3.8 Rank of a Matrix

Definition 3.8.1. Let A be an $m \times n$ matrix with entries from a field F. Define:

- The *Row Space of* A *(RSp(A))* as the span of the rows of A. This is a subspace of F^n.
- The *Row Rank of* A is $dim(RSp(A))$.
- The *Column Space of* A *(CSp(A))* as the span of the columns of A. This is a subspace of F^m.
- The *Column Rank of* A is $dim(CSp(A))$.

Example 3.8.2. Let $F = \mathbb{R}$ and $A = \begin{pmatrix} 3 & 1 & 2 \\ 0 & -1 & 1 \end{pmatrix}$. Then,

$$\begin{aligned} RSp(A) &= Span\{(3 \quad 1 \quad 2), (0 \quad -1 \quad 1)\}, \\ CSp(A) &= Span\left\{\begin{pmatrix} 3 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \end{pmatrix}\right\}. \end{aligned}$$

Now the row vectors $(3 \quad 1 \quad 2)$ and $(0 \quad -1 \quad 1)$ are linearly independent so $dim(RSp(A)) = 2$, so the column rank is 2. The set

$$\left\{\begin{pmatrix} 3 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \end{pmatrix}\right\}$$

is linearly dependent as

$$\begin{pmatrix} 3 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix} + \begin{pmatrix} 2 \\ 1 \end{pmatrix}.$$

So

$$CSp(A) = Span\{\left\{\begin{pmatrix} 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \end{pmatrix}\right\},$$

which is linearly independent, so $dimCSp(A) = 2$.

Procedure 3.8.3.

Calculating the row rank of a matrix A.

- Step 1: Reduce A to row echelon form using row operations:

$$A_{ech} = \begin{pmatrix} 1 & * & * & * & * & \dots \\ 0 & 0 & 1 & * & * & \dots \\ 0 & 0 & 0 & 1 & * & *\dots \\ \vdots & & & & & \\ 0 & \dots & & & & \end{pmatrix}$$

(Actually it doesn't matter whether the leading entries in each row are 1s or not.)

- Step 2: The row rank of A is the number of non-zero rows in A_{ech}. In fact it the non-zero rows of A_{ech} form a basis for $RSp(A)$.

Justification

It will be enough to show:

1. $RSp(A) = RSP(A_{ech})$
2. The rows of A_{ech} are linearly independent.

To show 1., not that to obtain A_{ech} from A we use row operations:

$$\begin{cases} r_i \mapsto r_i + \lambda r_j & \lambda \in F, \quad i \neq j \\ r_i \mapsto \lambda r_i & \lambda \in F \setminus \{0\} \\ r_i \mapsto r_j & i \neq j \end{cases}$$

Let A' be obtained from A by one row operation. then clearly every row of A' lies in $RSp(A)$ and so $RSp(A') \subseteq RSp(A')$. Also every row operation is invertibl;e by another row operation:

$$\begin{cases} r_i \mapsto r_i + \lambda r_j & \text{has inverse} & r_i \mapsto r_i - \lambda r_j \\ r_i \mapsto \lambda r_i & \text{has inverse} & r_i \mapsto \frac{1}{\lambda} r_i \\ r_i \mapsto r_j & \text{has inverse} & r_i \mapsto r_j \end{cases}$$

It follwos that A is obtained from A' by row operations, so $RSp(A) \subseteq RSp(A')$. Hence $RSp(A) = RSp(A')$.

In other words row operations have no effect on the row space. In particular $RSp(A) = RSp(A_{ech})$.

