

## Chapter 4

# Singular Value Decomposition

In order to solve linear systems with a general rectangular coefficient matrix, we introduce the singular value decomposition. It is one of the most important tools in numerical linear algebra, because it contains a lot of information about a matrix, including: rank, distance to singularity, column space, row space, and null spaces.

**Definition 4.1 (SVD).** *If  $A \in \mathbb{C}^{m \times n}$ . If  $m \geq n$  then a singular value decomposition (SVD) of  $A$  is a decomposition*

$$A = U \begin{pmatrix} \Sigma \\ 0 \end{pmatrix} V^*, \quad \text{where } \Sigma = \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \end{pmatrix}, \quad \sigma_1 \geq \dots \geq \sigma_n \geq 0,$$

and  $U \in \mathbb{C}^{m \times m}$  and  $V \in \mathbb{C}^{n \times n}$  are unitary.

*If  $m \leq n$  then a SVD of  $A$  is*

$$A = U \begin{pmatrix} \Sigma & 0 \end{pmatrix} V^*, \quad \text{where } \Sigma = \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_m \end{pmatrix}, \quad \sigma_1 \geq \dots \geq \sigma_m \geq 0,$$

and  $U \in \mathbb{C}^{m \times m}$  and  $V \in \mathbb{C}^{n \times n}$  are unitary.

*The matrix  $U$  is called a left singular vector matrix,  $V$  is called a right singular vector matrix, and the scalars  $\sigma_j$  are called singular values.*

### Remark 4.2.

- $A m \times n$  matrix has  $\min\{m, n\}$  singular values.
- The singular values are unique, but the singular vector matrices are not. Although a SVD is not unique, one often says “the SVD” instead of “a SVD”.
- Let  $A \in \mathbb{C}^{m \times n}$  with  $m \geq n$ . If  $A = U \begin{pmatrix} \Sigma \\ 0 \end{pmatrix} V^*$  is a SVD of  $A$  then  $A^* =$

$V(\Sigma \ 0)U^*$  is a SVD of  $A^*$ . Therefore  $A$  and  $A^*$  have the same singular values.

- $A \in \mathbb{C}^{n \times n}$  is nonsingular if and only if all singular values are nonzero, i.e.  $\sigma_j > 0$ ,  $1 \leq j \leq n$ .  
If  $A = U\Sigma V^*$  is a SVD of  $A$  then  $A^{-1} = V\Sigma^{-1}U^*$  is a SVD of  $A^{-1}$ .

**Example 4.3** The  $2 \times 2$  matrix

$$A = \begin{pmatrix} 1 & \alpha \\ 0 & 1 \end{pmatrix}$$

has a smallest singular value equal to

$$\sigma_2 = \left( \frac{2}{2 + |\alpha|^2 + |\alpha|\sqrt{4 + |\alpha|^2}} \right)^{1/2}.$$

As  $|\alpha| \rightarrow \infty$ , the smallest singular value approaches zero,  $\sigma_2 \rightarrow 0$ , so that the absolute distance of  $A$  to singularity decreases. ■

