University of California, Berkeley

Literally everything I know about

Linear Algebra

Warren Alphonso

A very reductionist summary of Linear Algebra and its Applications by Lay, Lay, and McDonald, as well as Linear Algebra Done Wrong by Treil.

May 30, 2019

Contents

Chapter 1

Basic Notations

1.1 Vector Spaces

Definition 1.1.1. A **vector space** V is a collection of vectors, along with vector addition and scalar multiplication defined such that for vectors u, v, and w:

1. Commutative: $v + w = w + v$
2. Associative: $(u + v) + w = u + (v + w)$
3. Zero vector: $v + 0 = v$
4. Additive inverse: $v + w = 0$
5. Multiplicative identity: $1v = v$
6. Multiplicative associative: $(\alpha\beta)v = \alpha(\beta v)$
7. Distribution of scalars: $\alpha(u + v) = \alpha u + \alpha v$
8. Distribution of vectors: $(\alpha + \beta)u = \alpha u + \beta u$

These properties are simply to ensure that vector spaces are **abelian groups**.

Definition 1.1.2. An $m \times n$ **matrix** is an array with m rows and n columns. Elements of a matrix are called *entries*. Given a matrix A, its **transpose** is defined as the matrix whose columns are A's rows, so A^T is an $n \times m$ matrix.

1.2 Linear Combinations

Definition 1.2.1. A **linear combination** of vectors $v_1, \cdots, v_p \in V$ is a sum of the form

$$\alpha_1 v_1 + \cdots + \alpha_p v_p = \sum_{k=1}^{p} \alpha_k v_k$$

Definition 1.2.2. A set of vectors $v_1, \cdots, v_n$ is said to be **linearly independent** if the equation

$$\alpha_1 v_1 + \cdots + \alpha_n v_n = 0$$

has only the trivial solution.

Definition 1.2.3. A **basis** is a set of vectors $v_1, \cdots, v_n \in V$ such that any vector $u \in V$ has a *unique* representation as a linear combination

$$u = \alpha_1 v_1 + \cdots + \alpha_n v_n$$

The coefficients $\alpha_1, \cdots, \alpha_n$ are called *coordinates* of u.

Fundamentally, our definition of basis states that a basis must be spanning and unique. In order for a representation to be unique, we know the basis must be linearly independent.

Theorem 1.2.1. *A set of vectors $v_1, \cdots, v_p \in V$ is a basis if and only if it is linearly independent and complete (spanning).*

Proof. We already know basis must be linearly independent and spanning, so we just need to prove the other direction.
Suppose the set $v_1, \cdots, v_p$ is linearly independent. Then we know for some vector $u \in V$:

$$u = \sum_{k=1}^{n} \alpha_k v_k$$

All that is remaining is to prove this representation is unique.
Suppose there is another representation, $u = \sum_{k=1}^{n} \beta_k v_k$. Then,

$$\sum_{k=1}^{n} (\alpha_k - \beta_k) v_k = v - v = 0$$

Since the set is linearly independent, we know $\alpha_k - \beta_k = 0$. Thus, the representation is unique. ■

1.3 Linear Transformations

Definition 1.3.1. A **transformation** T from set X to set Y assigns a value $y \in Y$ for every value $x \in X$: $y = T(x)$. X is called the *domain* of T, Y is called the *codomain* of T, and the set of all $T(x)$ is called the *range* of T.
Let V, W be vector spaces. A transformation $T : V \to W$ is **linear** if:

1. $T(u + v) = T(u) + T(v)$
2. $T(\alpha v) = \alpha T(v)$

A mapping $T : \mathbb{F}^n \to \mathbb{F}^m$ is *onto* $\mathbb{F}^m$ if each b in $\mathbb{F}^m$ is the image of at least one x in $\mathbb{F}^n$.

A mapping $T : \mathbb{F}^n \to \mathbb{F}^m$ is *one-to-one* if each b in $\mathbb{F}^m$ is the image of at most one x in $\mathbb{F}^n$.

We can represent linear transformations with matrices. To represent a transformation $T : \mathbb{F}^n \to \mathbb{F}^m$, we need to only how our n basis vectors are transformed. To see this, note that any vector $u = \alpha_1 v_1 + \cdots + \alpha_n v_n$. So $T(x) = \alpha 1 T(v_1) + \cdots + \alpha_n T(v_n)$. If we join the vectors $T(v_1), \cdots, T(v_n)$ in a matrix $A = \begin{bmatrix} T(v_1) & \cdots & T(v_n) \end{bmatrix}$, we have captured all the information about T.

Definition 1.3.2. There are two ways to approach **matrix-vector multiplication**:
Column by coordinate rule: Multiply each column of the matrix by the corresponding coordinate of the vector and add.

$$\begin{bmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = 1 \begin{bmatrix} 1 \\ 3 \end{bmatrix} + 2 \begin{bmatrix} 2 \\ 2 \end{bmatrix} + 3 \begin{bmatrix} 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 14 \\ 10 \end{bmatrix}$$

Row by column rule: To get entry k of the result, multiply row k of the matrix with the column.

$$\begin{bmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 \cdot 1 + 2 \cdot 2 + 3 \cdot 3 \\ 3 \cdot 1 + 2 \cdot 2 + 1 \cdot 3 \end{bmatrix} = \begin{bmatrix} 14 \\ 10 \end{bmatrix}$$

Definition 1.3.3. The natural extension to **matrix multiplication** of two matrices AB is to multiply A by each column of B.

$$AB = A \begin{bmatrix} b_1 & \cdots & b_n \end{bmatrix} = \begin{bmatrix} Ab_1 & \cdots & Ab_n \end{bmatrix}$$

Using the *row by columns rule*, we can see that the $(AB)_{j,k} =$ (row j of A)$\cdot$(column k of B). This also means AB is only defined if A is $m \times n$ and B is $n \times r$.
Let A be an $m \times n$ matrix and let B and C have sizes for which the indicated sums and products are defined.

