ and $B \in \mathbb{R}^{n\times n}$ such that $A_{i,j} = B_{i,j} = 0$ for $i > j$, AB should also be an upper triangular matrix.

Exercise 1.4. How many diagonal $A \in \mathbb{R}^{n\times n}$ are there such that

$$A^2 = I$$

where I is the **identity matrix**, a diagonal matrix with only 1s in the diagonal entry.

Definition 1.5 (Transpose)**.** For $A \in \mathbb{R}^{m\times n}$, **transpose** $A^T \in \mathbb{R}^{n\times m}$ is defined as

$$\left(A^T\right)_{i,j} = A_{j,i}$$

Example 1.4 (Transpose)**.**

$$\begin{pmatrix} 3 & 0 & 1 \\ 4 & 2 & 5 \end{pmatrix}^T = \begin{pmatrix} 3 & 4 \\ 0 & 2 \\ 1 & 5 \end{pmatrix}$$

Remark 1.6.

- Taking transpose of a row vector turns it into a column vector.

- Taking a transpose of a column vector turns it into a row vector.
- For any $A \in \mathbb{R}^{m\times n}$, $\left(A^T\right)^T = A$

Exercise 1.5 (Symmetric Triangular Matrix is Diagonal)**.** Show that if an upper triangular matrix $A \in \mathbb{R}^{n\times n}$ satisfies $A^T = A$, then it must be a diagonal matrix.

Exercise 1.6 (Antisymmetric Matrix Does Not Have Diagonal Entries)**.** Show that if $A \in \mathbb{R}^{n\times n}$ satisfies $A^T = -A$, then $A_{i,i} = 0$ for all $i = 1, \cdots n$.

Exercise 1.7 (Transpose of Product of Matrix)**.** Show that for any $A \in \mathbb{R}^{m\times n}$ and $B \in \mathbb{R}^{n\times \ell}$,

$$(AB)^T = B^T A^T$$

Using this result, show that product of two symmetric matrices ($A^T = A$) is also a symmetric matrix.

Remark 1.7. For two column vectors v_1 and v_2, the dot product can be written as $v_1^T v_2$. This becomes a useful notation for dot product, in factf [1]

2 System of Simultaneous Linear Equations to Matrix Equations

We've seen a system like (1.1), (1.2), and (1.3).

We can turn the three examples of linear system of equations to the following **matrix equations** of the form $\underbrace{A}_{\text{Matrix}} \underbrace{x}_{\text{Column Vector}} = \underbrace{b}_{\text{Column Vector}}$:

$$\begin{pmatrix} 3 & 5 \\ 4 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -1 \\ 10 \end{pmatrix} \tag{2.1}$$

$$\begin{pmatrix} 2 & 1 \\ 4 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{2.2}$$

$$\begin{pmatrix} 8 & -7 & 6 \\ 1 & 0 & 2 \\ 3 & 2 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 59 \\ 9 \\ 11 \end{pmatrix} \tag{2.3}$$

[1] Technically 1×1 matrix, but there is no problem in taking that as just a scalar.

One could do even better and skip writing the name of the variables:

$$\left(\begin{array}{cc|c} 3 & 5 & -1 \\ 4 & -1 & 10 \end{array}\right) \tag{2.4}$$

$$\left(\begin{array}{cc|c} 2 & 1 & 0 \\ 4 & 2 & 1 \end{array}\right) \tag{2.5}$$

$$\left(\begin{array}{ccc|c} 8 & -7 & 6 & 59 \\ 1 & 0 & 2 & 9 \\ 3 & 2 & -1 & 11 \end{array}\right) \tag{2.6}$$

(the bar is decorative to show that the system is augmented.)

Our task is to solve it, that is, we wish to put $(A|b)$ into the form $\left(I\middle|\tilde{b}\right)$. This is unfortunately not possible to do[2].

For example, the following system is over-determined (more equations[3] than variables):

$$\begin{cases} 3x + y = -4 \\ 2x - 5y = 3 \\ x + 2y = 8 \end{cases}$$

and the following is under-determined (less equations than variables):

$$\begin{cases} x + 2y = 5 \end{cases}$$

Also (1.2) is an inconsistent system, as the two equations violate each other. As you could see, there are many ways things can go wrong.

A systematic approach to make analysis easier by performing "row operations" is the **Gaussian elimination**.

3 EROs and Gaussian Elimination

Definition 3.1 (Elementary Row Operations (ERO))**.** For a matrix $A \in \mathbb{R}^{m\times n}$, following are the three elementary row operations.

- For $1 \le r < s \le m$, $R_s \leftrightarrow R_r$ (Exchange row r and row s)
- For $1 \le r \le m$ and $\lambda \neq 0$, $R_r \to \lambda R_r$ (Multiply row r by λ)
- For $1 \le r, s \le m$ where $r \neq s$ and $\lambda \in \mathbb{F}$, $R_s \to R_s + \lambda R_r$ (Add to row s the row r multiplied by λ)

[2] A colloquial saying is that with n variables to solve, you need n independent equations.
[3] Conditions

Remark 3.1. It is **paramount** that each operation is reversible.

