MAT 307, Multivariable Calculus with Linear Algebra – Fall 2024

1 Gauss-Jordan Elimination

Recall that a solution to $A\vec{x} = \vec{b}$ means expressing $\vec{b}$ as a linear combination of the column vectors $\vec{v}_i$ of A, that is

$$A\vec{x} = x_1\vec{v}_1 + \cdots + x_n\vec{v}_n = \vec{b}.$$

Thus, a solution exists if and only if $\vec{b} \in \text{span}\{\vec{v}_1, \ldots, \vec{v}_n\}$. Also recall that the ***rank*** of a system of m linear equations in n unknowns, $A\vec{x} = \vec{b}$, is the maximal number of linearly independent rows in the augmented matrix $[A \mid \vec{b}]$.

You should feel concerned about this definition. To compute the rank, you would have to decide if every subcollection of rows is dependent or independent and then find the largest linearly independent one. That's a lot of work!

We will now devise an algorithm that help us to both compute the rank and to solve the system of linear equations simultaneously.

Definition 1. Two linear systems are ***equivalent*** if they have the same solution set.

We wish ***reversible*** operations taking a system to an ***equivalent*** (simpler) system.

Example 1. Solve the system

$$\begin{array}{lrcrcrcl} eq_1: & x & + & 3y & + & z & = & 1 \\ eq_2: & 2x & + & 4y & + & 7z & = & 2 \\ eq_3: & 4x & + & 10y & + & 9z & = & 4 \end{array}$$

We will "modify" the equations in such a way as to simplify them while preserving the solution set. Note, after each step below, we relabel the equations by order.

$$\begin{array}{rcrcrcl} x & + & 3y & + & z & = & 1 \\ 2x & + & 4y & + & 7z & = & 2 \\ 4x & + & 10y & + & 9z & = & 4 \end{array} \xrightarrow{\begin{bmatrix} eq_1 \\ eq_2 - 2eq_1 \\ eq_3 - 4eq_1 \end{bmatrix}} \begin{array}{rcrcrcl} x & + & 3y & + & z & = & 1 \\ 0 & - & 2y & + & 5z & = & 0 \\ 0 & - & 2y & + & 5z & = & 0 \end{array} \xrightarrow{\begin{bmatrix} eq_1 \\ eq_2 \\ eq_3 - eq_2 \end{bmatrix}}$$

$$\begin{array}{rcrcrcl} x & + & 3y & + & z & = & 1 \\ 0 & - & 2y & + & 5z & = & 0 \\ 0 & + & 0 & + & 0 & = & 0 \end{array} \xrightarrow{\begin{bmatrix} 2eq_1 \\ 3eq_2 \\ eq_3 \end{bmatrix}} \begin{array}{rcrcrcl} 2x & + & 6y & + & 2z & = & 2 \\ 0 & - & 6y & + & 15z & = & 0 \\ 0 & + & 0 & + & 0 & = & 0 \end{array} \xrightarrow{\begin{bmatrix} eq_1 + eq_2 \\ eq_2 \\ eq_3 \end{bmatrix}}$$

$$\begin{array}{rcrcrcl} 2x & + & 0 & + & 17z & = & 2 \\ 0 & - & 6y & + & 15z & = & 0 \\ 0 & + & 0 & + & 0 & = & 0 \end{array} \Longrightarrow \boxed{\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 - \frac{17}{2}z \\ \frac{5}{2}z \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + z \begin{bmatrix} -\frac{17}{2} \\ \frac{5}{2} \\ 1 \end{bmatrix}}$$

Notice that the solution is a line in $\mathbb{R}^3$. The vector $\begin{bmatrix}1 & 0 & 0\end{bmatrix}^t$ is called the ***translation vector*** and $\begin{bmatrix}-\frac{17}{2} & \frac{5}{2} & 1\end{bmatrix}^t$ is called a ***spanning vector***. Notice that z can be assigned any real value to get a solution for the system.

The above operations are called ***elementary row operations*** and can be applied to the matrix $[A \mid \vec{b}]$ directly. ***Gauss-Jordan or Gaussian Elimination*** is a process where successive ***elementary row operations*** bring the matrix into a ***(reduced) row echelon form*** — we will define this shortly.

The elimination process consists of three ***reversible*** elementary row operations:

1. Swapping two rows.
2. Scaling a row (i.e. multiply a row by a non-zero scalar).
3. Adding a multiple of one row to another.