1. $A(BC) = (AB)C$
2. $A(B + C) = AB + AC$
3. $(B + C)A = BA + CA$
4. $\alpha(AB) = (\alpha A)B = A(\alpha B)$
5. $I_m A = A = AI_n$

Warnings:

1. In general, $AB \neq BA$
2. If $AB = AC$, then it is **not true** in general that $B = C$
3. If $AB = 0$, then it is **not true** always that $A = 0$ or $B = 0$

Definition 1.3.4. The **transpose** of A is the matrix whose columns are formed from the corresponding rows of A, denoted by A^T.

1. $(A^T)^T = A$
2. $(A+B)^T = A^T + B^T$
3. $(cA)^T = cA^T$
4. $(AB)^T = B^T A^T$

To understand the final property, let AB denote a $n \times m$ matrix so that

$$AB = \begin{bmatrix} A_{1*} \cdot B_1 & A_{1*} \cdot B_2 & \cdots & A_{1*} \cdot B_m \\ A_{2*} \cdot B_1 & A_{2*} \cdot B_2 & \cdots & A_{2*} \cdot B_m \\ \vdots & \vdots & \vdots & \vdots \\ A_{n*} \cdot B_1 & A_{n*} \cdot B_2 & \cdots & A_{n*} \cdot B_m \end{bmatrix}$$

$$(AB)^T = \begin{bmatrix} A_{1*} \cdot B_1 & A_{2*} \cdot B_1 & \cdots & A_{n*} \cdot B_1 \\ A_{1*} \cdot B_2 & A_{2*} \cdot B_2 & \cdots & A_{n*} \cdot B_2 \\ \vdots & \vdots & \vdots & \vdots \\ A_{1*} \cdot B_m & A_{2*} \cdot B_m & \cdots & A_{n*} \cdot B_m \end{bmatrix}$$

where $A_{\alpha *}$ denotes the αth row of a A. Note that $(AB)^T_{jk}$ = (row k of A) $\cdot$ (column j of B) = (row j of B^T) $\cdot$ (column k of A^T).

Definition 1.3.5. For a *square* matrix A, its **trace** is the sum of its diagonal entries.

$$trace(A) = \sum_{k=1}^{n} a_{k,k}$$

Theorem 1.3.1. *Let A and B be size $m \times n$ and $n \times m$, respectively. Then*

$$trace(AB) = trace(BA)$$

Proof. We need only show that the diagonal entries of AB are the same as the diagonal entries of BA. Using the *row by columns rule*, we know $(AB)_{j,k}$ = (row j of A)$\cdot$(column k of B), so when applying to diagonal entries we get $(AB)_{k,k} = (BA)_{k,k}$ = (row k of A) $\cdot$ (column k of B) ■

1.4 Invertible Transformations and Isomorphisms

Definition 1.4.1. An $n \times n$ matrix A is **invertible** if there is an $n \times n$ matrix A^{-1} such that $A^{-1}A = I$.

Theorem 1.4.1. *If A and B are invertible (and such that AB is defined), then the product AB is invertible and*

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

Proof. Direct computation:

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

and

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

■

Theorem 1.4.2. *If A is invertible, then A^T is also invertible and*

$$(A^{-1})^T = (A^T)^{-1}$$

Proof. Using $(AB)^T = B^T A^T$,

$$(A^{-1})^T(A^T) = (AA^{-1})^T = I$$

and

$$A^T(A^{-1})^T = (A^{-1}A)^T = I$$

■

Definition 1.4.2. An invertible linear transformation $A : V \to W$ is called an **isomorphism**. The two vector spaces V and W for which A is defined are called **isomorphic**, denoted $V \cong W$.

Isomorphic spaces can be considered different representations of the *same* space. The following theorem is an example.

Theorem 1.4.3. *Let $A : V \to W$ be an isomorphism, and let $v_1, \cdots, v_n$ be a basis in V. Then $Av_1, \cdots, Av_n$ is a basis in W.*

Proof. Because V and W are isomorphic, every $w \in W$ can be represented as some $v \in V$ by applying A^{-1}. For arbitrary $w \in W$

$$A^{-1}w = v = \sum_{k=1}^{n} \alpha_k v_k$$

Then we apply A to get

$$Av = w = \sum_{k=1}^{n} \alpha_k A v_k$$

■

1.5 Subspaces

Definition 1.5.1. A **subspace** of vector space V is a non-empty subset $V_0 \subset V$ which is also a vector space.

For any linear transformation $A : V \to W$, we can associate the following subspaces:

1. The *null space*, or *kernel*, of A which consists of all vectors $v \in V$ such that $Av = 0$.
2. The *range* of A which is the set of all vectors $w \in W$ which can be represented as $w = Av$ for $v \in V$.

By the *column by coordinate rule*, we know that any vector in $Range(A)$ can be represented as a weighted sum of the column vectors of A, which is why the term Column Space is sometimes used to refer to Range.