A Second Course in Linear Algebra

Raymond Bian

January 23, 2024

Contents

Lecture 1: Review

1 Vectors and Matrices

For the time being, everything indicated in this course is in $\mathbb{R}$.

Definition 1. A **vector** will be defined as a column vector, e.g.

$$u = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in \mathbb{R}^3.$$

Notation. Sometimes, they will be written as a column vector lying down, e.g. $(x_1, x_2, x_3) \in \mathbb{R}^3$

Definition 2. Let a be a scalar. Then multiplication between vector and scalar is defined as

$$au = \begin{bmatrix} a \cdot x_1 \\ a \cdot x_2 \\ a \cdot x_3 \end{bmatrix}.$$

Definition 3. Let $u = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ and $v = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$. Then addition between vectors is defined as

$$u + v = \begin{bmatrix} x_1 + y_1 \\ x_2 + y_2 \\ x_3 + y_3 \end{bmatrix}.$$

Definition 4. If u, v are vectors and a, b are scalars, then any $au + bv$ is a **linear combination** of u and v.

Remark. A **vector space** V is a set of objects u, v such that $au + bv \in V$.

Example. Polynomials of degree ≤ 2 in one variable can form a vector space.

Explanation. Let $p(x) = a_0 + a_1x + a_2x^2$, and $q(x) = b_0 + b_1x + b_2x^2$. Multiplying by scalars and adding are defined. Note that $p(x) \to \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix}$.

Example. Let $f(x) : [0, 1] \to \mathbb{R}$ be a continuous function. We can multiply such functions by scalars and add together such functions, so they form a vector space as well.

Suppose we have two vectors $u, v \in \mathbb{R}^3$. Looking at the set of all linear combinations of u, v,

- if both u and v are the zero vectoor, then $W = \{0\}$.
- if $u = \lambda v$, $v \neq 0$, then W is the line of all multiples of v.
- if u and v are **linearly independent**, then W is a plane in $\mathbb{R}^3$.

Definition 5. Vectors u_1, u_2, u_3 are **linearly independent** if and only if

$$a_1u_1 + a_2u_2 + a_3u_3 = 0 \Rightarrow a_1 = a_2 = a_3 = 0.$$

Definition 6. Let V, W be a vector spaces such that $W \subseteq V$. Then, W is called a **subspace** of V.

Example. Let $W = \{\begin{bmatrix} x_1 \\ x_2 \\ 0 \end{bmatrix} : x_1, x_2 \in \mathbb{R}\}$. Then, W is a subspace of $\mathbb{R}^3$.

Theorem 1. If $u, v \in V$, then the set of linear combinations of u and v is a subspace.

Proof. Let $W = \text{span}\{u, v\}$. We must show that $w_1, w_2 \in W \Rightarrow c_1 w_1 + c_2 w_2 \in W$. By assumption, $w_1 = a_1 u + b_1 v$, and $w_2 = a_2 u + b_2 v$, such that $w = (c_1 a_1 + c_2 a_2)u + (c_1 b_1 + c_2 b_2)v$. Therefore, w is a linear combination of u, v. $\square$

Example. Let $u = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$, and $v = \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix}$. Then, $\text{span}\{u, v\}$ is a proper subspace of $\mathbb{R}^3$.

Definition 7. $u \cdot v = x_1 y_1 + x_2 y_2 + x_3 y_3$ is the dot product of the vectors $u = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ and $v = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$

Definition 8. We say that $u \perp v$ if $u \cdot v = 0$.

Definition 9. The length or **norm** of a vector u is $\sqrt{u \cdot u} = \|u\|$

Theorem 2. The **Cauchy–Schwarz inequality** states that $|u \cdot v| \leq \|u\|\|v\|$.

Proof.

$$(u + \lambda v) \cdot (u + \lambda v) \geq 0$$
$$u \cdot u + \lambda^2 v \cdot v + 2\lambda u \cdot v \geq 0.$$

The minimum lambda is $\frac{-b}{2a} = \frac{-u \cdot v}{v \cdot v}$, which results in this inequality being true. Therefore, all greater values for lambda will result in this inequality being true. $\square$

Theorem 3. The **triangle inequality theorem** states that $\|u + v\| \leq \|u\| + \|v\|$.

Definition 10. The **unit vector** of a vector u, $\hat{u}$ is given by $\frac{u}{\|u\|}$.