For 2. let $i_1, ..., i_k$ be the numbers of the columns of A_{ech} containing the leading entries:

$$A_{ech} = \begin{array}{c} \begin{pmatrix} 1 & * & * & * & * & ... \\ 0 & 0 & 1 & * & * & ... \\ 0 & 0 & 0 & 1 & * & *... \\ \vdots & & & & & \\ 0 & \dots & & & & \end{pmatrix} \\ \begin{array}{cccccc} i_1 & & i_2 & i_3 & ... & \end{array} \end{array}$$

Let $r_1, ...r_k$ are the rows of A_{ech}. Suppose $\lambda_1 r_1 + ... + \lambda_k r_k = 0$ for scalars λ_i. We see that the i_1^{th} entry of $\lambda_1 r_1 + ... + \lambda_k r_k$ is $\lambda_1 \cdot 1 = \lambda_1$ hence $\lambda_1 = 0$. Therefore $\lambda_1 r_1 + ... + \lambda_k r_k = \lambda_2 r_2 + ... + \lambda_k r_k$, similarly the I_2^{th} entry of $\lambda_2 r_2 + ... + \lambda_k r_k$ is λ_2, so $\lambda_2 = 0$. By induction we can show that $a_i = 0$ for all i. So $\{r_1, ..., r_k\}$ is linearly independent.

Example 3.8.4. Find the row rank of $A = \begin{pmatrix} 1 & 2 & 5 \\ 2 & 1 & 0 \\ -1 & 4 & 15 \end{pmatrix}$

Answer:

$$A \mapsto \begin{pmatrix} 1 & 2 & 5 \\ 0 & -3 & -10 \\ 0 & 6 & 20 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 2 & 5 \\ 0 & 1 & \frac{10}{3} \\ 0 & 0 & 0 \end{pmatrix} = A_{ech}$$

A_{ech} has 2 non-zero rows, so the row rank of A is 2.

Example 3.8.5. Find the dimension of

$$W = Span\{\begin{pmatrix} -1 & 1 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 2 & 3 & 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 2 & 3 \end{pmatrix}\} \subseteq \mathbb{R}^4.$$

Answer

We can work this out by seeing our vectors as the rows of a matrix:

Let $A = \begin{pmatrix} -1 & 1 & 0 & 1 \\ 2 & 3 & 1 & 0 \\ 0 & 1 & 2 & 3 \end{pmatrix}$. The span we want is the row span of this matrix, which we work out:

$$\begin{aligned} A &\mapsto \begin{pmatrix} -1 & 1 & 0 & 1 \\ 0 & 5 & 1 & 2 \\ 0 & 1 & 2 & 3 \end{pmatrix} \mapsto \begin{pmatrix} -1 & 1 & 0 & 1 \\ 0 & 5 & 1 & 2 \\ 0 & 5 & 10 & 15 \end{pmatrix} \\ &\mapsto \begin{pmatrix} -1 & 1 & 0 & 1 \\ 0 & 5 & 1 & 2 \\ 0 & 0 & 9 & 13 \end{pmatrix} = A_{ech} \end{aligned}$$

A_{ech} has 3 non-zero rows so $RSp(A)$ has dimension 3. So $sim(W) = 3$.

We can find the column rank of a matrix in a very similar way to finding the row rank of a matrix.

Procedure 3.8.6. The columns of A are the rows of A^T so we can apply Procedure 3.8.3 to A^T.

Alternatively: use column operations to resuce A to "column echelon form and then count the non-zero columns.

Example 3.8.7. Let $A = \begin{pmatrix} 1 & 2 & 5 \\ 2 & 1 & 0 \\ -1 & 4 & 15 \end{pmatrix}$. Find the column rank of A. This equals the row rank of A^T.

$$A^T = \begin{pmatrix} 1 & 2 & -1 \\ 2 & 1 & 4 \\ 5 & 0 & 15 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 2 & -1 \\ 0 & -3 & 6 \\ 0 & -10 & 20 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 2 & -1 \\ 0 & -3 & 6 \\ 0 & 0 & 0 \end{pmatrix} = A^T_{ech}$$

So the column rank of A is 2. A basis for $RSp(A^T)$ is $\{\begin{pmatrix} 1 & 2 & -1 \end{pmatrix}, \begin{pmatrix} 0 & -3 & 6 \end{pmatrix}\}$. So a basis for $CSp(A)$ is $\left\{ \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ -3 \\ 6 \end{pmatrix} \right\}$

Theorem 3.8.8. For any matrix A the row rank of A is equal to the column rank of A.