Proposition 1. *Elementary row operations do not alter the solution set of the system.*

Proof. Swapping two equations in a system does not change the solution set. Similarly, eq_i and $a \cdot eq_i$ (for $a \neq 0$) define the same hyperplane, so scaling rows doesn't change the solution set. Lastly, let eq_i and eq_j are two different equations in the system. Notice that any solution $\vec{x}$ to eq_i ***and*** eq_j is also a solution to $eq_i' = eq_i + a \cdot eq_j$ ***and*** eq_j for any $a \in \mathbb{R}$. Similarly, any solution $\vec{x}'$ to eq_i' and eq_j is also a solution to $eq_i = eq_i' - a \cdot eq_j$ and eq_j for any $a \in \mathbb{R}$. Thus, we are done. □

The proof on the next theorem is rather technical and we omit it here.

Theorem 2. *Elementary row operations do not alter the rank of a matrix.*

Definition 2. A matrix R is in ***row echelon form*** if

- the first nonzero entry in any nonzero row occurs to the right of first such entry in the row directly above it.
- all zero rows are grouped together at the bottom.

The first nonzero entry in a row is called a ***pivot***. The column containing a pivot is called a ***pivot column***. If R represents a linear system, then the variable corresponding to that entry is called a ***pivot variable***.

Theorem 3. *Every matrix is (row) equivalent to a matrix in echelon form.*

Proof. (sketch) We can assume that the first column of A is nonzero because a zero column does not affect elementary row operations. By swapping, it is possible to make $a_{11} \neq 0$. Then, by adding/subtracting multiples $\vec{r}_1$, we can make and $a_{i1} = 0$ for $i > 1$. Now, we apply the reduction method to the submatrix of A with first row and first column removed and then to successively smaller submatrices of A. This process is guaranteed to terminate and will give a matrix in row echelon form. □

One caveat of row echelon form is that it is ***not*** unique. In particular, scaling a row of a matrix in echelon form keeps it in echelon form. Similarly, subtracting a lower from from an upper row still leaves the matrix in row echelon form. We define:

Definition 3. A matrix R is in ***reduced row echelon form*** when

1. R is in echelon form
2. all the pivots are 1.
3. all entries above (and below) the pivots are 0.

Example 2.

$$\left[\begin{array}{cccccc|c} 0 & 1 & 0 & 0 & * & 0 & * \\ 0 & 0 & 1 & 0 & * & 0 & * \\ 0 & 0 & 0 & 1 & * & 0 & * \\ 0 & 0 & 0 & 0 & 0 & 1 & * \end{array}\right]$$

Theorem 4. *Every matrix is (row) equivalent to exactly* **one** *matrix R in reduced row echelon form.*

Thus, we can define $\mathrm{rref}(C) = R$ for $C \in M(m,n)$. The algorithm of Gauss-Jordan elimination allows us to "efficiently" compute $\mathrm{rref}(C)$.

Exercise 1. Show that $\mathrm{rank}(\mathrm{rref}(C))$ is exactly the number of pivots.

Definition 4 (Better version)**.** The number of pivots in $\mathrm{rref}(C)$ is called the ***rank*** of C. If $C = [A \mid \vec{b}]$ corresponds to a system $A\vec{x} = \vec{b}$, then $\mathrm{rref}(C)$ is the ***rank of the system*** and the variables corresponding to pivots are called ***pivot variables***, while the other variables are called ***non-pivot variables***.

Corollary 1. *The rank $r(C)$ of an $m \times n$ matrix C must satisfy $r(C) \leq \min(m,n)$.*

Proof. The number of pivots in $\mathrm{rref}(C)$, which can only occur at most once in each row or column, must be be less than or equal to both m and n. □

Remark. This definition of rank is consistent with our previous definition (i.e. the maximal number of linearly independent rows). Since elementary row operations do not affect the (old) rank of a matrix by Theorem 2, $\mathrm{rank}(A) = \mathrm{rank}(\mathrm{rref}(A))$.

Example 3. The pivot variables in Example 2 are x_2, x_3, x_4, x_6, the non-pivot variables are x_1, x_5, and the rank is 4.

Example 4. Consider the systems $A\vec{x} = \vec{b}$ with augmented matrices:

1. $$[A \mid \vec{b}] = \left[\begin{array}{ccc|c} 0 & 1 & 2 & 2 \\ 1 & -1 & 1 & 5 \\ 2 & 1 & -1 & -2 \end{array}\right]$$

2.

$$[A \mid \vec{b}] = \left[\begin{array}{ccc|c} 0 & 1 & 2 & 2 \\ 1 & -1 & 1 & 5 \\ 2 & -1 & 4 & 3 \end{array}\right]$$

3.

$$[A \mid \vec{b}] = \left[\begin{array}{ccc|c} 0 & 1 & 2 & 2 \\ 1 & -1 & 1 & 5 \\ 2 & -1 & 4 & 12 \end{array}\right]$$

Find the row reduced echelon form of $[A \mid \vec{b}]$, determine the rank of rank(A) and rank$([A \mid \vec{b}])$, and the number of the solutions to the system.