Theorem 4. If u and v are vectors such that $\|u\| = \|v\| = 1$, then $u \cdot v = \cos(\theta)$ where θ is the angle between u and v.

Corollary. If u and v are vectors, then $u \cdot v = \|u\|\|v\| \cos(\theta)$. Note that $u \cdot v = 0$ when $\theta = \frac{\pi}{2}$ or $\frac{3\pi}{2}$.

Lecture 2: Matrices

Example.

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

is a matrix. We can also write $A = \{a_{ij}\}$ such that $i = 1 \dots n$ and $j = 1 \dots m$.

What does it mean to take a product between a matrix and a vector?

Definition 11. This product is defined as

$$\begin{pmatrix} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 \\ a_{21}x_1 + a_{22}x_2 + a_{13}x_3 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 \end{pmatrix}.$$

i.e. a collection of dot products between the rows and x.

We can also see the product as a linear combination of the columns of the matrix A.

Definition 12. Let the columns of A be A_1, A_2, A_3. Then, $Ax = x_1 A_1 + x_2 A_2 + x_3 A_3$.

Notation. A's columns are denoted A_1, A_2, A_3, while A's rows are denoted A^1, A^2, A^3.

If we look at the linear equation $Ax = b$, we can say that b is a linear combination of the columns of A. Instead, looking at it like an equation, "can b be written as a linear combination of the columns of A"?

Looking at $A^1 x = b_1$, there are two free variables, such as this is a plane in $\mathbb{R}^3$. The only time this is not a plane is if a_{11}, a_{12}, a_{13} are all zero, and b_1 is nonzero.

If we have x, y, $A^1 x = 0$ and $A^1 y = 0$ implies $ax + by = z$, which solves $A^1 z = 0$. The set of solutions is a subspace.

Now, suppose we have all solutions of $A^1 x = 0$. Call this V. How do we then write the solutions to $A^1 x = b$? We find any such c such that $A^1 c = b_1$. Then, we claim that the set of solutions of $A^1 x = b_1$ is $V + c = \{x + c | x \in V\}$. Checking our solution, $A^1 \cdot (x + c) = \underbrace{A^1 \cdot x}_{0} + \underbrace{A^1 \cdot c}_{b_1} = b_1$.

Let $W = V + c$. We want to show if $x \in W \Rightarrow A^1 \cdot x = 0$. Assume $A^1 z = b_1$. If we set $x = z - c$, then $A^1 x = A^1 z - A^1 c = 0$. Therefore, $z = x + c \in W$.

All in all, solving all three equations $A^1 x = b_1, A^2 x = b_2, A^3 x = b_3$ is now just finding the intersection of three translated planes. **This is what solving $Ax = b$ means.**

Another viewpoint is this. Consider the equation $A_1x_1 + A_2x_2 + A_3x_3 = b$. Consider the span of A_1, A_2, A_3. Does this span contain b?

Example. Let's say that

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

Solving $Ax = b$, we have $x_3 = b_3$, $x_2 = b_2 + b_3$, and $x_1 = b_1 + b_2 + b_3$ such that

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

Let C denote this matrix. Then, $Ax = b \Leftrightarrow Cb = x$, such that $C = A^{-1}$. Then, C is the **inverse** of A.

Definition 13. We want to say that every $n \times n$ matrix can be written as the product as an upper triangular and lower triangular matrix, called **LU factorization**.

Definition 14. **Matrix multiplication** is defined as $(AB)_{ij} = \sum_k a_{ik} + b_{kl}$ where $A = \{a_{ij}\}$ and $B = \{b_{kl}\}$

The other way to see AB is if $B = \begin{pmatrix} B_1 & B_2 & \dots & B_n \end{pmatrix}$, then $AB = \begin{pmatrix} AB_1 & AB_2 & \dots & AB_n \end{pmatrix}$. In other words, $(AB)_{ij} = A^i \cdot B_j$.

