# Linear Algebra

Warren Alphonso

May 18, 2019

A very reductionist summary of key Linear Algebra concepts.

## 1 Systems of Linear Equations

**Definition 1.1.** A **linear equation** is an equation that can be written in the form

$$a_1x_1 + ... + a_nx_n = b$$

where $b$ and the coefficients $a_k$ are real or complex numbers.

We can record the important information of a system of linear equations in a matrix. Given the system

$$x_1 - 2x_2 + x_3 = 0$$

$$2x_2 - 8x_3 = 8$$

$$5x_1 - 5x_3 = 10$$

we place the coefficients of each variable aligned in columns

$$\begin{bmatrix} 1 & -2 & 1 \\ 0 & 2 & -8 \\ 5 & 0 & -5 \end{bmatrix}$$

This is called the **coefficient matrix** and

$$\begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ 5 & 0 & -5 & 10 \end{bmatrix}$$

is called the **augmented matrix**. The size of a matrix is described as **m x n** where $m$ denotes the number of rows and $n$ the number of columns.

**Definition 1.2. Elementary Row Operations:**

1. (Replacement) Replace one row by the sum of itself and a multiple of another row.

2. (Interchange) Interchange two rows.

3. (Scaling) Multiply all entries in a row by a nonzero constant.

**Definition 1.3.** **Row Echelon form** denotes a matrix with:

1. All nonzero rows are above any rows of all zeros.
2. Each leading entry of a row is in a column to the right of the leading entry of the row above it.
3. All entries in column below a leading entry are zeros.

**Reduced Row Echelon form** means the leading entry in nonzero rows is 1.

**Parallelogram Rule for Addition:** If **u** and **v** in $\mathbb{R}^2$ are represented as points on the plane, then $\mathbf{u} + \mathbf{v}$ corresponds to the fourth vertex of the parallelogram whose other vertices are 0, **u**, and **v**.

**Definition 1.4.** **Span**$\{v_1, ..., v_p\}$ denotes the set of all vectors formed by $c_1 v_1 + ... + c_p v_p$.

**Definition 1.5.** A set of vectors $\{v_1, ..., v_p\}$ is said to be **linearly independent** if the equation

$$c_1 v_1 + ... + c_p v_p = 0$$

has only the trivial solution.

**Theorem 1.** *If a set contains more vectors than there are entries in each vector, then the set is linearly independent. That is, the set $\{v_1, ..., v_p\}$ in $\mathbb{R}^n$ is linearly dependent if $p > n$.*

*Proof.* Let $A = [v_1 \cdots v_p]$. Then $A$ is $n \times p$, and the equation $Ax = 0$ corresponds to a system of $n$ equations in $p$ unknowns. In $Ax = b$, the $x$ vector must have dimension $p$, so if $p > n$, then there are more variables than equations, so $Ax = 0$ has a nontrivial solution, and the columns of $A$ are linearly dependent. $\square$

An alternate way to conceptualize matrix multiplication: A **transformation** (or **function** or **mapping**) $T$ from $\mathbb{R}^n$ to $\mathbb{R}^m$ is a rule that assigns to each vector $x$ in $\mathbb{R}^n$ a vector $T(x)$ in $\mathbb{R}^m$. The set $\mathbb{R}^n$ is called the **domain** of $T$ and $\mathbb{R}^m$ is called the **codomain** of T. The set of all images $T(x)$ is called the **range** of $T$.

**Definition 1.6.** A transformation $T$ is **linear** if they preserve vector addition and scalar multiplication. That is:

1. $T(u + v) = T(u) + T(v)$
2. $T(cu) = cT(u)$ for all scalars $c$

Every matrix transformation is a linear transformation. These two requirements mean that $T(0) = 0$ for linear transformations.