1. We successively apply elementary row operations to $[A \mid \vec{b}]$

$$\left[\begin{array}{ccc|c} 0 & 1 & 2 & 2 \\ 1 & -1 & 1 & 5 \\ 2 & 1 & -1 & -2 \end{array}\right] \xrightarrow{\begin{bmatrix} \vec{r}_2 \\ \vec{r}_1 \\ \vec{r}_3 \end{bmatrix}} \left[\begin{array}{ccc|c} 1 & -1 & 1 & 5 \\ 0 & 1 & 2 & 2 \\ 2 & 1 & -1 & -2 \end{array}\right] \xrightarrow{\begin{bmatrix} \vec{r}_1 \\ \vec{r}_2 \\ \vec{r}_3 - 2\vec{r}_1 \end{bmatrix}} \left[\begin{array}{ccc|c} 1 & -1 & 1 & 5 \\ 0 & 1 & 2 & 2 \\ 0 & 3 & -3 & -12 \end{array}\right] \xrightarrow{\begin{bmatrix} \vec{r}_1 \\ \vec{r}_2 \\ \frac{1}{3}\vec{r}_3 \end{bmatrix}}$$

$$\left[\begin{array}{ccc|c} 1 & -1 & 1 & 5 \\ 0 & 1 & 2 & 2 \\ 0 & 1 & -1 & -4 \end{array}\right] \xrightarrow{\begin{bmatrix} \vec{r}_1 \\ \vec{r}_2 \\ \vec{r}_3 - \vec{r}_2 \end{bmatrix}} \left[\begin{array}{ccc|c} 1 & -1 & 1 & 5 \\ 0 & 1 & 2 & 2 \\ 0 & 0 & -3 & -6 \end{array}\right] \xrightarrow{\begin{bmatrix} \vec{r}_1 \\ \vec{r}_2 \\ -\frac{1}{3}\vec{r}_3 \end{bmatrix}} \left[\begin{array}{ccc|c} 1 & -1 & 1 & 5 \\ 0 & 1 & 2 & 2 \\ 0 & 0 & 1 & 2 \end{array}\right] \xrightarrow{\begin{bmatrix} \vec{r}_1 \\ \vec{r}_2 - 2\vec{r}_3 \\ \vec{r}_3 \end{bmatrix}}$$

$$\left[\begin{array}{ccc|c} 1 & -1 & 1 & 5 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 2 \end{array}\right] \xrightarrow{\begin{bmatrix} \vec{r}_1 - \vec{r}_3 \\ \vec{r}_2 \\ \vec{r}_3 \end{bmatrix}} \left[\begin{array}{ccc|c} 1 & -1 & 0 & 3 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 2 \end{array}\right] \xrightarrow{\begin{bmatrix} \vec{r}_1 + \vec{r}_2 \\ \vec{r}_2 \\ \vec{r}_3 \end{bmatrix}} \boxed{\left[\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 2 \end{array}\right] = \operatorname{rref}([A \mid \vec{b}])}$$

All three variables are pivot variables, there are no non-pivot varibles. There is one solution: $x = 1$, $y = -2$, and $z = 2$. Also, rank$(A) = 3 =$ rank$([A \mid \vec{b}])$.

2.

$$\operatorname{rref}([A \mid \vec{b}]) = \left[\begin{array}{ccc|c} 1 & 0 & 3 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right].$$

Here, x and y are pivot variables, z is a non-pivot variable. The system is inconsistent because of the last row. Note, rank$(A) = 2$ and rank$([A \mid \vec{b}]) = 3$.

3.

$$\operatorname{rref}(B) = \left[\begin{array}{ccc|c} 1 & 0 & 3 & 7 \\ 0 & 1 & 2 & 2 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

Here, x and y are pivot variables, z is a non-pivot variable. For $z \in \mathbb{R}$, we have

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 7 \\ 2 \\ 0 \end{bmatrix} + z \begin{bmatrix} -3 \\ -2 \\ 1 \end{bmatrix}.$$

Solution set is a line starting from $[7, 2, 0]^t$ (***translation vector***) in the direction $[-3, -2, 1]^t$ (***spanning vector***). Note, $\text{rank}(A) = 2$ and $\text{rank}([A \mid \vec{b}\,]) = 2$.

From the above examples, we see that a general solution to $A\vec{x} = \vec{b}$ has the form

$$\vec{x} = T + x_{j_1}X_1 + \cdots + x_{j_k}X_k,$$

where $x_{j_1}, \ldots, x_{j_k}$ are the ***non-pivot*** variables, T is the ***translation vector*** (obtained by setting the non-pivot variables to zero), and $X_1, \ldots, X_k$ are the ***spanning vectors***. We will see more on this in the next lecture.

Also, notice that you can choose any real number for the value of a non-pivot variable and obtain a solution. For this reason, non-pivot variables are examples of ***free valuable***. We say that a variable x_i is ***free*** if for every $t \in \mathbb{R}$ there exists a solution of $A\vec{x} = \vec{b}$ such that $x_i = t$.