Remark 3.2. It is **recommended** that you do not compose bunch of EROs as one ERO, as tempting as it is. Using pure EROs is useful for determinant analysis later.

Now we need the notion of "simplified" matrix under these operations.

Definition 3.2 (Echelon Form)**.** $M\mathbb{R}^{m\times n}$ is in **echelon form** if

- If row r of M has any nonzero entries, the first of these is 1.
- If $1 \le r < s \le m$, and row r and row s contain nonzero entries ($M_{r,j}$ and $M_{s,k}$ respectively), then $j < k$.
  - "Leading entry of lower rows should come to the right of the ones on the higher rows."
- If row r of M contains nonzero entries and row s does not, then $r < s$
  - "Zero rows are below all nonzero rows."

Example 3.1 (Echelon Form)**.** The following matrix in echelon form:

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

3.1 Gaussian Elimination

Now we need an algorithm to perform Gaussian elimination to reduce matrix. The idea is to take a matrix A, and eliminate the leading order entries one by one to echelon form.

Example 3.2 (Gaussian Elimination)**.**

$$\begin{aligned}\left(\begin{array}{cc|c} 3 & 5 & -1 \\ 4 & -1 & 10 \end{array}\right) &\xrightarrow[R_2\to\frac{1}{4}R_2]{R_1\to\frac{1}{3}R_1} \left(\begin{array}{cc|c} 1 & \frac{5}{3} & -\frac{1}{3} \\ 1 & -\frac{1}{4} & \frac{5}{2} \end{array}\right) \\ &\xrightarrow{R_2\to R_2-R1} \left(\begin{array}{cc|c} 1 & \frac{5}{3} & -\frac{1}{3} \\ 0 & -\frac{23}{12} & \frac{17}{6} \end{array}\right) \\ &\xrightarrow{R_2\to-\frac{12}{23}R_2} \underbrace{\left(\begin{array}{cc|c} 1 & \frac{5}{3} & -\frac{1}{3} \\ 0 & 1 & -\frac{34}{23} \end{array}\right)}_{\text{echelon}}\end{aligned}$$

Remark 3.3. If there was a zero in the leading entry of the first column, one could swap the rows.

Example 3.3 (Gaussian Elimination (Cont'd))**.** We could take it even further and get **reduced row echelon (RRE) form**.

$$\left(\begin{array}{cc|c} 1 & \frac{5}{3} & -\frac{1}{3} \\ 0 & 1 & -\frac{34}{23} \end{array}\right) \xrightarrow{R_1 \to R_1 - \frac{5}{3}R_2} \underbrace{\left(\begin{array}{cc|c} 1 & 0 & -\frac{49}{23} \\ 0 & 1 & \frac{34}{23} \end{array}\right)}_{\text{RRE}}$$

So $x = -\frac{49}{23}$ and $y = \frac{34}{23}$ for (1.1).

Exercise 3.1. Given $A \in \mathbb{R}^{n\times n}$, show that it can be reduced to RRE within n^2 EROs.

3.2 Inverse Matrix

Definition 3.3 (Inverse Matrix)**.** Note that if it is possible to identify $B \in \mathbb{R}^{n\times n}$ such that $AB = BA = I$, then we call this B the **inverse matrix** of A, and write $A^{-1} := B$. We also say that A is invertible in this case.

Remark 3.4. For a well-posed matrix equation $Ax = b$ where $A \in \mathbb{R}^{n\times n}$ being an invertible matrix, one could write the solution as $x = A^{-1}b$.

Exercise 3.2 (Inverse is Unique)**.** If $A \in \mathbb{R}^{n\times n}$ is invertible, show that the inverse A^{-1} is well-defined.

Exercise 3.3 (Inverse of Product)**.** If $A, B \in \mathbb{R}^{n\times n}$ are both invertible matrices, then show that

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

Exercise 3.4 (Inverse Transpose)**.** If $A \in \mathbb{R}^{n\times n}$ is invertible, then show that A^T is also invertible, and that:

$$\left(A^T\right)^{-1} = \left(A^{-1}\right)^T$$

Remark 3.5. The above exercise shows that one could write A^{-T} as a shorthand for inverse transpose of A.

It is possible to compute the inverse matrix (should it exist) via Gaussian elimination. Consider $A \in \mathbb{R}^{n\times n}$, an invertible matrix. Perform Gaussian elimination on $(A|I) \in \mathbb{R}^{n\times 2n}$, then we should end up with: $(I|A^{-1})$.

Remark 3.6. For the proof, refer to the actual lecture note. Try showing that if $A \in \mathbb{R}^{n\times n}$ is invertible, then RRE is the identity matrix.

Remark 3.7. In practice, computing the inverse matrix is in fact not a feasible thing to do (considering data sets are enormous, and we are often only interested in solution to one single solution to a matrix equation).