### Exercises

- Let  $A \in \mathbb{C}^{n \times n}$ . Show: All singular values of  $A$  are the same if and only if  $A$  is a multiple of a unitary matrix.
  - Show that the singular values of a Hermitian idempotent matrix are 0 and 1.
  - Show:  $A \in \mathbb{C}^{n \times n}$  is Hermitian positive definite if and only if it has a SVD  $A = V\Sigma V^*$  where  $\Sigma$  is nonsingular.
  - Let  $A, B \in \mathbb{C}^{m \times n}$ . Show:  $A$  and  $B$  have the same singular values if and only if there exist unitary matrices  $Q \in \mathbb{C}^{n \times n}$  and  $P \in \mathbb{C}^{m \times m}$  such that  $B = PAQ$ .
  - Let  $A \in \mathbb{C}^{m \times n}$ ,  $m \geq n$ , with QR decomposition  $A = Q \begin{pmatrix} R \\ 0 \end{pmatrix}$ , where  $Q \in \mathbb{C}^{m \times m}$  is unitary and  $R \in \mathbb{C}^{n \times n}$ . Determine a SVD of  $A$  from a SVD of  $R$ .
  - Determine a SVD of a column vector, and a SVD of a row vector.
  - Let  $A \in \mathbb{C}^{m \times n}$  with  $m \geq n$ . Show that the singular values of  $A^*A$  are the squares of the singular values of  $A$ .
- If  $A \in \mathbb{C}^{n \times n}$  is Hermitian positive-definite and  $\alpha > -\sigma_n$  then  $A + \alpha I_n$  is also Hermitian positive-definite with singular values  $\sigma_j + \alpha$ .
  - Let  $A \in \mathbb{C}^{m \times n}$  and  $\alpha > 0$ . Express the singular values of  $(A^*A + \alpha I)^{-1}A^*$  in terms of  $\alpha$  and the singular values of  $A$ .
  - Let  $A \in \mathbb{C}^{m \times n}$ ,  $m \geq n$ . Show that the singular values of  $\begin{pmatrix} I_n \\ A \end{pmatrix}$  are equal to  $\sqrt{1 + \sigma_j^2}$ ,  $1 \leq j \leq n$ .

## 4.1 Extreme Singular Values

The smallest and largest singular values of a matrix provide information about the two norm of the matrix, the distance to singularity, and the two norm of the inverse.

**Fact 4.4 (Extreme Singular Values)** If  $A \in \mathbb{C}^{m \times n}$  has singular values  $\sigma_1 \geq \dots \geq \sigma_p$ , where  $p = \min\{m, n\}$ , then

$$\|A\|_2 = \max_{x \neq 0} \frac{\|Ax\|_2}{\|x\|_2} = \sigma_1, \quad \min_{x \neq 0} \frac{\|Ax\|_2}{\|x\|_2} = \sigma_p.$$

**Proof.** The two norm of  $A$  does not change when  $A$  is multiplied by unitary matrices, see Exercise (iv) in 2.6. Hence  $\|A\|_2 = \|\Sigma\|_2$ . Since  $\Sigma$  is a diagonal matrix, Exercise (i) in 2.6 implies  $\|\Sigma\|_2 = \max_j |\sigma_j| = \sigma_1$ .

To show the expression for  $\sigma_p$ , assume that  $m \geq n$ , so  $p = n$ . Then  $A$  has a SVD  $A = U \begin{pmatrix} \Sigma \\ 0 \end{pmatrix} V^*$ . Let  $z$  be a vector so that  $\|z\|_2 = 1$  and  $\|Az\|_2 = \min_{\|x\|_2=1} \|Ax\|_2$ . With  $y = V^*z$  we get

$$\min_{\|x\|_2=1} \|Ax\|_2 = \|Az\|_2 = \|\Sigma V^*z\|_2 = \|\Sigma y\|_2 = \left( \sum_{i=1}^n \sigma_i^2 |y_i|^2 \right)^{1/2} \geq \sigma_n \|y\|_2 = \sigma_n.$$

Thus,  $\sigma_n \leq \min_{\|x\|_2=1} \|Ax\|_2$ . As for the reverse inequality,

$$\sigma_n = \|\Sigma e_n\|_2 = \|U^* A V e_n\|_2 = \|A(V e_n)\|_2 \geq \min_{\|x\|_2=1} \|Ax\|_2.$$

The proof for  $m < n$  is analogous.  $\square$

The extreme singular values are useful because they provide information about the two norm condition number with respect to inversion, and the distance to singularity.

The expressions below show that the largest singular value determines how much a matrix can stretch a unit-norm vector and the smallest singular value determines how much a matrix can shrink a unit-norm vector.