Let $r' = \text{rank}(A)$ for the ***coefficient matrix*** A. Observe that

- $r' \leq r$ as $\text{rref}([A \mid \vec{b}\,])$ is obtained from $\text{rref}(A)$ by adding an extra column.
- $r' \leq m$ and $r \leq m$ by Corollary 1.
- $r' \leq n$ and $r \leq n + 1$ by Corollary 1.

Proposition 5. *Let $A\vec{x} = \vec{b}$ be a linear system with $A \in M(m, n)$. Then,*

- *If $r' < r$, then there are no solutions (i.e. the system is* **inconsistent** *and $\vec{b}$ is* **not** *in the span of the columns of A).*
- *If $r' = r < n$, then there are infinitely many solutions.*
- *If $r = r' = n$, then there is* **one** *solution and $\vec{b}$ is a* **unique** *linear combination of the column vectors of A.*

Proof. We do not prove it here, but the observation that $\text{rref}([A \mid \vec{b}\,])$ is obtained from $\text{rref}(A)$ by adding an extra column should make it clear. □

2 Row, Column, and Null Spaces

Recall that a ***real vector space*** is a set together with rules for addition and scalar multiplication by real numbers so that the space is closed under taking linear combinations. Addition is commutative, associative, and an additive identity and inverses exist. Scalar multiplication is associative, distributive, and a scalar multiplicative identity exists.

Some examples of vector spaces were:

- $\mathbb{R}^n$
- A hyperplane in $\mathbb{R}^n$
- $M(n, m)$

Many vector spaces are ***subspaces*** of other vector spaces.

Definition 5. A subset $\mathcal{W}$ of a vector space $\mathcal{V}$ is called a ***subspace*** for any $X, Y \in \mathcal{W}$ and $s \in \mathbb{R}$ one has $X + Y \in \mathcal{W}$ and $s\,X \in \mathcal{W}$. We say that $\mathcal{W}$ is closed under addition and scalar multiplication.

Exercise 2. Show that $\mathcal{W} \subset \mathcal{V}$[1] is a subspace if and only if $s\,X + t\,Y \in \mathcal{W}$ for any $X, Y \in \mathcal{W}$ and $s, t \in \mathbb{R}$.

We don't need to check that all the vector space properties hold for $\mathcal{W}$ — they hold for $\mathcal{V}$, so the laws of addition and scalar multiplication are consistent in W.

The set $\{\mathfrak{o}\}$ consisting of the zero element (i.e. the additive identity) of a vector space $\mathcal{V}$ is a subspace of $\mathcal{V}$ called the ***trivial subspace***. Some examples of subspace of the respective vectors spaces enumerated above are:

- set of vectors in $\mathbb{R}^n$ whose first component is 0 — subspace of $\mathbb{R}^n$
- set of matrices in $M(m, m)$ with zeros on the diagonal — subspace of $M(m, m)$

Since we know that $\mathbb{R}^n, M(n, m)$ are vector spaces, all you have to check in the exercise above is closure under addition and scalar multiplication.

Proposition 6. *Show that* $\operatorname{span}\{X_1, \ldots, X_k\}$ *is a subspace of a vector space* $\mathcal{V}$ *where* $X_i \in \mathcal{V}$ *for* $i = 1, \ldots, k$*.*

[1]The notation $\subset$ mean "subset." So "$\mathcal{W} \subset \mathcal{V}$ is a subspace" reads as "the subset $\mathcal{W}$ of $\mathcal{V}$ is a subspace." In general, subsets are ***not*** subspaces.

Proof. Let $\mathcal{W} = \text{span}\{X_1, \ldots, X_k\}$. Elements $A, B \in \mathcal{W}$ can be written as $A = \sum_{i=1}^{k} a_i\, X_i$ and $B = \sum_{i=1}^{k} b_i\, X_i$ for real number $a_i, b_i \in \mathbb{R}$. For any $s, t \in \mathbb{R}$, we have

$$s\,A + t\,B = \sum_{i=1}^{k} s\,a_i X_i + \sum_{i=1}^{k} t\,b_i X_i = \sum_{i=1}^{k} (s\,a_i + t\,b_i) X_i.$$

Since this is again a linear combination of X_i's, it follows that $s\,A + t\,B \in \mathcal{W}$. Thus, $\mathcal{W}$ is a subspace of $\mathcal{V}$. □

2.1 Subspaces of $\mathbb{R}^m$ arising from systems

Definition 6. Let A be an $m \times n$. Define the ***column space*** $\mathcal{C}(A)$ to be the span of column vectors A. Define the ***row space*** $\mathcal{R}(A)$ to be the span of row vectors of A. Lastly, define the ***nullspace*** $\mathcal{N}(A)$ to be the solution set of $A\vec{x} = \vec{0}$.