3.3 Special Square Matrices

Here is a list of some notable matrices, some of which are already mentioned:

- $I \in \mathbb{R}^{n\times n}$ such that $I_{ij} = \delta_{i,j}$ is an identity matrix.
- $T \in \mathbb{R}^{n\times n}$ such that $T_{i,j} = 0$ for $i > j$ is an upper triangular matrix.
- $T \in \mathbb{R}^{n\times n}$ such that $T_{i,j} = 0$ for $i < j$ is a lower triangular matrix.
- $Q \in \mathbb{R}^{n\times n}$ such that $Q^TQ = I = QQ^T$ is an orthogonal matrix.
- $A \in \mathbb{R}^{n\times n}$ such that $A^T = A$ is a symmetric matrix.
- $B \in \mathbb{R}^{n\times n}$ such that $B^T = -B$ is an antisymmetric matrix.
- $U \in \mathbb{C}^{n\times n}$ such that $U^*U = I = UU^*$ is a unitary matrix[4].

4 Vector Spaces

Definition 4.1 (Vector Space)**.** Let $\mathbb{F}$ be a field. A **vector space** over $\mathbb{F}$ is a non-empty set V with addition and scalar multiplication operation defined, satisfying the vector space axioms:

- Addition is commutative.
  - $u + v = v + u$.
- Addition is associative.
  - $(u + v) + w = u + (v + w)$
- There exists an additive identity.
  - $0 \in V$
- There exists additive inverse for every element.
  - $\forall v \in V : \exists(-v) \in V$
- Scalar multiplication is distributive over vector addition.
  - $\lambda (u + v) = \lambda u + \lambda v$
- Scalar multiplication is distributive over scalar multiplication.

[4] The star denotes conjugate transpose

  - $(\lambda + \mu) v = \lambda v + \mu v$

- Scalar multiplication interacts well with field multiplication.
  - $(\lambda\mu) v = \lambda (\mu v)$

- Identity for scalar multiplication.
  - $1v = v$

Example 4.1 (Vector Space: $\mathbb{R}^n$)**.** $\mathbb{R}^n$ is a vector space. Honestly, this is the most important vector space, so I suggest you get used to working in here. Note that often you can restrict your attention to $\mathbb{R}^3$, and most of your intuition also applies in higher dimensions.

Example 4.2 (Other Vector Spaces)**.** Other vector spaces include:

- $\mathbb{C}$ is a vector space over field $\mathbb{R}$ (essentially $\mathbb{C}$ is the same thing as $\mathbb{R}^2$).
- $\mathcal{P}_n$, the space of polynomials of degree n or less is a vector space over $\mathbb{R}$.

Exercise 4.1 (Uniqueness of Additive Identity and Inverse)**.** For vector space V,

- show that the additive identity element is unique.
- show that for any $v \in V$, the additive inverse $(-v)$ is well-defined.

Definition 4.2 (Subspaces)**.** Let V be a vector space over $\mathbb{F}$. $U \subset V$ is a **subspace** of V if it is nonempty, closed under addition and scalar multiplication.

Exercise 4.2 (Subspace Test)**.** Let V be a vector space over $\mathbb{F}$. Let U be a subset of V. Show that U is a subspace of V if and only if $0 \in U$ and $\lambda u_1 + u_2 \in U$ for all $u_1, u_2 \in U$ and $\lambda \in \mathbb{F}$.

Remark 4.1. If you think about subspace test for a long time, it kinda becomes obvious; the fact that $\lambda u_1 + u_2 \in U$ captures the idea of closure under both addition and scalar multiplication.

Exercise 4.3 (Subspace is also a Vector Space)**.** If V is a vector space and U is a subspace of V, show that

- U is a vector space.
- If W is a subspace of U, then W is a subspace of V.

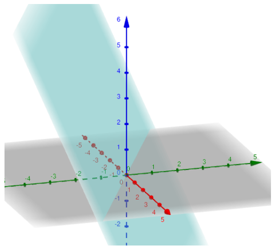


Figure 4.1: $x + 2y + z = 0$

Example 4.3 (Hyperplane is a Subspace of $\mathbb{R}^n$)**.** $U := \{(x, y, z) \in \mathbb{R}^3 | x + 2y + z = 0\} \leq \mathbb{R}^3$.

Definition 4.3 (Sum Space)**.** For $A, B \subset V$ where V is a vector space, define $A + B := \{a + b | a \in A, b \in B\}$

Exercise 4.4 (Sum Space and Intersection are Subspaces)**.** For V vector space, and $U, W \leq V$, show that

- $U + W$ is a subspace of V.
- $U \cap W$ is a subspace of V.

Construct an example where $U \cup W$ is not a subspace.

5 Bases

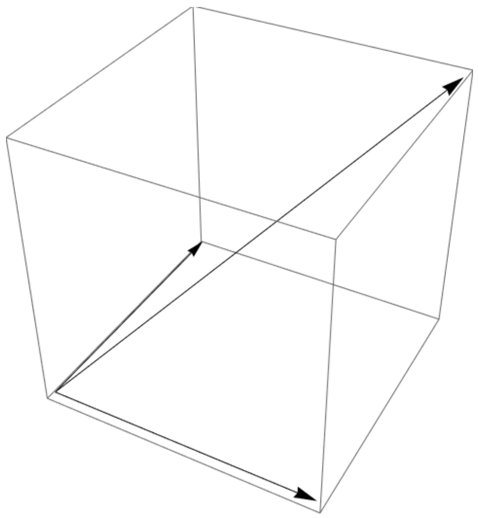
Definition 5.1 (Span)**.** **Span** of vectors $v_1, v_2, \cdots, v_k$ is defined as $\operatorname{span}(\{v_1, \cdots, v_k\}) := \left\{ \sum_{i=1}^k a_i v_i \middle| a_i \in \mathbb{F} \right\}$.