Proof:

Let $A = (a_{ij}) \in M_{m\times n}(F)$. Let the rows of A be $r_1, ..., r_m$, so $r_i = (a_{i1}, ..., a_{in})$.Let the columns of A be $c_1, ..., c_n$, so $c_j = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix}$.

Let k be the row rank of A. Then $RSp(A)$ has a basis $\{v_1, ..., v_k\}$. Every row r_i is a linear combination of $v_1, ..., v_k$. Say:

$$r_i = \lambda_{i1}v_1 + ... + \lambda_{ik}v_k(\dagger)$$

Suppose that $v_i = (b_{i1}, b_{i2}, ..., b_{in})$ then looking at the j^{th} coordinate in $(\dagger)$ we get:

$$a_{ij} = \lambda_{i1}b_{1j} + \lambda_{i2}b_{2j} + ... + \lambda_{ik}b_{kj}$$

Now

$$\begin{aligned} c_j &= \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix} = \begin{pmatrix} \lambda_{11}b_{1j} + \lambda_{12}b_{2j} + ... + \lambda_{1k}b_{kj} \\ \lambda_{21}b_{1j} + \lambda_{22}b_{2j} + ... + \lambda_{2k}b_{kj} \vdots \\ \lambda_{m1}b_{1j} + \lambda_{m2}b_{2j} + ... + \lambda_{mk}b_{kj} \end{pmatrix} \\ &= \begin{pmatrix} \lambda_{11} \\ \vdots \\ \lambda_{m1} \end{pmatrix} b_{1j} + \begin{pmatrix} \lambda_{12} \\ \vdots \\ \lambda_{m2} \end{pmatrix} b_{2j} + ... + \begin{pmatrix} \lambda_{1k} \\ \vdots \\ \lambda_{mk} \end{pmatrix} b_{kj} \end{aligned}$$

So c_j is a linear combination of the vectors:

$$\begin{pmatrix} \lambda_{11} \\ \vdots \\ \lambda_{m1} \end{pmatrix}, \begin{pmatrix} \lambda_{12} \\ \vdots \\ \lambda_{m2} \end{pmatrix}, ..., \begin{pmatrix} \lambda_{1k} \\ \vdots \\ \lambda_{mk} \end{pmatrix}$$

Hence $CSp(A)$ is spanned by these vectors, thus $sim(CSp(A) \leq k = dim(RSP(A))$. Equally the column rank of A^T is at most the row rank of A^T (by the same argument). The column rank of A^T is the row rank of A, and the row rank of A^T is the Column rank of A. Thus we have $dim(RSp(A)) \leq dim(CSp(A))$, and hence $dim(RSp(A)) = dim(CSp(A))$.

Example 3.8.9. Let $A = \begin{pmatrix} 1 & 2 & -1 & 0 \\ -1 & 1 & 0 & 1 \\ 0 & 3 & -1 & 1 \end{pmatrix}$

Note that $r_3 = r_1 + r_2$, so a basis for $RSp(A)$ is

$$\{\underbrace{(1,2,-1,0)}_{v_1} \quad , \quad \underbrace{(-1,1,0,1)}_{v_2}\}$$

Write the rows as linear combinations of v_1 and v_2:

$$\begin{aligned} r_1 &= 1v_1 + 0v_2 \\ r_2 &= 0v_1 + 1v_2 \\ r_3 &= 1v_1 + 1v_2 \end{aligned}$$

These co-efficients are the λ_{ij}'s from the proof:

$$\begin{array}{ll} \lambda_{11} = 1 & \lambda_{12} = 0 \\ \lambda_{21} = 0 & \lambda_{22} = 1 \\ \lambda_{31} = 1 & \lambda_{32} = 1 \end{array}$$

According to the proof, a spanning set for $CSp(A)$ is:

$$\begin{pmatrix} \lambda_{11} \\ \lambda_{21} \\ \lambda_{31} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} \lambda_{11} \\ \lambda_{21} \\ \lambda_{31} \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$$

Check this is really a spanning set for $CSP(A)$: Let $w_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$, $w_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$

Now we have:

$$\begin{aligned} c_1 &= \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} &= w_1 - w_2 \\ c_2 &= \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix} &= 2w_1 + w_2 \\ c_3 &= \begin{pmatrix} -1 \\ 0 \\ -1 \end{pmatrix} &= -w_1 \\ c_4 &= \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} &= w_2 \end{aligned}$$

So it is indeed the case that $\{w_1, w_2\}$ spans $CSp(A)$.