Proposition 7. *For $A \in M(m, n)$, the column space is a subspace of $\mathbb{R}^m$ and the row space is a subspaces of $\mathbb{R}^n$.*

Proof. This follows from Proposition 6 as the column and row spaces are spans of sets of vectors. □

Proposition 8. *For $A \in M(m, n)$, nullspace is a subspace of $\mathbb{R}^n$.*

Proof. We can prove this two ways. We could show that the nullspace must be the span of spanning vectors of the system $A\,\vec{x} = \vec{0}$ (see Theorem 10 below), so it is a subspace by Proposition 6. A more nuanced way is to use the following augment.

Let $\mathcal{N} = \{\vec{x} \in \mathbb{R}^n \mid A\,\vec{x} = \vec{0}\}$ denote the nullspace. We must show that for every $\vec{u}, \vec{w} \in \mathcal{N}$ and $s, t \in \mathbb{R}$ we have that $s\vec{u} + t\vec{w} \in \mathcal{N}$. Recall that

$$A\,\vec{x} = x_1\vec{v}_1 + \cdots + x_n\vec{v}_n,$$

where $\vec{v}_i$ are the column vectors of A. So,

$$A(s\vec{u} + t\vec{w}) = (su_1 + tw_1)\vec{v}_1 + \cdots + (su_n + tw_n)\vec{v}_n =$$

$$= s\,(u_1\vec{v}_1 + \cdots + u_n\vec{v}_n,) + t\,(w_1\vec{v}_1 + \cdots + w_n\vec{v}_n) =$$
$$= s\,(A\,\vec{u}) + t\,(A\,\vec{w}) = s \cdot 0 + t \cdot 0 = 0.$$

Above, we heavily used the distributive property of vector spaces and the fact that $\vec{u}, \vec{w} \in \mathcal{N}$. Since we just computed that $A(s\vec{u} + t\vec{w}) = 0$, it follows that $s\vec{u} + t\vec{v} \in \mathcal{N}$. Thus, $\mathcal{N}$ is a subspace (of $\mathbb{R}^n$). □

Let's do a concrete example to help us understand the nullspace a little better.

Example 5.

$$A = \begin{bmatrix} 1 & 2 & 3 & 0 \\ 2 & 4 & 6 & 0 \\ 3 & 2 & 1 & -12 \end{bmatrix}, \qquad \vec{b} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}.$$

$$\text{rref}([A \mid \vec{b}]) = \left[\begin{array}{cccc|c} 1 & 0 & -1 & -6 & 1 \\ 0 & 1 & 2 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right].$$

The pivot variables are x_1 and x_2, and the non-pivot variables are x_3 and x_4. By setting the non-pivot variables to 0, we get the solution $\vec{x}_p = [1\ 0\ 0\ 0]^t$. We called $\vec{x}_p$ the ***translation vector*** (or ***particular solution***) to the system $A\vec{x} = \vec{b}$. Notice that for any other solution $\vec{x}_s$, we have $A(\vec{x}_s - \vec{x}_p) = \vec{0}$. Let $\vec{x}_{null} = \vec{x}_s - \vec{x}_p$, then $\vec{x}_{null}$ lies in null space of A. Thus, $\vec{x}_s = \vec{x}_p + \vec{x}_{null}$. Conversely, ***any*** choice of vector $\vec{x}_{null}$ in the nullspace will give a solution $\vec{x}_p + \vec{x}_{null}$. So, the general solution $\vec{x}$ is given by

$$\vec{x} = \vec{x}_p + \vec{x}_{null}.$$

We call $\vec{x}_{null}$ a ***homogeneous solution***. The system $A\vec{x} = \vec{0}$, of which $\vec{x}_{null}$ is a solution, is called the ***homogenous system*** corresponding to $A\vec{x} = \vec{b}$.

Recall that each non-pivot variable has a corresponding spanning vector, which we can compute from $\text{rref}([A \mid \vec{b}])$. The spanning vector for x_3 is $[1\ -2\ 1\ 0]^t$ and for x_4 we have $[6\ -3\ 0\ 1]^t$. Since every solution to $A\vec{x} = \vec{b}$ has the form (translation vector) + (linear combination of spanning vectors), we see that $\mathcal{N}$ has to be spanned by the spanning vectors. So, in this example, the nullspace is

$$\mathcal{N} = \text{span}\left\{ \begin{bmatrix} 1 \\ -2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 6 \\ -3 \\ 0 \\ 1 \end{bmatrix} \right\}.$$

We can see that $\mathcal{N}$ is a "2-plane" in $\mathbb{R}^4$ and the solution set of $A\vec{x} = \vec{b}$ corresponds to ***translating*** $\mathcal{N}$ by the translation vector $\vec{x}_p = [1\ 0\ 0\ 0]^t$.

