Linear Algebra: Systems of Linear Equations

Arjun Vardhan

†

Created: 11th July 2022
Last updated: 2nd August 2022

1 Introduction

- Consider a system of m linear equations with n unknowns: $\sum_{k=1}^{n} a_{1,k}x_k = b_1$, $\sum_{k=1}^{n} a_{2,k}x_k = b_2$, ..., $\sum_{k=1}^{n} a_{m,k}x_k = b_m$. Taking the coefficients from the equations, we can form a matrix $A = \begin{pmatrix} a_{1,1} & a_{1,2} & ... & a_{1,n} \\ a_{2,1} & a_{2,2} & ... & a_{2,n} \\ ... & ... & ... & ... \\ a_{m,1} & a_{m,2} & ... & a_{m,n} \end{pmatrix}$. Then the linear system can be written in the form of a matrix equation $Ax = b$, where $x = (x_1, x_2, ..., x_n)^T$ and $b = (b_1, b_2, ..., b_m)^T$.
- A here is the coefficient matrix of the system. If we join the coefficient matrix to the vector b, we get the augmented matrix $\begin{pmatrix} a_{1,1} & a_{1,2} & ... & a_{1,n} & b_1 \\ a_{2,1} & a_{2,2} & ... & a_{2,n} & b_2 \\ ... & ... & ... & ... & ... \\ a_{m,1} & a_{m,2} & ... & a_{m,n} & b_m \end{pmatrix}$, which contains all the information necessary to solve the system.

2 Echelon Form and Reduced Echelon Form

- Linear systems can be solved by Gauss-Jordan elimination, also known as row reduction. By performing elementary row operations on the augmented matrix, we can bring it into echelon form.
- There are three types of elementary row operations:
  1. Row exchange: Interchanging two rows of the matrix.
  2. Scaling: Multiplying a row with a scalar.
  3. Row replacement: Replacing a row by its sum with a constant multiple of another row.

  None of these operations alter the solution set of the linear system.
- An elementary matrix is one that is obtained by performing exactly one elementary row operation to the identity matrix. Performing elementary row operations on a matrix is equivalent to left-multiplying it by an elementary matrix.
- Two matrices are said to be row equivalent if one can be obtained from the other through a finite number of elementary row operations. Alternatively, A and B are row equivalent if there exist elementary matrices $E_1, E_2, ..., E_n$ such that $A = E_1E_2...E_nB$.
- For each row in a matrix, the leftmost nonzero entry is called the pivot entry or just pivot.
- A matrix is said to be in echelon form or row echelon form when all zero rows are below all nonzero entries and the pivot entry of each nonzero row is strictly to the right of the pivot of the row above it.
- A matrix is said to be in reduced echelon form or reduced row echelon form when it is in echelon form, all pivot entries equal 1 and all entries above pivots are 0.

- When in echelon form, the variables corresponding to columns without pivots are called free variables.