Lecture 3: Matrix Algebra

Example. Solve

$$\underbrace{\begin{bmatrix} 2 & 4 & -2 \\ 4 & 9 & 4 \\ -2 & -3 & 7 \end{bmatrix}}_{A} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \underbrace{\begin{bmatrix} 2 \\ 8 \\ 10 \end{bmatrix}}_{b}.$$

Explanation.

$$x = \begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix}.$$

Let

$$E_{12} = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Then, we have

$$E_{12} \begin{bmatrix} 2 \\ 8 \\ 10 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \\ 10 \end{bmatrix}.$$

Note that this is also $E_{12}(Ax) = E_{12}b = (E_{12}A)x$

Definition 15. AB is such that

$$A(Bx) = (AB)x.$$

for every vector x. It is defined as

$$AB = [AB^1, AB^2, \dots, AB^n].$$

where B^i is the i-th column of B.

Theorem 5. $Ax = b \Rightarrow (CA)x = Cb$

Theorem 6. Let $\mathbb{R}^n$ be a vector space and $A, B : \mathbb{R}^n \to \mathbb{R}^n$ linear mappings. Then,

$$A \circ B : \mathbb{R}^n \to \mathbb{R}^n.$$

is also a linear transformation. Also

$$A \circ B(x) = ABx.$$

Theorem 7. If $\hat{A}$ is a linear map from $\mathbb{R}^n \to \mathbb{R}^n$ then $\hat{A}(x) = Ax$ for a matrix A.

Proof. For a linear map, we have $\hat{A}(x + y) = \hat{A}(x) + \hat{A}(y)$ and $\hat{A}(\alpha x) = \alpha\hat{A}(x)$. We want to show that any linear mapping is a matrix multiplication. Let

$$e_i = \begin{bmatrix} 0 \\ 0 \\ \dots \\ 1 \\ \dots \\ 0 \end{bmatrix}$$

where the 1 is in the ith place. Let $A^i = \hat{A}(e_i)$. Let $A = \begin{bmatrix} A^1 & A^2 & \dots & A^n \end{bmatrix}$. Then, by construction

$$\begin{aligned} \hat{A}(x) &= \hat{A}(x_1e_1 + x_2e_2) + \dots + x_ne_n) \\ &= x_1\hat{A}(e_1) + x_2\hat{A}(e_2) + \dots + x_n\hat{A}e_n) \\ &= x_1A^1 + x_2A^2 + \dots + x_nA^n \\ &= Ax. \end{aligned}$$

$\square$

We can also calculate matrix multiplication as $(AB)_{i,j} = \sum_k A_{i,k} \cdot B_{k,j}$.

Theorem 8. Suppose we take a third matrix C. Then,

$$A(BC) = (AB)C.$$

This is the **associative property**.

Proof. We saw that

$$A(Bx) = (AB)x.$$

Applying this, we have:

$$\begin{aligned}(AB)C &= \begin{bmatrix}(AB)C^1 & \dots & (AB)C^n\end{bmatrix} \\ &= \begin{bmatrix}A(BC^1) & \dots & A(BC^n)\end{bmatrix} \\ &= A\begin{bmatrix}BC^1 & \dots & BC^n\end{bmatrix} \\ &= A(BC).\end{aligned}$$

□

With this information, row reduction is just a series of matrix multiplications. Note that in row reduction, we can also have permutation matrices that switches the rows.

Theorem 9. $AB \neq BA$.

Proof.

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

but not the other way around. □

To summarize matrix operations, we have

$$\begin{aligned}A + B &= B + A \\ \alpha(A + B) &= \alpha A + \alpha B\end{aligned}$$

$$\begin{aligned}(AB)C &= A(BC) \\ (A + B)C &= AC + BC \\ C(A + B) &= CA + CB.\end{aligned}$$

By these properties, space of matrices is a vector space, and an algebra. However, we are missing division (the inverse)!

Note that a mapping $A : \mathbb{R}^n \to \mathbb{R}^m$ such that $m < n$ cannot be invertible, as there are many solutions to $Ax = b$ and therefore cannot be a bijection. The same can be said when $n > m$, because $Ax = b$ will have no solutions. Therefore, A is an invertible if $n = m$.

Definition 16. The **inverse** A^{-1} of A is defined such that

$$A^{-1}Ax = x \quad \forall x.$$

as well as $AA^{-1} = I$ and A^{-1} must be unique.

Theorem 10.