Proposition 9. *A* **consistent** *system $A\vec{x} = \vec{b}$ has a unique solution if and only if the nullspace of A is $\{\vec{0}\}$.*

Proof. Since A is consistent, it has at least one solution $\vec{x}_1$. Let $\vec{x}_2$ be any other solution. Then $A(\vec{x}_1 - \vec{x}_2) = A\vec{x}_1 - A\vec{x}_2 = \vec{b} - \vec{b} = \vec{0}$, so $\vec{x}_1 - \vec{x}_2$ is in the nullspace of A. Therefore, if the nullspace of A is $\vec{0}$, then $\vec{x}_1 - \vec{x}_2 = \vec{0}$, so $\vec{x}_1 = \vec{x}_2$ and there can only be one solution. Conversely, if the null space of A is ***not*** $\{\vec{0}\}$, then it must contain some vector $\vec{x}_{null} \neq \vec{0}$. However, this gives $\vec{x}_3 = \vec{x}_1 + \vec{x}_{null}$ as a solution to $A\vec{x} = \vec{b}$ and $\vec{x}_3 \neq \vec{x}_1$. □

The behavior we see above happens in general:

- $\vec{x}_p$ is obtained by setting all the non-pivot variables to 0.
- the spanning vectors (i.e. the vectors whose coefficients are non-pivot variables) span the nullspace of A.

Theorem 10. *Let A be an $m \times n$ matrix. The spanning vectors of the system $A\vec{x} = 0$ span $\mathcal{N}(A)$ and are linearly independent.*

Proof. Since the last column of $\mathrm{rref}([A \mid \vec{0}\,])$ is all zeros, we see that the translation vector for the system $A\vec{x} = \vec{0}$ is $\vec{0}$. Let $x_{j_1}, \ldots, x_{j_k}$ be the non-pivot variables of $\mathrm{rref}([A \mid \vec{b}\,])$. By setting $x_{j_i} = 1$ and the other non-pivot variables to 0, we can compute the spanning vector X_i for x_{j_i} from the reduced row echelon form. We obtain that a solution of $\mathrm{rref}([A \mid \vec{b}\,])$, and therefore $A\vec{x} = \vec{0}$ (since row reduction preserves solution sets), must have the form

$$\vec{x} = x_{j_1} X_1 + \cdots + x_{j_k} X_k.$$

Thus, we see that the null space of A is $\mathrm{span}\{X_1, \ldots, X_k\}$.

It remains to show linear independence. Notice that the j_i^{th} entry of X_i is always 1, since it corresponds to $x_{j_i} = 1$. Additionally, the j_i^{th} entry of X_t for $t \neq i$ is always 0. Focusing on the entries corresponding to the non-pivot variables, we see that there cannot be a nontrivial linear combination between the X_i's. $\square$

2.2 Orthogonality in $\mathbb{R}^n$

Definition 7. Recall that the ***dot product*** of $\vec{u} = [u_1\ u_2\ \cdots\ u_n]^t \in \mathbb{R}^n$ and $\vec{v} = [v_1\ v_2\ \cdots\ v_n]^t \in \mathbb{R}^n$ is

$$\vec{u} \cdot \vec{v} = \vec{u}^{\,t}\ \vec{v} = [u_1\ u_2\ \cdots\ u_n] \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = \sum_{i=1}^{n} u_i\, v_i.$$

Recall the dot product is distributive over vector addition and scalar multiplication.

Note. The dot product is special to $\mathbb{R}^n$, other vector spaces may not have one!

Definition 8. Suppose $\mathcal{W}$ is a subspace of $\mathbb{R}^n$. We define the ***orthogonal complement*** of $\mathcal{W}$ by

$$\mathcal{W}^{\perp} = \{\vec{u} \in \mathbb{R}^n \mid \vec{u} \cdot \vec{v} = 0 \text{ for all } \vec{v} \in \mathcal{W}\}$$

Proposition 11. *Let $\mathcal{W}$ be a subspace of $\mathbb{R}^n$, then $\mathcal{W}^{\perp}$ is also subspace of $\mathbb{R}^n$*

Proof. Let us check that $\mathcal{W}^\perp$ satisfies the definition of subspace. For $\vec{u}, \vec{w} \in \mathcal{W}^\perp$ and $a, b \in \mathbb{R}$, let $\vec{y} = a\vec{u} + b\vec{w}$. We must show that $\vec{y} \in \mathcal{W}^\perp$. For any $\vec{v} \in \mathcal{W}$, we have that

$$\vec{y} \cdot \vec{v} = (a\vec{u} + b\vec{w}) \cdot \vec{v} \xlongequal[\text{dist. prop.}]{\text{dot prod.}} a(\vec{u} \cdot \vec{v}) + b(\vec{w} \cdot \vec{v}) = a\,0 + b\,0 = 0.$$

