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 $\mathbf{u}$ and $\mathbf{v}$ 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, $\mathbf{u}$, and $\mathbf{v}$.

Definition 1.4. **Span**$\{v_1, ..., v_p\}$ denotes the set of all vectors formed by $c_1v_1 + ... + c_pv_p$.

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

$$c_1v_1 + ... + c_pv_p = 0$$

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. □

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 jth column is the vector $T(e_j)$, where e_j is the jth column of the identity matrix in $\mathbb{R}^n$:

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

Proof. Write $x = I_n x = [e_1 \cdots e_n]x = x_1 e_1 + \cdots + x_n e_n$, and use the linearity of T to compute

$$T(x) = T(x_1 e_1 + \cdots x_n e_n) = x_1 T(e_1) + \cdots + x_n T(e_n)$$

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

□

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. $I_m A = A = AI_n$

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 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:

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

□

3 Determinants

Definition 3.1. Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. If $ad - bc \neq 0$, then A is invertible and

$$A^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

The quantity $ad - bc$ is the **determinant** of the matrix. If the determinant is 0, the matrix A is not invertible.

Definition 3.2. To generalize, the determinant of an $n \times n$ matrix A can be computed using a **cofactor expansion** across any row or down any column. The expansion across the ith row is

$$det(A) = a_{i1}C_{i1} + a_{i2}C_{i2} + \cdots + a_{in}C_{in}$$

where $C_{ij} = (-1)^{i+j} det(A_{ij})$

Theorem 4. *If A is an upper triangular matrix, then det(A) is the product of the entries on the main diagonal.*

Proof. Cofactoring an upper triangular matrix by using the first column ultimately leads to continuously multiplying the upper left item by the determinant of the smaller matrix. For example,

$$A = \begin{bmatrix} 3 & 2 & 9 \\ 0 & 4 & -1 \\ 0 & 0 & -8 \end{bmatrix}$$

Then,

$$det(A) = 3 \cdot det(\begin{bmatrix} 4 & -1 \\ 0 & -8 \end{bmatrix} = 3 \cdot -32 = -96 = 3 \cdot 4 \cdot -8$$

□

Definition 3.3. Determinants after Row Operations

1. If a multiple of a row in matrix A is added to another row to produce matrix B, then $det(B) = det(A)$
2. If two rows in A are swapped to produce B, then $det(B) = -det(A)$
3. If one row in A is multiplied by k to produce B, then $det(B) = k \cdot det(A)$

These identities can be used to easily find determinants of square matrices. Once we reduce a matrix A to upper triangular form B, we know $det(B) = (-1)^r det(A)$ if r is the number of row swaps we performed. If we cannot reduce to row echelon form, we know the determinant must be 0 since A must not be invertible.

Theorem 5. *If A is an $n \times n$ matrix, then $det(A^T) = det(A)$.*

Proof. We proceed by induction. The theorem is trivially true for $n = 1$. Assume the theorem is true for $k \times k$ matrices. We will show it holds for $n = k+1$. The cofactor of a_{1j} in A equals the cofactor of a_{j1} in A^T because it is a $k \times k$ determinant. Thus, the cofactor of $det(A^T)$ down the first *column* equals the cofactor of $det(A)$ across the first *row*, so A and A^T have equal determinants. Thus, the statement is true for all n. ☐

Theorem 6. *If A and B are $n \times n$ matrices, then $det(AB) = det(A)det(B)$.*

Theorem 7. *Cramer's Rule Let A be an invertible $n \times n$ matrix. For any b in $\mathbb{R}^n$, the unique solution x of $Ax = b$ has entries given by*

$$x_i = \frac{det(A_i(b))}{det(A)} \text{ for } i = 1, 2, ..., n$$

where $A_i(b)$ denotes the matrix obtained by replacing A's ith column with b.

Proof. Denote the columns of A by $a_1, ..., a_n$ and the columns of the $n \times n$ identity matrix by $e_1, ..., e_n$. If $Ax = b$, the definition of matrix multiplication tells us

$$A \cdot I_i(x) = A \begin{bmatrix} e_1 & \cdots & x & \cdots & e_n \end{bmatrix} = \begin{bmatrix} Ae_1 & \cdots & Ax & \cdots & Ae_n \end{bmatrix}$$

$$= \begin{bmatrix} a_1 & \cdots & b & \cdots & a_n \end{bmatrix} = A_i(b)$$

Using the multiplicative property of determinants,

$$(det(A))(det(I_i(x))) = det(A_i(b))$$

Since $det(I_i(x))$ is x, we just divide by $det(A)$. ☐