**Fact 4.5** If  $A \in \mathbb{C}^{n \times n}$  is nonsingular with singular values  $\sigma_1 \geq \dots \geq \sigma_n > 0$ , then

$$\|A^{-1}\|_2 = 1/\sigma_n, \quad \kappa_2(A) = \sigma_1/\sigma_n.$$

The absolute distance of  $A$  to singularity is

$$\sigma_n = \min \{ \|E\|_2 : A + E \text{ is singular} \}$$

and the relative distance is

$$\frac{\sigma_n}{\sigma_1} = \min \left\{ \frac{\|E\|_2}{\|A\|_2} : A + E \text{ is singular} \right\}.$$

**Proof.** Remark 4.2 implies that  $1/\sigma_j$  are the singular values of  $A^{-1}$ , so that  $\|A^{-1}\|_2 = \max_j 1/|\sigma_j| = 1/\sigma_n$ . The expressions for the distance to singularity follow from Fact 2.29 and Corollary 2.30.  $\square$

Fact 4.5 implies that a nonsingular matrix is almost singular in the absolute sense, if its smallest singular value is close to zero. If the smallest and largest singular values are far apart, i.e. if  $\sigma_1 \gg \sigma_n$ , then the matrix is ill-conditioned with respect to inversion in the normwise relative sense, and it is almost singular in the relative sense.

The singular values themselves are well-conditioned in the normwise absolute sense. We show this below for the extreme singular values.

**Fact 4.6** Let  $A, A + E \in \mathbb{C}^{m \times n}$ ,  $p = \min\{m, n\}$ ,  $\sigma_1 \geq \dots \geq \sigma_p$  the singular values of  $A$ , and  $\tilde{\sigma}_1 \geq \dots \geq \tilde{\sigma}_p$  the singular values of  $A + E$ . Then

$$|\tilde{\sigma}_1 - \sigma_1| \leq \|E\|_2, \quad |\tilde{\sigma}_p - \sigma_p| \leq \|E\|_2.$$

**Proof.** The inequality for  $\sigma_1$  follows from  $\sigma_1 = \|A\|_2$  and Fact 2.13, which states that norms are well-conditioned.

Regarding the bound for  $\sigma_p$ , let  $y$  be a vector so that  $\sigma_p = \|Ay\|_2$  and  $\|y\|_2 = 1$ . Then the triangle inequality implies

$$\tilde{\sigma}_p = \min_{\|x\|_2=1} \|(A + E)x\|_2 \leq \|(A + E)y\|_2 \leq \|Ay\|_2 + \|Ey\|_2 = \sigma_p + \|Ey\|_2 \leq \sigma_p + \|E\|_2.$$

Hence  $\tilde{\sigma}_p - \sigma_p \leq \|E\|_2$ . To show that  $-\|E\|_2 \leq \tilde{\sigma}_p - \sigma_p$ , let  $y$  be a vector so that  $\tilde{\sigma}_p = \|(A + E)y\|_2$  and  $\|y\|_2 = 1$ . Then the triangle inequality yields

$$\begin{aligned} \sigma_p &= \min_{\|x\|_2=1} \|Ax\|_2 \leq \|Ay\|_2 = \|(A + E)y - Ey\|_2 \leq \|(A + E)y\|_2 + \|Ey\|_2 \\ &= \tilde{\sigma}_p + \|Ey\|_2 \leq \tilde{\sigma}_p + \|E\|_2. \end{aligned}$$

$\square$

## Exercises

1. Extreme Singular Values of a Product.

Let  $A \in \mathbb{C}^{k \times m}$ ,  $B \in \mathbb{C}^{m \times n}$ ,  $q = \min\{k, n\}$ , and  $p = \min\{m, n\}$ . Then

$$\sigma_1(AB) \leq \sigma_1(A)\sigma_1(B), \quad \sigma_q(AB) \leq \sigma_1(A)\sigma_p(B).$$

2. Appending a column to a tall and skinny matrix does not increase the smallest singular value but can decrease it, because the new column may depend linearly on the old ones. The largest singular value does not decrease but it can increase, because more “mass” is added to the matrix.

Let  $A \in \mathbb{C}^{m \times n}$  with  $m > n$ ,  $z \in \mathbb{C}^m$ , and  $B = \begin{pmatrix} A & z \end{pmatrix}$ . Show that  $\sigma_{n+1}(B) \leq \sigma_n(A)$ , and  $\sigma_1(B) \geq \sigma_1(A)$ .