Example 5.1 (Span)**.** Span of $\{e_1, e_2\}$ is $\mathbb{R}^2$, where e_i are i^{th} canonical vector. Span of $\left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\}$ is also $\mathbb{R}^2$.

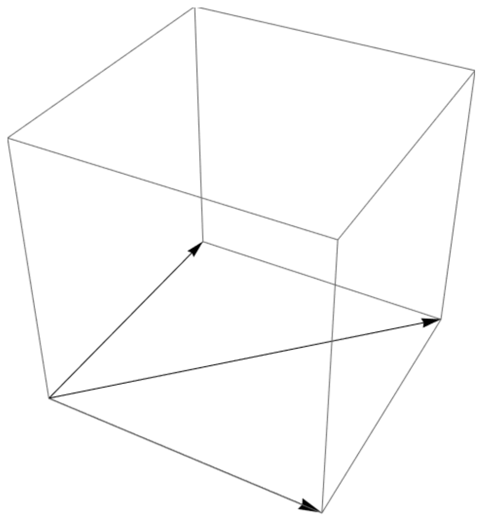
Exercise 5.1 (Span is Subspace)**.** Show that span of finite number of vectors is a subspace.

Definition 5.2 (Linear Combination)**.** u is a **linear combination** of $v_1, \cdots, v_k$ if $u \in \operatorname{span}(\{v_1, \cdots, v_k\})$

Remark 5.1. Note that linear combination only refers to sum of **finitely many** elements.



(a) These three vectors are linearly independent, and they span $\mathbb{R}^3$.



(b) Because e_1 and e_2 can be linearly combined to produce $e_1 + e_2$, linearly dependent.They span $\left\{(x, y, 0)^T \middle| x, y \in \mathbb{R}\right\}$ in this case.

Figure 5.1: Linear Independence & Linear Dependence

Definition 5.3 (Linear Independence)**.** Let V be vector space over $\mathbb{F}$. $v_1, \cdots, v_m \in V$ are **linearly independent** if $\sum_{i=1}^{m} \alpha_i v_i = 0$ implies $\alpha_1 = \cdots = \alpha_m = 0$. Otherwise **linearly dependent**.

Remark 5.2. Might be jumping the gun a bit, but intuitionistically, if there are the minimal number of vectors needed to describe their span, they are linearly independent. (See Figure 5.1)

Exercise 5.2 (Additional "Nonparallel" Vector Preserves Linear Independence)**.** Suppose $v_1, \cdots, v_m$ are linearly independent in vector space V over $\mathbb{F}$. Let $v_{m+1} \in V$ be such that $v_{m+1} \notin \operatorname{span}(\{v_1, \cdots, v_m\})$. Show that $v_1, \cdots, v_m, v_{m+1}$ are also linearly independent.

Remark 5.3. General strategy to showing linear independence is to assume $\sum_{i=1}^{m} \alpha_i v_i = 0$, and show that $\alpha_1 = \cdots = \alpha_m = 0$

Definition 5.4 (Bases)**.** Let V be a vector space. A basis of V is a set of linearly independent spanning set.

Remark 5.4 (Basis is Nonunique). For $\mathbb{R}^3$, $\{e_1, e_2, e_3\}$ is a basis, but $\{e_1, e_2, e_1 + e_2 + e_3\}$ is also a basis.

Exercise 5.3 (Basis = Any Vector Has Unique Representation)**.** Show that $\{v_1, \cdots, v_n\} \subset V$, then $\{v_1, \cdots, v_n\}$ is a basis of V if and only if for any $v \in V$, there exists unique linear combination representation $v = \sum_{i=1}^n a_i v_i$

Example 5.2. Suppose $v = (1,2,3)^T \in \mathbb{R}^3$. In $\{e_1, e_2, e_3\}$ basis, $v = 1e_1 + 2e_2 + 3e_3$. In $\left\{ \underbrace{e_1 + 2e_3}_{v_1}, \underbrace{e_2}_{v_2}, \underbrace{e_3}_{v_3} \right\}$ basis, $v = 1v_1 + 0v_2 + v_3$.

Remark 5.5 (Coordinate). Think of the coefficients (α_i) in the linear combination expansion $v = \sum_{i=1}^n \alpha_i v_i$ as coordinates.

Exercise 5.4 (Finite Spanning Set Implies Existence of Basis)**.** Show that if V has finite spanning set S, then S contains a linearly independent spanning set. (Hint: Let $T \subset S$ be the largest linearly independent set. Then show that T spans V)

Exercise 5.5 (Steinitz Exchange Lemma)**.** Let V be vector space over $\mathbb{F}$. Take $X \subset V$. Suppose $u \in \operatorname{span}(X)$, but $u \notin \operatorname{span}(X \setminus \{v\})$ for some $V \in X$. Let $Y = (X \setminus \{v\}) \cup \{u\}$ ("exchange u for v"). Then $\operatorname{span}(Y) = \operatorname{span}(X)$.

