Linear Algebra I & II Summary

Paul Kim

June 10, 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, <u>there are many ways things can go wrong</u>.

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 $1 \leq r < s \leq m$, $R_s \leftrightarrow R_r$ (Exchange row r and row s)
- For $1 \leq r \leq m$ and $\lambda \neq 0$, $R_r \to \lambda R_r$ (Multiply row r by λ)
- For $1 \leq r, s \leq m$ 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).

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

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

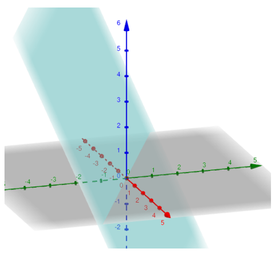

Figure 4.1: $x + 2y + z = 0$