3. Appending a row to a tall and skinny matrix does not decrease the smallest singular value but can increase it. Intuitively, this is because the columns become longer which gives them an opportunity to become more linearly independent. The largest singular value does not decrease but can increase because more “mass” is added to the matrix.

Let  $A \in \mathbb{C}^{m \times n}$  with  $m \geq n$ ,  $z \in \mathbb{C}^n$ , and  $B = \begin{pmatrix} A \\ z^* \end{pmatrix}$ . Show that

$$\sigma_n(B) \geq \sigma_n(A), \quad \sigma_1(A) \leq \sigma_1(B) \leq \sqrt{\sigma_1(A)^2 + \|z\|_2^2}.$$

## 4.2 Rank

For a nonsingular matrix, all singular values are nonzero. For a general matrix, the number of nonzero singular values measures how much ‘information’ is contained in a matrix, while the number of zero singular values indicates the amount of ‘redundancy’.

**Definition 4.7 (Rank).** *The number of nonzero singular values of a matrix  $A \in \mathbb{C}^{m \times n}$  is called the rank of  $A$ . A  $m \times n$  zero matrix has rank 0.*

### Example 4.8

- If  $A \in \mathbb{C}^{m \times n}$  then  $\text{rank}(A) \leq \min\{m, n\}$ .  
This follows from Remark 4.2.
- If  $A \in \mathbb{C}^{n \times n}$  is nonsingular then  $\text{rank}(A) = n = \text{rank}(A^{-1})$ .  
A nonsingular matrix  $A$  contains the maximum amount of information, because it can reproduce any vector  $b \in \mathbb{C}^n$  by means of  $b = Ax$ .
- For any  $m \times n$  zero matrix  $0$ ,  $\text{rank}(0) = 0$ .  
The zero matrix contains no information. It can only reproduce the zero vector, because  $0x = 0$  for any vector  $x$ .
- If  $A \in \mathbb{C}^{m \times n}$  has  $\text{rank}(A) = n$  then  $A$  has a SVD  $A = U \begin{pmatrix} \Sigma \\ 0 \end{pmatrix} V^*$ , where  $\Sigma$  is nonsingular. This means, all singular values of  $A$  are nonzero.
- If  $A \in \mathbb{C}^{m \times n}$  has  $\text{rank}(A) = m$  then  $A$  has a SVD  $A = U \begin{pmatrix} \Sigma & 0 \end{pmatrix} V^*$ , where  $\Sigma$  is nonsingular. This means, all singular values of  $A$  are nonzero.

■

A nonzero outer product  $uv^*$  contains little information: because  $uv^*x = (v^*x)u$ , the outer product  $uv^*$  can produce only multiples of the vector  $u$ .

**Remark 4.9 (Outer Product).** *If  $u \in \mathbb{C}^m$  and  $v \in \mathbb{C}^n$  with  $u \neq 0$  and  $v \neq 0$  then  $\text{rank}(uv^*) = 1$ .*

*To see this, determine a SVD of  $uv^*$ . Let  $U \in \mathbb{C}^{m \times m}$  be a unitary matrix so that  $U^*u = \|u\|_2 e_1$ , and  $V \in \mathbb{C}^{n \times n}$  be a unitary matrix so that  $V^*v = \|v\|_2 e_1$ . Substituting these expressions into  $uv^*$  shows that  $uv^* = U\Sigma V^*$  is a SVD, where  $\Sigma \in$*

$\mathbb{R}^{m \times n}$  and  $\Sigma = \|u\|_2 \|v\|_2 e_1 e_1^*$ . Therefore the singular values of  $uv^*$  are  $\|u\|_2 \|v\|_2$ , and  $(\min\{m, n\} - 1)$  zeros. In particular,  $\|uv^*\|_2 = \|u\|_2 \|v\|_2$ .