Example 5.3. While mouthful to state, it is in fact an obvious statement. Consider $\mathbb{R}^2$. Let $X = \{e_1, e_2\}$. Take $u = e_1 + e_2 \notin \operatorname{span}\left(X \setminus \left\{\underbrace{e_2}_{v}\right\}\right)$. Then $Y = (X \setminus \{v\}) \cap \{u\} = \{e_1, e_1 + e_2\}$, which has the same span as X.

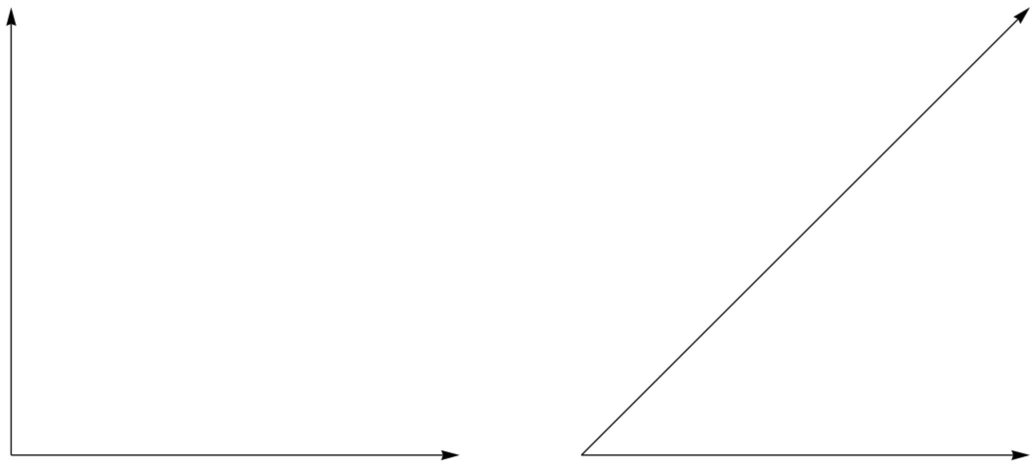
Steinitz exchange lemma gives justification as to well-definededness of the notion of dimension.

Definition 5.5 (Dimension)**.** For finite dimensional vector space V, **dimension** of V, denoted as $\dim V$ is the size of any basis of V.

Exercise 5.6. Show that

- $\mathbb{R}^n$ has dimension n.
- space of polynomial of degree upto n has dimension $n+1$.
- space of matrices $\mathbb{R}^{m \times n}$ has dimension mn.

Definition 5.6 (Row Space and Rank)**.** Given $A \in \mathbb{R}^{m \times n}$, the span of the rows of A is **row space**, and **row rank** is defined as the dimension of the row space.



(a) X (b) Y

Figure 5.2: Steinitz exchange lemma says you can swap a vector in a basis so long as it preserves the span.

Exercise 5.7. Show that the row rank of the following matrix is 3:

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

(Hint: EROs do not change the row space by Steinitz exchange lemma!)

Exercise 5.8. Show that if $U \leq V$ where V is finite dimensional, then

- $\dim U \leq \dim V$
- If $\dim U = \dim V$, then $U = V$.

Exercise 5.9 (Extension of Subspace Basis to Vector Space Basis)**.** Show that if V is finite dimensional and $U \leq V$, then any basis of U can be extended to the basis of V.

Show by an example that $U \leq V$ does not imply that basis of V includes a basis of U.

Exercise 5.10 (Dimension Formula)**.** Show that $U, W \leq V$ (V finite dimensional vector space), then $\dim(U + W) + \dim(U \cap W) = \dim U + \dim W$

Example 5.4 (Dimension Formula)**.** Consider $U = \text{span}(\{e_1\})$ and $W = \text{span}(\{e_1, e_1 + e_2\})$. Then $U + W = \mathbb{R}^2$, so $\dim(U + W) = 2$, and $U \cap V = U$, so $\dim(U \cap V) = 1$. This satisfies the dimension formula.

Exercise 5.11. Using dimension formula or otherwise[5], show that given $V_1, V_2 \leq V$, $\dim V_1 + \dim V_2 > \dim V$, $\dim(V_1 \cap V_2) \geq 1$.

Definition 5.7 (Diricet Sum)**.** If $U, W \leq V$ where V is finite dimensional vector space, and $U + W = V$, then V is a **direct sum** of U and W, and we write $V = U \oplus W$. (See Figure 5.3)

Example 5.5 (Direct Sum)**.** Let $V = \mathbb{R}^3$. Take $U = \text{span}\left(\left\{(1,1,0)^T\right\}\right)$, and $W = \text{span}\left(\left\{(1,1,1)^T, (0,1,0)^T\right\}\right)$. Then $V = U \oplus W$.