Thus, $\vec{y} \in \mathcal{W}$ by definition and $\mathcal{W}$ is a subspace of $\mathbb{R}^n$. □

Proposition 12. *Let $\mathcal{R}$ be the row space of A and $\mathcal{N}$ its nullspace. Then,*

$$\mathcal{N} = \mathcal{R}^\perp$$

Proof. Let $\vec{r}_1, \cdots, \vec{r}_m$ be the rows of A. Since $\vec{x} \in \mathcal{N}$ if and only if $A\,\vec{x} = \vec{0}$, it follows that $\vec{x} \in \mathcal{N}$ if and only if $\vec{r}_i \cdot \vec{x} = 0$ for all i. Thus, $\vec{x} \in \mathcal{R}^\perp$ if and only if $\vec{x} \in \mathcal{N}$. □

We will show later that that $\mathcal{R} = \mathcal{N}^\perp$ (i.e. $(\mathcal{R}^\perp)^\perp = \mathcal{R}$)

2.3 Brief summary of the content so far

We have introduced the notions of vector space $\mathcal{V}$ and (several definitions of) linear independence. The most useful one is that $S = \{X_1, \ldots, X_n\} \subset \mathcal{V}$ is linearly independent if $\mathsf{o} = c_1 X_1 + \cdots + c_n X_n$ for some $c_i \in \mathbb{R}$ implies that $c_i = 0$ for all i. In general, to test linear independence we turn the expression $\mathsf{o} = c_1 X_1 + \cdots + c_n X_n$ into a linear system of equations $A\,\vec{x} = \vec{0}$ where $A \in M(m,n)$ and $x_i = c_i$ are our unknowns. The set S is then linearly independent if the only solution is $\vec{x} = \vec{0}$.

Our method for solving $A\,\vec{x} = \vec{0}$ (and, more generally, $A\vec{x} = \vec{b}$) is to use row reduction to obtain a unique reduced row echelon form $\mathrm{rref}([A \mid \vec{b}\,])$. Here, the pivots of $\mathrm{rref}([A \mid \vec{b}\,])$ give us an easy way to (1) compute the rank, (2) analyze the number of solutions and (3) show that the solution set has the form: (transition vector) + (span of spanning vectors).

The span of the spanning vectors turned out to be the solution set of the homogenous system $A\,\vec{x} = \vec{0}$. This solution set, called the nullspace $\mathcal{N}(A)$, has the property that it is a vector subspace of $\mathbb{R}^n$. We introduced two other important vector subspaces: the columns space $\mathcal{C}(A) \subset \mathbb{R}^m$ and the row space $\mathcal{R}(A) \subset \mathbb{R}^n$. Lastly, we saw that $\mathcal{N}(A) = \mathcal{R}(A)^\perp$.

2.4 Basis

Previously, we showed that the spanning vectors of a linear system are linearly independent. They form a special type of set called a ***basis*** of $\mathcal{N}(A)$.

Definition 9. A set $S \subset \mathcal{V}$ is a ***basis*** of vector space $\mathcal{V}$ if $\mathrm{span}(S) = \mathcal{V}$ ***and*** S is linearly independent.

Example 6. $S = \{\hat{e}_1, \ldots, \hat{e}_n\}$ is a (standard) basis for $\mathbb{R}^n$.

Example 7. The spanning vectors for the system $A\vec{x} = \vec{b}$ are a basis for $\mathcal{N}(A)$.

Proposition 13. *The pivot rows of* $\text{rref}(A)$ *form a basis for* $\mathcal{R}(A)$.

Proof. Since the rows of $\text{rref}(A)$ are linear combinations of the rows of $\mathcal{R}(A)$ and ***vice versa***, we see that they have the same span[2]. Further, we known that the rows of $\text{rref}(A)$ are either pivot rows or all zero. Since zero rows don't contribute to the span, we can throw the out. Lastly, the pivot rows of $\text{rref}(A)$ must be linearly independent since each pivot appears in a column where all other entries are zero. Thus, they form a basis for $\mathcal{R}(A)$. □

We have a somewhat more difficult result about $\mathcal{C}(A)$.

Proposition 14. *The columns of A corresponding to pivots (variables) of* $\text{rref}(A)$ *form a basis for $\mathcal{C}(A)$. There columns are called* **pivot columns** *of A.*

Proof. Let $\vec{v}_1, \ldots, \vec{v}_n$ be the columns of A. First, we show that the pivot columns span $\mathcal{C}(A)$. To do this, we need to show that all ***nonpivot*** columns are linear combinations of the ***pivot*** columns. Let $x_{a_1}, \ldots, x_{a_k}$ be the nonpivot variables A and let $x_{b_1}, \ldots, x_{b_t}$ be the pivot variables of A. For a given $i = 1, \ldots, k$, let $\vec{c}$ be a solution to $A\vec{x} = \vec{0}$ such that $x_{a_i} = -1$ and $x_{a_j} = 0$ for $j \neq i$[3]. Then,

$$\vec{0} = A\vec{c} = -\vec{v}_{a_i} + \sum_{j=1}^{t} c_{b_j}\, \vec{v}_{b_j},$$

which means that the nonpivot column $\vec{v}_{a_i}$ is a linear combination of pivot columns. As we can do this for all $i = 1, \ldots, k$, we see that $\mathcal{C}(A) = \text{span}\{\vec{v}_{b_1}, \ldots, \vec{v}_{b_t}\}$.