The above example demonstrates that a nonzero outer product has rank one. Now we show that a matrix of rank  $r$  can be represented as a sum of  $r$  outer products. To this end we distinguish the columns of the left and right singular vector matrices.

**Definition 4.10 (Singular Vectors).** Let  $A \in \mathbb{C}^{m \times n}$ , with SVD  $A = U \begin{pmatrix} \Sigma \\ 0 \end{pmatrix} V^*$  if  $m \geq n$ , and SVD  $A = U \begin{pmatrix} \Sigma & 0 \end{pmatrix} V^*$  if  $m \leq n$ . Set  $p = \min\{m, n\}$  and partition

$$U = (u_1 \ \dots \ u_m), \quad V = (v_1 \ \dots \ v_n), \quad \Sigma = \begin{pmatrix} \sigma_1 & & & \\ & \ddots & & \\ & & \ddots & \\ & & & \sigma_p \end{pmatrix},$$

where  $\sigma_1 \geq \dots \geq \sigma_p \geq 0$ .

We call  $\sigma_j$  the  $j$ th singular value,  $u_j$  the  $j$ th left singular vector, and  $v_j$  the  $j$ th right singular vector.

Corresponding left and right singular vectors are related to each other.

**Remark 4.11.** Let  $A$  have a SVD as in Definition 4.10. Then

$$Av_i = \sigma_i u_i, \quad A^* u_i = \sigma_i v_i, \quad 1 \leq i \leq p.$$

This follows from the fact that  $U$  and  $V$  are unitary, and  $\Sigma$  is Hermitian.

Now we are ready to derive an economical representation for a matrix, where the size of the representation is proportional to the rank of the matrix. Fact 4.12 below shows that a matrix of rank  $r$  can be expressed in terms of  $r$  outer products. These outer products involve the singular vectors associated with the nonzero singular values.

**Fact 4.12 (Reduced SVD)** Let  $A \in \mathbb{C}^{m \times n}$  have a SVD as in Definition 4.10. If  $\text{rank}(A) = r$  then

$$A = \sum_{j=1}^r \sigma_j u_j v_j^*.$$

**Proof.** From  $\text{rank}(A) = r$  follows  $\sigma_1 \geq \dots \geq \sigma_r > 0$ . Confine the nonzero singular values to the matrix  $\Sigma_r$ , so that

$$\Sigma_r = \begin{pmatrix} \sigma_1 & & & \\ & \ddots & & \\ & & \ddots & \\ & & & \sigma_r \end{pmatrix}, \quad \text{and} \quad A = U \begin{pmatrix} \Sigma_r & 0 \\ 0 & 0 \end{pmatrix} V^*$$

is a SVD of  $A$ . Partitioning the singular vectors conformally with the nonzero singular values,

$$U = \begin{pmatrix} r & m-r \\ U_r & U_{m-r} \end{pmatrix}, \quad V = \begin{pmatrix} r & n-r \\ V_r & V_{n-r} \end{pmatrix},$$

yields  $A = U_r \Sigma_r V_r^*$ . Using  $U_r = (u_1 \ \dots \ u_r)$  and  $V_r = (v_1 \ \dots \ v_r)$ , and viewing matrix multiplication as an outer product, as in View 4 of §1.7, shows

$$A = U_r \Sigma_r V_r^* = (\sigma_1 u_1 \ \dots \ \sigma_r u_r) \begin{pmatrix} v_1^* \\ \vdots \\ v_r^* \end{pmatrix} = \sum_{j=1}^r \sigma_j u_j v_j^*.$$

□

For a nonsingular matrix, the thin SVD is equal to the ordinary SVD.

Based on the above outer product representation of a matrix, we will now show that the singular vectors associated with the  $k$  largest singular values of  $A$  determine the rank  $k$  matrix that is closest to  $A$  in the two-norm. Moreover, the  $(k+1)$ st singular value of  $A$  is the absolute distance of  $A$ , in the two norm, to the set of rank  $k$  matrices.