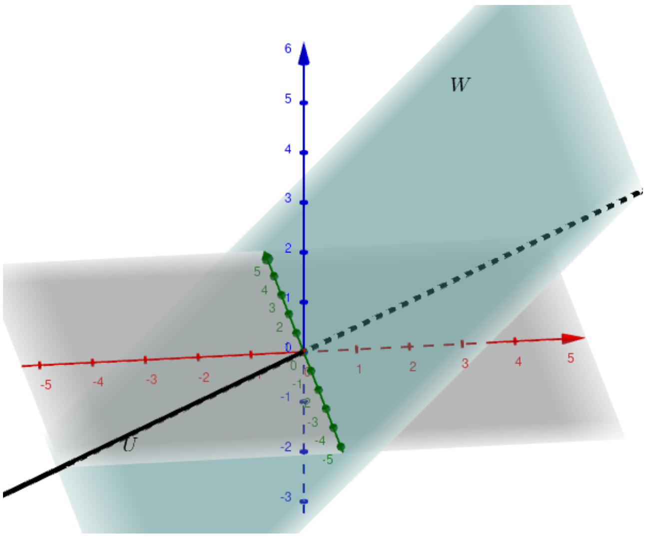


Figure 5.3: U and W form direct sum of V as the only intersection is the origin.

Exercise 5.12 (Characterization of Direct Sum)**.** Show that the following are equivalent:

- $V = U \oplus W$
- Every $v \in V$ has unique expression $v = u + w$ where $u \in U$ and $w \in W$.

[5] Another way to argue is to consider the bases of the two subspaces, put them in a matrix, and note that the matrix is "fat", meaning it must have free variable in a matrix equation.

- $\dim V = \dim U + \dim W$ and $V = U + W$
- $\dim V = \dim U + \dim W$ and $U \cap W = \{0_V\}$
- Given bases $\mathcal{B}_U$ of U and $\mathcal{B}_W$ of W, $\mathcal{B}_U \cap \mathcal{B}_W$ is a basis of V.

Exercise 5.13. Find an example that shows $V = U \oplus W$ does not imply every basis of V is a union of basis of U and a basis of W.

6 Linear Transformation

Linear transformation is the reason why we define matrix multiplication in such a funky way.

Definition 6.1 (Linear Transformation)**.** Let V, W be vector spaces over $\mathbb{F}$. Map $T : V \to W$ is a **linear transformation** if

- $T(v_1 + v_2) = T(v_1) + T(v_2)$.
- $T(\lambda v) = \lambda T(v)$.

Remark 6.1. It is as if T can be replaced by a matrix…Hm…

Example 6.1.

- Identity map $\mathrm{id}_V : V \to V$ defined by $\mathrm{id}_V(v) = v$ is a linear map.
- Zero map $z : V \to W$ defined by $z(v) = 0$ is a linear map.
- Given a matrix $A \in \mathbb{R}^{m\times n}$, left multiplication map $L : \mathbb{R}^n \to \mathbb{R}^m$ defined by $L(v) = Av$ is a linear map.
- Given $V = U \oplus W$, we know we can uniquely write $v = u + w$ where $u \in U$ and $w \in W$. Projection map $P : V \to V$ defined by $P(v) = w$ is a linear map.

Exercise 6.1. Given U, V, W be vector spaces, if $S : U \to V$ and $T : V \to W$ are linear, then $TS \coloneqq T \circ S : U \to W$ is also a linear transformation.

Exercise 6.2. Given U, V be vector spaces, show that linear map $T : U \to V$ is uniquely defined by the image of T for a basis of U, that is, if $\{u_1, \cdots, u_m\}$ is a basis of U, map is uniquely determined provided the outputs of $T(u_i)$.

Remark 6.2 (Matrix Multiplication and Linear Transform). **This was a revelation for me when I learned this.** Consider matrix

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$$

Then notice that $Ae_1 = (1,4,7)^T$, $Ae_2 = (2,5,8)^T$, and $Ae_3 = (3,6,9)^T$. Note that these are columns of the matrix A. Now consider

$$A\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = A(e_1 + 2e_2 + 3e_3) = \begin{pmatrix} 1 \\ 4 \\ 7 \end{pmatrix} + 2\begin{pmatrix} 2 \\ 5 \\ 8 \end{pmatrix} + 3\begin{pmatrix} 3 \\ 6 \\ 9 \end{pmatrix}$$

This and Exercise 6.2 suggests that **any linear map in a finite dimensional vector space can be written as a left matrix multiplication.** This is also why I stick by remark 1.3, unless you know for sure that you will be working with right multiplication.

Example 6.2. To describe a counterclockwise rotation by angle θ in $\mathbb{R}^2$, note that we can map $\begin{pmatrix} 1 \\ 0 \end{pmatrix} \mapsto \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}$, and $\begin{pmatrix} 0 \\ 1 \end{pmatrix} \mapsto \begin{pmatrix} -\sin\theta \\ \cos\theta \end{pmatrix}$. This linear transform can be described by a left multiplication by the orthogonal matrix:

$$Tv := \underbrace{\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}}_{\text{orthogonal}} v$$

Check again that it maps the basis e_1, e_2 to their respective images.

Exercise 6.3. Construct a linear map in $\mathbb{R}^3$ that describes $\frac{2\pi}{3}$ rotation around the axis $(1,1,1)^T$.

Remark 6.3. While it is helpful to get intuition in terms of canonical basis, left multiplication by matrix describes linear transformation in any given basis in their "coordinates". See Example 7.1.