3 Analyzing the Pivots

- A linear system is said to be inconsistent if it has no solutions.
- A linear system is inconsistent if and only if there is a pivot in the last column of an echelon form of the augmented matrix. I.e, if the echelon form has a row of the type (0 0 0 ... 0 b), with $b \neq 0$. In this case one of the equations ends up being $0x_1 + 0x_2 + ... + 0x_n = b$, which obviously has no solution in $\mathbb{C}$.
- If a linear system has a solution, the solution is unique if and only the system has no free variables, i.e, when the echelon form of the coefficient matrix has a pivot in every column.
- The equation $Ax = b$ has a solution for any $b \in \mathbb{F}^m$ if and only if the echelon form of A has a pivot in every row. So a linear system is consistent only when its coefficient matrix has a pivot in every row when in echelon form.
- The equation $Ax = b$ has a unique solution for any $b \in \mathbb{F}^m$ if and only if the echelon form of A has a pivot in every row and every column. So a linear system has a unique solution only when its coefficient matrix has a pivot in every row and every column when in echelon form.
- **Let $v_1, v_2, ..., v_m \in \mathbb{F}^n$. Let $A = [v_1, v_2, ..., v_n]$ be the matrix with these vectors as its columns. Then:**
  1. **The system $v_1, v_2, ..., v_m$ is linearly independent if and only if the echelon form of A has a pivot in every column.** *Proof:* The system $v_1, v_2, ..., v_m$ is linearly independent if the equation $x_1v_1 + x_2v_2 + ... + x_nv_n = 0$ has $x_1 = x_2 = ... = x_n = 0$ as its only solution, or equivalently, if the matrix equation $Ax = 0$ has $x = 0$ as its only solution. But $Ax = 0$ has a unique solution if and only if the echelon form of A has a pivot in every column. ■
  2. **The system $v_1, v_2, ..., v_m$ spans $\mathbb{F}^n$ if and only if the echelon form of A has a pivot in every row.** *Proof:* The system $v_1, v_2, ..., v_m$ is spans $\mathbb{F}^n$ if and only if $x_1v_1 + x_2v_2 + ...x_nv_n = b$ has a solution for any $b \in \mathbb{F}^n$. But this happens only when the echelon form of A has a pivot in every row. ■
  3. **Corollary: The system $v_1, v_2, ..., v_m$ is a basis in $\mathbb{F}^n$ if and only if the echelon form of A has a pivot in every row and every column.**
- **Any linearly independent system of vectors in $\mathbb{F}^n$ cannot have more than n vectors in it.** *Proof 1:* Let $v_1, v_2, ..., v_m \in \mathbb{F}^n$ be a linearly independent system. Let $A = [v_1, v_2, ..., v_m]$ be an $n \times m$ matrix. Then the echelon form of A must have a pivot in every column, which is impossible if $m > n$, as the number of pivots cannot be greater than the number of rows. ■ *Proof 2:* Let $e_1, e_2, ..., e_n$ be the standard basis of $\mathbb{F}^n$, and let $v_1, v_2, ..., v_n, v_{n+1}$ be a linearly independent system in $\mathbb{F}^n$. Then $v_1 = \sum_{k=1}^{n} a_k e_k$. As $v_1 \neq 0$ there is at least one $a_k \neq 0$. Without loss of generality, suppose $a_1 \neq 0$. Then $e_1 = \frac{1}{a_1}(v_1 - \sum_{k=2}^{n} a_k e_k)$. Thus $v_1, e_2, ..., e_n$ spans $\mathbb{F}^n$. The assumed linear independence of $v_1, v_2, ..., v_{n+1}$ allows us to repeat this step until we get that $v_1, v_2, ..., v_n$ spans $\mathbb{F}^n$. Then v_{n+1} is a linear combination of $v_1, ..., v_n$ which contradicts the linear independence of the system. ■
- **Any two bases in a vector space V have the same number of vectors in them.** *Proof:* Let $v_1, v_2, ..., v_n$ and $w_1, w_2, ..., w_m$ be bases in V. Without loss of generality let $n \leq m$. Let $T : \mathbb{F}^n \to V$, $T(e_k) = v_k$ for all k, where $e_1, ..., e_n$ is the standard basis in $\mathbb{F}^n$. Clearly T is an isomorphism, and so T^{-1} is also an isomorphism. Then $T^{-1}(w_1), T^{-1}(w_2), ..., T^{-1}(w_m)$ is a basis in $\mathbb{F}^n$, and since bases are linearly independent, $m \leq n$. Therefore $n = m$. ■
- **Any basis in $\mathbb{F}^n$ must have exactly n vectors in it.** *Proof 1:* As the standard basis in $\mathbb{F}^n$ has n vectors in it, every other basis must also have n elements. ■ *Proof 2:* Let $v_1, v_2, ..., v_m$ be a

basis in $\mathbb{F}^n$. Let $A = [v_1, v_2, ..., v_m]$ be the $n \times m$ matrix with those vectors as its columns. As this system is a basis, the echelon form of A has a pivot in every row and every column. Since there are n rows, there must be n pivots. As each column can have at most one pivot, there are n columns in A. ■

- **Any spanning set in $\mathbb{F}^n$ must have at least n vectors in it.** *Proof 1:* Let $v_1, v_2, ..., v_m$ span $\mathbb{F}^n$. Then the $n \times m$ matrix A with those vectors as its columns has a pivot in every row in its echelon form. As the number of pivots cannot exceed the number of columns, $n \leq m$. ■ *Proof 2:*

- **A matrix A is invertible if and only if the echelon form of A has a pivot in every column and every row.** *Proof:* We know that A is invertible if and only if $Ax = b$ has a unique solution for every b, which is true if and only if every row and column has a pivot in A's echelon form. ■

- **Corollary: An invertible matrix must be a square matrix.**

- **If a square matrix is either left invertible or right invertible, then it is invertible.** *Proof:* Let A be an $n \times n$ matrix. If there exists a matrix B such that $BA = I$, then $Ax = 0$ has unique solution as $BAx = B0 \implies x = 0$. Thus there are no free variables and the echelon form of A has pivot in each column. Since it is a square matrix it also has a pivot in each row, and thus A is invertible. Now let there exist a matrix C such that $AC = I$. Let $x = Cb$, for some $b \in \mathbb{F}^n$. Then $Ax = ACb = Ib = b$. Thus for any $b \in \mathbb{F}^n$, the equation $Ax = b$ has a solution $x = Cb$. Thus A has a pivot in every row in echelon form, and since it is square it also has a pivot in every column. Thus A is invertible. ■

4 Finding A^{-1} by Row Reduction

- Since an invertible matrix must be square, and its echelon form must have a pivot in every column and every row, its reduced echelon form is an identity matrix. Every invertible matrix is thus row-equivalent to an identity matrix.