To show that $\{\vec{v}_{b_1}, \ldots, \vec{v}_{b_t}\}$ is linearly independent, notice that any expression

$$\vec{0} = r_1\, \vec{v}_{b_1} + \cdots + r_t\, \vec{v}_{b_t}$$

can be turned into a solution to $A\,\vec{x} = 0$ by letting $x_{b_j} = r_j$ for $j = 1, \ldots, t$ and zero otherwise. However, this solution has all the nonpivot variables set to zero, but the only such solution is, in fact, $\vec{0}$. So $r_j = 0$ for all j and $\{\vec{v}_{b_1}, \ldots, \vec{v}_{b_t}\}$ is linearly independent. This means that the pivot columns of A are a basis for $\mathcal{C}(A)$. □

Example 8. Let

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 & 0 \\ 1 & 2 & 7 & 8 & 4 \\ 1 & 2 & 11 & 12 & 8 \end{bmatrix} \Longrightarrow \text{rref}(A) = \begin{bmatrix} 1 & 2 & 0 & 1 & -3 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

[2] That is, any linear combination of rows of A can be turned into a linear combination of the rows of $\text{rref}(A)$ and vice versa.

[3] We can do this since the nonpivot variables are free.

Then,

$$\left\{\begin{bmatrix}-2\\1\\0\\0\\0\end{bmatrix}, \begin{bmatrix}-1\\0\\-1\\1\\0\end{bmatrix}, \begin{bmatrix}3\\0\\-1\\0\\1\end{bmatrix}\right\} \text{ is a basis for } \mathcal{N}(A)$$

$$\left\{\begin{bmatrix}1 & 2 & 0 & 1 & -3\end{bmatrix}, \begin{bmatrix}0 & 0 & 1 & 1\end{bmatrix}\right\} \text{ is a basis for } \mathcal{R}(A)$$

$$\left\{\begin{bmatrix}1\\1\\1\end{bmatrix}, \begin{bmatrix}3\\7\\11\end{bmatrix}\right\} \text{ is a basis for } \mathcal{C}(A)$$

Note that the nonpivot columns are linear combinations of the pivot columns. Also, the basis vectors for $\mathcal{N}(A)$ are perpendicular to both basis vectors for $\mathcal{R}(A)$.

Let A be an $m \times n$ matrix. We just saw how to compute bases for $\mathcal{R}(A), \mathcal{N}(A)$, and $\mathcal{C}(A)$. Letting $k = \text{rank}(A)$, we can summarize out findings as

Space	Dimension	Basis
$\mathcal{R}(A)$	k	pivot row vectors of rref(A)
$\mathcal{N}(A)$	$n - k$	spanning vectors computed from rref(A)
$\mathcal{C}(A)$	k	columns of A corresponding to pivot columns of rref(A)

All that hard work from before pays off in making the following results clear.

Theorem 15. *[Rank] For any matrix A, we have that* $\dim \mathcal{R}(A) = \dim \mathcal{C}(A)$ *and this dimension is* $\text{rank}(A)$.

Theorem 16. *[Rank-Nullity] For $A \in M(m,n)$, define* $\text{null}(A) = \dim \mathcal{N}(A)$. *Then,*

$$\text{rank}(A) + \text{null}(A) = n.$$

The dimension of the nullspace $\text{null}(A)$ is called the ***nullity*** of A.

Proposition 17. *Let A be an $m \times n$ matrix, then*

1. *the columns of A are linearly independent if and only if* $\text{null}(A) = 0$.
2. *the columns of A are linearly independent if and only if* $\text{rank}(A) = n$.
3. *the rows of A are linearly independent if and only if* $\text{rank}(A) = m$.

Recall that the ***transpose*** A^t of A is obtained by "flipping" A along the diagonal. That is, we take the rows of A and make them the columns of A^t. Notice that the columns of A become rows of A^t. In particular, we see that $\mathcal{R}(A^t) = \mathcal{C}(A)$ and $\mathcal{C}(A^t) = \mathcal{R}(A)$. From this, it is clear that

Proposition 18. *Let A be an $m \times n$ matrix, then* $\text{rank}(A^t) = \text{rank}(A)$.