**Theorem 2.** *Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Then there exists a* ***unique*** *matrix $A$ such that*

$$T(x) = Ax$$

*In fact, $A$ is the $m \times n$ matrix whose $j$th column is the vector $T(e_j)$, where $e_j$ is the $j$th column of the identity matrix in $\mathbb{R}^n$:*

$$A = [T(e_1) \cdots T(e_n)]$$

*Proof.* Write $x = I_n x = [e_1 \cdots e_n]x = x_1e_1 + \cdots + x_ne_n$, and use the linearity of $T$ to compute

$$T(x) = T(x_1e_1 + \cdots x_ne_n) = x_1T(e_1) + \cdots + x_nT(e_n)$$

$$\begin{bmatrix} T(e_1) \cdots T(e_n) \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = Ax$$

□

**Definition 1.7.** A mapping $T : \mathbb{R}^n \to \mathbb{R}^m$ is **onto** $\mathbb{R}^m$ if each $b$ in $\mathbb{R}^m$ is the image of at least one $x$ in $\mathbb{R}^n$.

**Definition 1.8.** A mapping $T : \mathbb{R}^n \to \mathbb{R}^m$ is **one-to-one** if each $b$ in $\mathbb{R}^m$ is the image of at most one $x$ in $\mathbb{R}^n$.

## 2 Matrix Algebra

**Definition 2.1.** If $A$ is an $m \times n$ matrix and $B$ is an $n \times p$ matrix, then

$$AB = A\begin{bmatrix} b_1 & b_2 & \cdots & b_p \end{bmatrix} = \begin{bmatrix} Ab_1 & Ab_2 & \cdots & Ab_p \end{bmatrix}$$

**Definition 2.2.** 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. $r(AB) = (rA)B = A(rB)$
5. $I_mA = A = AI_n$

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

- $(A^T)^T = A$
- $(A+B)^T = A^T + B^T$
- $(cA)^T = rA^T$
- $(AB)^T = B^T A^T$

**Definition 2.4.** An $n \times n$ matrix $A$ is **invertible** if there is an $n \times n$ matrix $A^{-1}$ such that $A^{-1}A = I$.

- $(A^{-1})^{-1} = A$
- If $A$ and $B$ are $n \times n$ invertible matrices then so is $AB$. And $(AB)^{-1} = B^{-1}A^{-1}$
- $(A^T)^{-1} = (A^{-1})^T$

To compute the inverse, solve the equation $AB = I$, by row-reducing the augmented matrix [A I], until you get [I B].

**Definition 2.5.** A **subspace** of $\mathbb{R}^n$ is any set $H$ in $\mathbb{R}^n$ that is closed under addition and scalar multiplication. That is:

1. The zero vector is in $H$
2. For each $u$ and $v$ in $H$, the sum $u + v$ is in $H$
3. For each $u$ in $H$ and each scalar $c$, the vector $cu$ is in $H$

**Definition 2.6.** The **column space** of matrix $A$ is the set of all linear combinations of the columns of $A$, denoted by $Col(A)$.

**Definition 2.7.** The **null space** of a matrix $A$ is the set of all solutions for $Ax = 0$, denoted by $Nul(A)$.

**Definition 2.8.** A **basis** for a subspace $H$ in $\mathbb{R}^n$ is a linearly independent set in $H$ that spans $H$.

Using the basis for a subspace $H$ is preferable because any vector in $H$ can only be written in one way as a linear combination of the basis vectors.

*Proof.* Suppose $\mathbb{B} = \{\ b_1\ ,\ ...,\ b_p\}$ is a basis for $H$, and suppose a vector $x$ in $H$ can be generated in two ways:

$$x = c_1 b_1 + \cdots + c_p b_p \text{ and } x = d_1 b_1 + \cdots + d_p b_p$$

Subtracting gives us:

$$0 = (c_1 - d_1)b_1 + \cdots + (c_p - d_p)b_p$$

Since $\mathbb{B}$ is linearly independent, the weights must all be zero, so $c_j = d_j$ so the two representations are really just the same. $\square$

**Definition 2.9.** The **dimension** of a nonzero subspace $H$ is the number of vectors in any basis for $H$. The dimension of the zero subspace $\{0\}$ is defined to be zero.

**Definition 2.10.** The **rank** of a matrix $A$ is the dimension of the column space of $A$.

**Theorem 3.** *If a matrix $A$ has $n$ columns, then $Rank(A) + Dim(Nul(A)) = n$.*

*Proof.* An intuitive understanding for this can be achieved by restating the theorem as follows:

$$\Big(\text{num of pivot columns}\Big) + \Big(\text{num of nonpivot columns}\Big) = \Big(\text{num of columns}\Big)$$

$\square$