- To find the inverse of an $n \times n$ matrix A, form an augmented $n \times 2n$ matrix $[A|I]$, i.e, the identity matrix of size n to the right of A. Perform row operations to get the reduced echelon form of A. Then the matrix I to the right of A will have been turned into A^{-1}.

5 Dimension

- The dimension of a vector space V, denoted $\dim V$, is the number of vectors in a basis. The dimension of $\{0\}$ is defined as 0. A vector space which does not have a finite basis is said to have dimension ∞.

- As every finite spanning system contains a basis, it follows that a vector space is finite dimensional if and only if it contains a finite spanning system.

- If we want to check whether a system of vectors in a finite dimensional vector space is a basis, spanning or linearly independent, then it is best to use an isomorphism $T : V \to \mathbb{R}^n$, $n = \dim V$, as in $\mathbb{R}^n$ such problems can be solved by row reduction. Such an isomorphism always exists, as we can define a linear transformation that maps a basis of V to the standard basis in $\mathbb{R}^n$.

- **Any linearly independent system in a vector space V cannot have more than $\dim V$ vectors in it.** *Proof:* Let $v_1, v_2, ..., v_m \in V$ be a linearly independent system, and let $T : V \to \mathbb{R}^n$ be an isomorphism, where $n = \dim V$. Then $T(v_1), T(v_2), ..., T(v_m)$ is a linearly independent system in $\mathbb{R}^n$, and so $m \leq n$. ■

- **Any spanning system in a finite dimensional vector space V has at least $\dim V$ vectors.** *Proof:* Let $v_1, v_2, ..., v_m$ span V, and let $T : V \to \mathbb{R}^n$ be an isomorphism, where $n = \dim V$. Then $T(v_1), T(v_2), ..., T(v_m)$ spans $\mathbb{R}^n$, so $m \geq n$. ■

- **A linearly independent system of vectors in a finite dimensional vector space can be completed to a basis, i.e, if $v_1, v_2, ..., v_r$ are linearly independent in V, the one can find vectors $v_{r+1}, v_{r+2}, ..., v_n$ such that the system $v_1, v_2, ..., v_n$ is a basis.** *Proof:* Let

$v_1, v_2, ..., v_r$ be linearly independent, and let $n = \dim V$. As the system does not span V, we can find a vector v_{r+1} such that it does not belong to $\text{span}(v_1, v_2, ..., v_r)$. Then $v_1, v_2, ..., v_r, v_{r+1}$ is linearly independent. (Suppose this system was linearly dependent, i.e, there existed a non-trivial linear combination $a_1v_1 + ... + a_rv_r + a_{r+1}v_{r+1} = 0$. Here, $a_{r+1} \neq 0$ as otherwise we would get $a_1v_1 + ... + a_rv_r = 0$. So we get $v_{r+1} = \frac{1}{a_{r+1}}(a_1v_1 + ... + a_rv_r)$, which is a contradiction). Thus we can repeat this procedure until we have a system that spans V. This process must terminate as a linearly independent system in V cannot have more than n vectors. $\blacksquare$

- **Let V be a subspace of W, with $\dim W < \infty$. Then $\dim V \leq \dim W$, and if $\dim V = \dim W$, then $V = W$.** *Proof:* If $V = \{0\}$, then the theorem is obviously true, so let $V \neq \{0\}$. Take $v_1 \in V$. If $V = \text{span}(v_1)$, then we are done. Otherwise suppose we have a set of r linearly independent vectors in V. As shown above this can be completed to a basis. Since $V \subseteq W$, this system can have at most n vectors, so $\dim V \leq \dim W$. Now suppose that it does indeed have n vectors, i.e, $\dim V = \dim W$. Suppose $V \neq W$, and there exists $w \in W$ such that $w \notin \text{span}(v_1, v_2, ..., v_n)$. Then the system $v_1, ..., v_n, w$ would be linearly independent, which is a contradiction. Thus $V = W$. $\blacksquare$

- **Let $\dim V = n$. A system of vectors $v_1, v_2, ..., v_n \in V$ is linearly independent if and only if it spans V.** *Proof:* Let $v_1, v_2, ..., v_n$ be linearly independent. Suppose that this system does not span V. Then we can find $v_{n+1} \in V$ such that $v_{n+1} \notin \text{span}(v_1, v_2, ..., v_n)$. But then $v_1, v_2, ..., v_{n+1}$ would end up being linearly independent. So $v_1, v_2, ..., v_n$ spans V. Conversely, let $v_1, v_2, ..., v_n$ spans V. Suppose $v_1, v_2, ..., v_n$ is linearly dependent. Without loss of generality let $v_1 = a_2v_2 + a_3v_3 + ...a_nv_n$. Then $\text{span}(v_2, v_3, ..., v_n) = V$, but a spanning set in V must have at least n vectors. So $v_1, ..., v_n$ is linearly independent. $\blacksquare$

- **Corollary: Let $\dim V = n$. Then any linearly independent system or spanning system of n vectors is a basis.**

6 Solution Set of a Linear System

- A linear system is called homogenous if the right hand side of every equation is 0.