6.1 Rank and Nullity

Definition 6.2 (Kernel and Image)**.** Given V, W to be vector spaces and $T : V \to W$ is linear, **kernel** (null space) of T is

$$\ker T := \{v \in V | T(v) = 0_W\}$$

and **image** of T is

$$\operatorname{Im} T := \{T(v) | v \in V\}$$

Remark 6.4. Kernel is the set of vectors that map to zero. Image is the set of vectors that end up being mapped to from V.

Exercise 6.4. For V, W vector spaces and $T : V \to W$ linear, T is injective if and only if $\ker T = \{0_V\}$.

Exercise 6.5 (Kernel and Image are Subspaces)**.** For V, W vector spaces and $T : V \to W$ linear, show

- $\ker T \leq V$ and $\operatorname{Im} T \leq W$
- if $\{v_1, \cdots, v_m\}$ spans V, then $T(\{v_1, \cdots, v_m\})$ spans $\operatorname{Im} T$

Since the kernel and image are subspaces, we are interested in their dimensions.

Definition 6.3 (Rank and Nullity)**.** V, W being vector spaces with V finite dimensional. $T : V \to W$ linear.

- **Nullity** of T is $\operatorname{null} T \coloneqq \dim(\ker T)$
- **Rank** of T is $\operatorname{rank} T \coloneqq \dim(\operatorname{Im} T)$

Exercise 6.6 (Rank-Nullity Theorem)**.** Given V, W being vector spaces with V finite dimensional, and $T : V \to W$ linear, show that $\dim V = \operatorname{rank} T + \operatorname{null} T$

7 Change of Basis

It was already mentioned at remark 6.2 that matrix can be used to describe a linear map **Note that in the discussion we were implicitly assuming that we were using canonical basis**[6]. Now we want to see the relationship between two matrices describing the same linear tranform, but in different basis.

Example 7.1. Consider linear map $T : \mathbb{R}^2 \to \mathbb{R}^2$ defined as $T(x, y) = \begin{pmatrix} 2x + y \\ 3x - 2y \end{pmatrix}$. Note that

$$T(1,0) = \begin{pmatrix} 2 \\ 3 \end{pmatrix} \qquad T(0,1) = \begin{pmatrix} 1 \\ -2 \end{pmatrix}$$

[6]The use of canonical basis seems like a natural choice, but from the POV of mathematics, it is an arbitrary choice..., like we use canonical basis only because they are easy to compute, most of the times.

So with respect to basis e_1, e_2, the matrix describing this transformation is

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

On the other hand, consider another basis $f_1 = (1, -2)^T, f_2 = (-2, 5)^T$ of $\mathbb{R}^2$. Then

$$T(f_1) = \begin{pmatrix} 0 \\ 7 \end{pmatrix} = 14f_1 + 7f_2 \qquad T(f_2) = \begin{pmatrix} 1 \\ -16 \end{pmatrix} = -27f_1 - 14f_2$$

So with respect to basis f_1, f_2, the matrix describing this transformation is

$$B = \begin{pmatrix} 14 & -27 \\ 7 & -14 \end{pmatrix}$$

These two matrices describe the *same* linear transform, just in different basis. We can actually explicitly construct how they are related.

Idea: Change basis from one to the other, perform the transform in the new basis, and turn it back to the original basis. Because of the relation of the two bases:

$$\begin{aligned} f_1 &= e_1 - 2e_2 \\ f_2 &= -2e_1 + 5e_2 \end{aligned}$$

Construct the change of basis matrix P, "used for mapping from f_i coordinates to e_i coordinates:

$$P = \begin{pmatrix} 1 & -2 \\ -2 & 5 \end{pmatrix}$$

The inverse matrix describes the mapping from e_i coordinates to f_i coordinates.

$$P^{-1} = \begin{pmatrix} 5 & 2 \\ 2 & 1 \end{pmatrix}$$

Then it must be that $P^{-1}AP$ describes the linear transformation in f_i coordinates[7], hence $P^{-1}AP = B$. (Indeed you can verify this.)

Remark 7.1. Explicit construction of $A = P^{-1}BP$ is not really the highlight of the course, but it is worth understanding the process.

Definition 7.1 (Similarity)**.** If there exists invertible $P \in \mathbb{R}^{n\times n}$ such that $A, B \in \mathbb{R}^{n\times n}$, then we say A and B are **similar**.

Exercise 7.1 (Similarity is an Equivalence Relation)**.** Verify that similarity is an equivalence relation by checking the three conditions of equivalence relation.

[7] In more detail, P first maps from f_i coordinates to e_i coordinates. Then A does the actual linear map in e_i coordinates. Finally P^{-1} maps the coordinates in e_i to ones in f_i.

8 Matrix and Rank

We've seen that matrix and linear transforms are very much related. In fact, we might as well start abusing the notion of "rank" of a matrix! We do need to check if the definition of rank for a matrix is well-defineded and natural[8]

To do so, it helps to show one result:

Exercise 8.1. For $A \in \mathbb{R}^{m\times n}$, there exists invertible $P \in \mathbb{R}^{n\times n}$ and $Q \in \mathbb{R}^{m\times m}$ such that $Q^{-1}AP$ can be written as:

$$Q^{-1}AP = \begin{pmatrix} I_r & 0_{r\times s} \\ 0_{t\times r} & 0_{t\times s} \end{pmatrix} \tag{8.1}$$

(Hint: The simplest argument is to note that elementary row operations are simply left multiplication by invertible matrices. Similarly one could consider elementary column operations, which must be right multiplication by invertible matrices. Then we consider the RRE/RCE[9] form.)