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

Proof.

$$\begin{aligned}(B^{-1}A^{-1})(AB) &= B^{-1}(A^{-1}A)B \\ &= B^{-1}IB \\ &= B^{-1}B \\ &= I.\end{aligned}$$

This is the only inverse. □

Example. The inverse of

$$E_{12} = \begin{pmatrix}1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1\end{pmatrix}.$$

is just

$$E_{12}^{-1} = \begin{pmatrix}1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1\end{pmatrix}.$$

(you add back the two first rows you subtracted from the second).

From the elimination example earlier, we have

$$E_{23}E_{13}E_{12}A = \begin{bmatrix}2 & 4 & -2 \\ 0 & 1 & 7 \\ 0 & 0 & -2\end{bmatrix} = U.$$

which is now upper triangular. Flipping this around, $A = \underbrace{E_{12}^{-1}E_{13}^{-1}E_{23}^{-1}}_{L}U$. Note that all $E_{i,j}$ are lower triangular, such that L is also lower triangular. This is **LU Factorization**.

We can use this to solve $Ax = b$ by first writing $A = LU \Rightarrow Ux = L^{-1}B$, from which you do backwards substitution to solve the problem, reducing the number of operations from a magnitude of n^3 to n^2. However, getting A^{-1} is still n^3, so it should only be precomputed if we solve equations $Ax = b$ n times.

Lecture 4: Transpose, Permutations, Spaces

Definition 17. If A is an $n \times m$ matrix, then the **transpose** A^T is

$$(A^T)_{ij} = A_{ji}.$$

If A is $n \times m$, then A^T is $m \times n$.

Example. If $x = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$, then $x^T = \begin{bmatrix} 1 & 2 & 3 \end{bmatrix}$

Proposition 1. $(AB)^T = B^T A^T$

Proof. How do we compute $(AB)^T$? Assume that B is just a vector x. This means that Ax is just a vector

$$Ax = x_1 A^1 + x_2 A^2 + \ldots + x_n A^n.$$

Subsequently,

$$(Ax)^T = x_1 (A^1)^T + \ldots + x_n (A^n)^T.$$

where $(A^3)^T$ is the transpose of the 3rd column, which is just the 3rd row. In other words,

$$\begin{aligned} (Ax)^T &= x_1 (A^T)_1 + \ldots + x_n (A^T)_n \\ &= x^T A^T. \end{aligned}$$

With this definition, then $(AB)^T = \begin{bmatrix} AB^1 & AB^2 & \ldots & AB^n \end{bmatrix}^T$ which equals

$$\begin{bmatrix} (AB^1)^T \\ (AB^2)^T \\ \ldots \\ (AB^n)^T \end{bmatrix} = \begin{bmatrix} (B^1)^T A^T \\ (B^2)^T A^T \\ \ldots \\ (B^n)^T A^T \end{bmatrix} = B^T A^T.$$

There is another way to prove this, by looking at the value at $(AB)^T_{ij}$. □

Note that this fact can be expanded, such that $(ABC)^T = C^T B^T A^T$.

Proposition 2. Let x and y be vectors. Then, $x^T y = (x \cdot y)$.

Proof. Let $x = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ and $y = \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix}$. Then, we have $x^T = \begin{bmatrix} 1 & 2 & 3 \end{bmatrix}$ such that

$$x^T y = 1 \cdot 2 + 2 \cdot 4 + 3 \cdot 6.$$

which is the dot product. □

What about the other way? Note that xy^T is 3×3 and is a rank 1 matrix. To elaborate, let $A = xy^T$ and z be any vector. Then, we have that

$$Az = x(y^T z) = (y \cdot z)x.$$

which is a multiple of x.

Definition 18. A is **rank** 1 because the image of A contains a line (x, dimension 1).

Proposition 3. $(A^{-1})^T = (A^T)^{-1}$

Proof. Proof with the identity. □

Proposition 4. $x^T(Ay) = (x \cdot Ay) = (Ay)^T x = (A^T y \cdot x)$ for every vector x, y. Note that this can be taken as the definition of the transpose.

Definition 19. An $n \times n$ matrix S is **symmetric** if $S^T = S$.

In row reduction, we saw that exchanging two rows is represented by the matrix P_{ij}.

Example.

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} = P_{34}.$$

Definition 20. In a **permutation matrix**, all entries are 1 or 0, and there is exactly one 1 on every row. More formally, for every i, the exists j such that $P_{ij} = 1$ and $P_{ij'} = 0$ for all $j' \neq j$.