- If $Ax = b$ is a linear system, then $Ax = 0$ is called the homogenous linear system associated with it.

- **Let $Ax = b$ be a linear system, and let x_1 be a vector such that $Ax_1 = b$. Let H be the set of all solutions of the associated linear system $Ax = 0$. Then, $\{x_h + x_1 : x_h \in H\}$ is the set of all solutions of $Ax = b$.** *Proof:* $A(x_h + x_1) = Ax_h + Ax_1 = 0 + b = b$, so any solution of the form $x_h + x_1$, $x_h \in H$ is a solution of $Ax = b$. Now suppose x_0 satisfies $Ax_0 = b$. Then for $x_h = x_0 - x_1$, $Ax_h = Ax_0 - Ax_1 = b - b = 0$, so $x_h \in H$. Therefore any solution x_0 of $Ax = b$ can be represented in the form $x_1 + x_h$, $x_h \in H$. $\blacksquare$

7 Fundamental Subspaces of a Matrix, Rank

- If A is an $m \times n$ matrix, i.e, representing a linear transformation $A : \mathbb{F}^n \to \mathbb{F}^m$, then for any $w \in \text{Ran}(A)$, w can be represented as a linear combination of the columns of A. Thus $\text{Ran}(A)$ is also called the column space of A, denoted $\text{Col}(A)$.

- $\text{Ran}(A^T)$ is called the row space of A, and $\text{Ker}(A^T)$ is sometimes called the left null space of A.

- $\text{Ran}(A)$, $\text{Ker}(A)$, $\text{Ran}(A^T)$ and $\text{Ker}(A^T)$ are called the fundamental subspaces of the matrix A.

- Given a linear transformation/matrix, its rank, denoted $\text{rank}(A)$, is given by $\text{rank}\, A = \dim \text{Ran}(A)$. Its nullity, denoted $\text{nullity}(A)$, is given by $\text{nullity}(A) = \dim \text{Ker}(A)$.

- Let A be a matrix and let A_e be its echelon form. Then,

  1. The pivot columns of the original matrix A, i.e, the columns of A where we will have pivots after row reduction, give us a basis for $\text{Ran}(A)$. Thus $\text{rank}(A)$ is the number of pivot columns in the echelon form of A.
  2. The pivot rows of A_e give us a basis for the row space $\text{Ran}(A^T)$.

3. A basis for the null space $\text{Ker}(A)$ can be found by solving the equation $Ax = 0$.

- **Rank Theorem:** $\text{rank}(A) = \text{rank}(A^T)$**. Or, the column rank of a matrix equals its row rank.** *Proof:*
- **Rank-Nullity Theorem: Let** A **be an** $m \times n$ **matrix, i.e, a linear transformation from** $\mathbb{F}^n$ **to** $\mathbb{F}^m$**. Then,**
  1. $\text{nullity}(A) + \text{rank}(A) = n$. *Proof:*
  2. $\dim \text{Ker}(A^T) + \text{rank}(A) = n$. *Proof:*
- **Let** A **be an** $m \times n$ **matrix. Then the equation** $Ax = b$ **has a solution for every** $b \in \mathbb{R}^n$ **if and only if the equation** $A^T x = 0$ **has a unique (only the trivial solution).** *Proof:*
-

8 Represenation of a Linear Transformation in Arbitrary Bases

- Let V be a vector space with basis $B = \{b_1, b_2, ..., b_n\}$. Any vector $v \in V$ can be uniquely represented as $v = x_1 b_1 + x_2 b_2 + ... + x_n b_n$. The scalars $x_1, x_2, ..., x_n$ are called the coordinates of v in B. The coordinate vector of v relative to B is $[v]_B = (x_1, x_2, ..., x_n)$. The function $f : V \to \mathbb{F}^n$, $f(v) = [v]_B$ is an isomorphism.
- Let $T : V \to W$ be a linear transformation, and let $A = \{a_1, a_2, ..., a_n\}$ and $B = \{b_1, b_2, ..., b_n\}$ be bases in V and W respectively.
- A matrix A is said to be similar to a matrix B if there exists an invertible matrix Q such that $A = Q^{-1} B Q$.
-