Using this, one could argue now the following result:

Exercise 8.2 (Column Rank is Equal to Row Rank)**.** Deduce from (8.1) that the column rank and the row rank of $A \times \mathbb{R}^{m\times n}$ are the same. (Hint: You can read off r to be both the column rank and the row rank.)

So finally we have rank of a matrix.

Definition 8.1 (Rank of a Matrix)**. Rank of matrix** $A \in \mathbb{R}^{m\times n}$, denoted by $\text{rank}(A)$, is the column rank (or row rank) of A.

Remark 8.1. Thinking of rank of a matrix as its column rank should reminisce you of the fact that matrices describe linear maps. (Since the image of the map by left multiplication by the matrix is the column space.)

9 Bilinear Forms and Inner Product Space

Definition 9.1 (Bilinear Form)**.** Given a vector space V over $\mathbb{F}$, a **bilinear form** on V is a function $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{F}$ such that

- $\langle \alpha_1 v_1 + \alpha_2 v_2, v_3 \rangle = \alpha_1 \langle v_1, v_3 \rangle + \alpha_2 \langle v_2, v_3 \rangle$
- $\langle v_1, \alpha_2 v_2 + \alpha_3 v_3 \rangle = \alpha_2 \langle v_1, v_2 \rangle + \alpha_3 \langle v_1, v_3 \rangle$

[8] Intuition may say well-defined, since change of basis does not change the actual span of them. . .

[9] Reduce column echelon

for all $v_1, v_2, v_3 \in V$ and $\alpha_1, \alpha_2, \alpha_3 \in \mathbb{F}$

Remark 9.1. Bilinear form is a generalization of dot product in a way. In fact, $\langle x, y \rangle = x^T y$ (usual dot product) for $x, y \in \mathbb{R}^n$ over $\mathbb{R}$ is indeed a bilinear form.

Example 9.1. Given $A \in \mathbb{F}^{n\times n}$, $\langle x, y \rangle := x^T A y$ is a bilinear form over $\mathbb{F}$.

Definition 9.2 (Inner Product Space)**.** Bilinear form $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{R}$ is

- **symmetric** if $\langle v_1, v_2 \rangle = \langle v_2, v_1 \rangle$ for all $v_1, v_2 \in V$.
- **positive definite** if $\langle v, v \rangle \geq 0$ for all $v \in V$ where equality if and only if $v = 0$.

Inner product space is a positive definite symmetric bilinear form on V.

Remark 9.2. Dot product is still an inner product over $\mathbb{R}^n$.

Definition 9.3. For real inner product space V, define **norm** of $v \in V$ is

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

Exercise 9.1. Consider $V = \mathbb{R}_n[x]$, a space of polynomials of degree n with real coefficients. Show that $\langle f, g \rangle := \int_{-1}^{1} f(x)g(x)\,\mathrm{d}x$ is an inner product over V.

Often we are interested in orthonormal set, as it makes the computation much easier.

Definition 9.4 (Orthonormal Set)**.** Given V an inner product space, $\{v_1, \cdots, v_n\} \subset V$ is an **orthonormal set** if for all i, j,

$$\langle v_i, v_j \rangle = \delta_{i,j}$$

Exercise 9.2. $\{v_1, \cdots, v_n\}$ be an orthonormal set in an inner product space V. Then show that $v_1, \cdots, v_n$ are linear independent.

Exercise 9.3. Show that canonical basis of $\mathbb{R}^n$ is an orthonormal set.

Exercise 9.4 (Legendre Polynomials)**.** Construct an orthonormal set of basis for the inner product space for $\mathbb{R}_2[x]$.

Exercise 9.5 (Orthogonality in Fourier Series)**.** For what $k > 0$ is the set $\{\cos(nx) | n = 1, 2, \cdots\}$ an orthonormal set under inner product $\langle f, g \rangle := \int_{-\pi}^{\pi} f(x)g(x)\,\mathrm{d}x$?

9.1 Orthogonal Matrices

Definition 9.5 (Orthogonal Matrix)**.** $Q \in \mathbb{R}^{n \times n}$ is an **orthogonal matrix** if $Q^T Q = QQ^T = I$. In other words, $Q^T = Q^{-1}$.

Exercise 9.6 (Characterization of Orthogonal Matrices)**.** Show that orthogonal matrices can be characterized in any of the following ways:

- $QQ^T = I$
- $Q^T Q = I$
- Rows of Q form an orthonormal basis of $\mathbb{R}^n$.
- Columns of Q form an orthonormal basis of $\mathbb{R}^n$.
- $(Qx)^T (Qy) = x^T y$ for all $x, y \in \mathbb{R}^n$

Exercise 9.7. Deduce that the Euclidean norms of each row and column are 1.

Exercise 9.8. If $U, V \in \mathbb{R}^{n \times n}$ are both orthogonal matrices, show that UV is also an orthogonal matrix.