Definition 3.8.10. Let A be a matrix. The *rank of* A written $rank(A)$ or $rk(A)$, is the row rank of A (or the column rank since they are the same).

Proposition 3.8.11. Let A be an $n \times n$ matrix with entried in F, then the following statements are equivalent:

1. $rank(A) = n$ ("A has full rank").
2. The rows of A form a basis for F^n.
3. The columns of A form a basis for F^n.
4. A is invertible (so $det(A) \neq 0$, etc.).

Proof:

- (1) $\Leftrightarrow$ (2):

$$\begin{aligned} rank(A) = n &\Leftrightarrow dim(RSp(A)) = n \\ &\Leftrightarrow RSp(A) = F^n \\ &\Leftrightarrow \text{the rows of } A \text{ form a basis for } F^n \end{aligned}$$

- (1) $\Leftrightarrow$ (3): The same, but with columns.
- (1) $\Leftrightarrow$ (4): $rank(A) = n$ if and only if $A_{ech} = \begin{pmatrix} 1 & & & & \\ & 1 & & * & \\ & & 1 & & \\ & 0 & & \ddots & \\ & & & & 1 \end{pmatrix}$

Now all of the $*$ entries can be eliminated using row operations and so A is reducible to Id using row operations. By 2.6.2 this is equivalent to A being invertible.

4 Linear Transformations

4.1 Introduction

Definition 4.1.1. Suppose V, W are vector spaces over a field F. Let $T : V \longrightarrow W$ be a function from V to W. We say:

- T *preserves addition* if for all $v_1, v_2 \in V$ we have $T(v_1 + v_2) = T(v_1) + T(v_2)$. (i.e. if $T(v_1) = w_1, T(v_2) = w_2$ for $w_1, w_2 \in W$ we have $T(v_1 + v_2) = w_1 + w_2$.
- T *preserves scalar multiplication* if for all $v \in V$, $\lambda \in F$, $T(\lambda v) = \lambda T(v)$.
- T is a *linear transformation* (or *linear map*) if it:
  1. preserves addition.
  2. preserves scalar multiplication

Example 4.1.2.

(a) The identity map $T : V \longrightarrow V$ is obviously a linear transformation.

(b) $T : \mathbb{R}^2 \longrightarrow \mathbb{R}$ defined by $T(x, y) = x + y$ is a linear transformation.
Check:

- $T((x_1, y_1) + (x_2, y_2)) = T((x_1 + x_2, y_1 + y_2)) = x_1 + x_2 + y_1 + y_2 = (x_1 + y_1) + (x_2 + y_2) = T((x_1, y_1)) + T((x_2, y_2))$ So T preserves addition.
- Let $\lambda \in \mathbb{R}$ then $T(\lambda(x, y)) = T((\lambda x, \lambda y)) = \lambda x + \lambda y = \lambda T((x, y))$. So T preserves scalar multiplication.

(c) Let V be the space of all polynomials in x over $\mathbb{R}$ (i.e. $V = \mathbb{R}[x]$). Defin $T : V \longrightarrow V$ by $T(f(x)) = \frac{d}{dx} f(x)$. Then T is a linear map.
Check:

- $T(f(x) + g(x)) = \frac{d}{dx}(f(x) + g(x)) = \frac{d}{dx}(f(x)) + \frac{d}{dx}(g(x)) = T(f(x)) + T(g(x))$ So T preserves addition.
- Let $\lambda \in \mathbb{R}$ then $T(\lambda f(x)) = \frac{d}{dx}\lambda f(x) = \lambda \frac{d}{dx} f(x) = \lambda T(f(x))$. So T preserves scalar multiplication.