For a permutation σ, the permutation matrix is defined as $P_{i\sigma(i)} = 1$ and otherwise $P_{ij} = 0$. Also note that if P, Q are permutation matrices, then PQ is also a permutation matrix.

Note that elimination and row exchange can be done in reverse order. Simply permute the rows, then eliminate, rather than eliminate the rows, then permute.

Proposition 5. If S is symmetric, we have

$$S = LDU.$$

and

$$S^T = U^T D L^T = S.$$

which means we can write S as

$$S = LDL^T.$$

2 Vector Spaces

Example. Let $f(x)$ be a continuous function from $[0, 1] \to R$. This is a vector space.

Example. Let $p(x)$ be a polynomial of degree $\leq n$. This is also a vector space.

In theory, we can have a vector space much more generally than $\mathbb{R}^n$.

Definition 21. Let $x_1, \ldots, x_m$ be vectors. Then, $V = \text{span}\{x_1, \ldots, x_m\}$ is a **subspace**.

Definition 22. Let V be a vector space such that $x_1, \ldots, x_m \in V$. Suppose that $\text{span}\{x_1, \ldots, x_m\} = V$. Then, V has **finite dimension**.

Definition 23. The vectors $x_1, \ldots, x_m$ are a **generating set**.

Example. V, the vector space of all continuous functions, is not finite.

Example. V, the vector space of all polynomials with degree $\leq n$, is finite. Consider the span of $1, x, x^2, \ldots, x^n$.

Lecture 5: Vector, Sub, Column and Null Spaces

Theorem 11. If AB is invertible, A, B is invertible.

Proof. if AB is invertible, there exists $C = (AB)^{-1}$ such that $(AB)C = I$. Then, $A(BC) = I$, and $BC = A^{-1}$. □

Proposition 6. For a permutation matrix,

$$P^{-1} = P^T.$$

which is also a permutation matrix.

Proposition 7. If A can be row reduced without row permutations, then

$$A = LU.$$

Proposition 8. If A is invertible, one can write

$$A = LDU_1.$$

Note. The product of two symmetric matrices are not necessarily symmetric.

Definition 24. V is a **vector space** if there is a function $V \times V \to V$ denoted $+$, which is commutative, associative, and has negation and null element and if there is another function $\mathbb{R} \times V \to V$ which is distributive and has a null element.

Note that $\mathbb{R}^n$ is a vector space. A subspace of $\mathbb{R}^n$ is also a vector space. Polynomials of degree $\leq n$ also form a vector space.

Definition 25. Let $B = \{x_1, x_2, \ldots x_n\}$. Then, the **span** of B is the set of linear combinations of all x_i. B is **generating** if $\text{span}\, B = V$.

Note. $\text{span}\, B$ is the smallest subspace of V that contains B.

Definition 26. We say that B is **linearly independent** if $\sum_i \alpha_i x_i = 0 \Rightarrow$ all $\alpha_i = 0$.

Proposition 9. Let B be generating. If B is not linearly independent, we can eliminate one element from B, and get smaller B' that is still generating.

Proof. Then some α_i is nonzero. Assuming α_1 is non-zero, $x_1 = \sum_{i \neq 1}^{n} \frac{\alpha_i}{\alpha 1} x_i$. Then, we have

$$\begin{aligned} y &= \sum_{i=1}^{n} \gamma_i x_i \qquad (B \text{ generating}) \\ &= \sum_{i \neq 1}^{n} \gamma_1 \frac{\alpha_i}{\alpha_1} x_i + \sum_{i=2}^{n} \gamma_i x_i \\ \beta_i &= \gamma_1 \frac{\alpha_i}{\alpha_1} + \gamma_i. \end{aligned}$$

Repeating this elimination process yields a set $D = \{x_1, \ldots, x_d\}$ that is minimal. This object D is called a **basis**. In other words, every vector x can be written as

$$x = \sum_{i=1}^{d} \alpha_i x_i \quad \alpha_i \in \mathbb{R}.$$

in a unique way. □

In other words, a basis is a mapping from $V \to \mathbb{R}^d$. The basis for polynomials is $B = \{1, x, x^2, \ldots, x^n\}$. The basis for vectors in $\mathbb{R}^3$ can be $\{\hat{\imath}, \hat{\jmath}, \hat{k}\}$, etc.

> **Note.** All bases for the same vector space have the same dimension.