(d) Let $V = \mathbb{C}$ (as a 1-dimensional vector space over $\mathbb{C}$). The map $T(z) = \bar{z}$ is ***not*** a linear map:

- $T(z_1 + z_2) = \overline{z_1 + z_2} = \bar{z_1} + \bar{z_2} = T(z_1) + T(z_2)$ So T *does* preserve addition.
- $T(\lambda z) = \overline{\lambda z} = \bar{\lambda}\bar{z} \neq \lambda \bar{z} = \lambda T(z)$ for $\lambda \notin \mathbb{R}$. So T *does **not*** preserve scalar multiplication.

(e) Let $T : \mathbb{R}^3 \longrightarrow \mathbb{R}$ be given by $T(x, y, z) = (xyz)^{\frac{1}{3}}$ then:

- $T(\lambda(x, y)) = T((\lambda x, \lambda y)) = (\lambda^3 xyz)^{\frac{1}{3}} = \lambda T((x, y, z))$. So T preserves scalar multiplication.
- $T((x_1, y_1, z_1) + (x_2, y_2, z_2)) = T((x_1 + x_2, y_1 + y_2, z_1 + z_2)) = ((x_1 + x_2)(y_1 + y_2)(z_1 + z_2))^{\frac{1}{3}} \neq ((x_1 + y_1 + z_1)^{\frac{1}{3}} + (x_2 + y_2 + z_2))^{\frac{1}{3}} = T((x_1, y_1, z_1)) + T((x_2, y_2, z_2))$. So T *does **not*** preserve addition.

(f) Lots of functions preserve neither addition nor scalar multiplication, e.g., for $\mathbb{R} \to \mathbb{R}$ the functions taking $x \mapsto x+1$, $x \mapsto x^2$, and $x \mapsto e^x$.

Proposition 4.1.3. Let A be an $m \times n$ matrix over F. Define $T : F^n \longrightarrow F^m$ (spaces of column vectors), by $T(v) = Av$ (for $v \in F^n$. Then T is a linear transformation.
Proof: We need to check:

- Preserves addition: Let $v_1, v_2 \in F^n$

$$T(v_1 + v_2) = A(v_1 + v_2) = Av_1 + Av_2 = T(v_1) + T(v_2) \quad \text{by M1GLA}$$

- Preserves scalar multiplicaiton: Let $v \in V$, $\lambda \in F$ then:

$$T(\lambda v) = A(\lambda v) = \lambda Av = \lambda T(v)$$

Proposition 4.1.4. *Basic Properties of linear transformations*

Let $T : V \longrightarrow W$ be a linear map. Write $0_V, 0_W$ for the zero vectors in V and W respectively. We have:

1. $T(0_v) = 0_W$
2. Suppose $v = \lambda_1 v_1 + ... + \lambda_k v_k$ for $\lambda_i \in F$, $v_i \in V$. Then $T(v) = \lambda_1 T(v_1) + ... + \lambda_k T(v_k)$.

Proof:

1. Since T preserves scalar multiplication we have $T(\lambda 0_v) = \lambda T(0_v)$ for $\lambda \in F$. Taking $\lambda = 0$, we have $T(00_v) = 0T(0_v)$, but $0 \cdot 0_v = 0_v$ and $0 \cdot T(0_v) = 0_W$. Hence $T(0_v) = 0_W$.
2. Induction on k.
   Base case. The case where $k = 1$ just says T preserves scalar multiplication, so it true.

   Inductive step: Suppose we know $T(\lambda_1 v_1 + ... + \lambda_{k-1} v_{k-1}) = \lambda_1 T(v_1) + ... + \lambda_{k-1} T(v_{k-1})$. Now
$$\begin{aligned} T(\lambda_1 v_1 + ... + \lambda_k v_k) &= T(\lambda_1 v_1 + ... + \lambda_{k-1} v_{k-1}) + T(\lambda_k v_k) \\ &= T(\lambda_1 v_1 + ... + \lambda_{k-1} v_{k-1}) + \lambda_k T(v_k) \\ &= T(\lambda_1 v_1 + ... + \lambda_{k-1} v_{k-1}) + \lambda_k T(v_k) \end{aligned}$$