**Fact 4.13 (Optimality of the SVD)** Let  $A \in \mathbb{C}^{m \times n}$  have a SVD as in Definition 4.9. If  $k < \text{rank}(A)$  then the absolute distance of  $A$  to the set of rank  $k$  matrices is

$$\sigma_{k+1} = \min_{B \in \mathbb{C}^{m \times n}, \text{rank}(B)=k} \|A - B\|_2 = \|A - A_k\|_2,$$

where  $A_k = \sum_{j=1}^k \sigma_j u_j v_j^*$ .

**Proof.** Write the SVD as

$$A = U \begin{pmatrix} \Sigma_1 & \\ & \Sigma_2 \end{pmatrix} V^*, \quad \text{where } \Sigma_1 = \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_{k+1} \end{pmatrix}$$

and  $\sigma_1 \geq \dots \geq \sigma_{k+1} > 0$ , so that  $\Sigma_1$  is nonsingular. The idea is to show that the distance of  $\Sigma_1$  to the set of singular matrices, which is  $\sigma_{k+1}$ , is a lower bound for the distance of  $A$  to the set of all rank  $k$  matrices.

Let  $C \in \mathbb{C}^{m \times n}$  be a matrix with  $\text{rank}(C) = k$ , and partition

$$U^* C V = \begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix},$$

where  $C_{11}$  is  $(k+1) \times (k+1)$ . From  $\text{rank}(C) = k$  follows  $\text{rank}(C_{11}) \leq k$  (although it is intuitively clear, it is proved rigorously in Fact 6.19), so that  $C_{11}$  is singular. Since the two norm is invariant under multiplication by unitary matrices we obtain

$$\|A - C\|_2 = \left\| \begin{pmatrix} \Sigma_1 & \\ & \Sigma_2 \end{pmatrix} - U^* C V \right\|_2 = \left\| \begin{pmatrix} \Sigma_1 - C_{11} & -C_{12} \\ -C_{21} & \Sigma_2 - C_{22} \end{pmatrix} \right\|_2 \geq \|\Sigma_1 - C_{11}\|_2.$$

Since  $\Sigma_1$  is nonsingular and  $C_{11}$  is singular, Facts 2.29 and 4.5 imply that  $\|\Sigma_1 - C_{11}\|_2$  is bounded below by the distance of  $\Sigma_1$  from singularity, and

$$\|\Sigma_1 - C_{11}\|_2 \geq \min\{\|\Sigma_1 - B_{11}\|_2 : B_{11} \text{ is singular}\} = \sigma_{k+1}.$$

A matrix  $C$  for which  $\|A - C\|_2 = \sigma_{k+1}$  is  $C = A_k$ . This is because

$$C_{11} = \begin{pmatrix} \sigma_1 & & & \\ & \ddots & & \\ & & \sigma_k & \\ & & & 0 \end{pmatrix}, \quad C_{12} = 0, \quad C_{21} = 0, \quad C_{22} = 0.$$

Since the  $\Sigma_1 - C_{11}$  has  $k$  diagonal elements equal to zero, and the diagonal elements of  $\Sigma_2$  are less than or equal to  $\sigma_{k+1}$  we obtain

$$\|A - C\|_2 = \left\| \begin{pmatrix} \Sigma_1 - C_{11} & 0 \\ 0 & \Sigma_2 \end{pmatrix} \right\|_2 = \left\| \begin{pmatrix} \sigma_{k+1} & \\ & \Sigma_2 \end{pmatrix} \right\|_2 = \sigma_{k+1}.$$

□

The singular values also help us to relate the rank of  $A$  to the rank of  $A^*A$  and  $AA^*$ . This will be important later on for the solution of least squares problems.

**Fact 4.14** For any matrix  $A \in \mathbb{C}^{m \times n}$

1.  $\text{rank}(A) = \text{rank}(A^*)$ .
2.  $\text{rank}(A) = \text{rank}(A^*A) = \text{rank}(AA^*)$ .
3.  $\text{rank}(A) = n$  if and only if  $A^*A$  is nonsingular.
4.  $\text{rank}(A) = m$  if and only if  $AA^*$  is nonsingular.

**Proof.**

1. This follows from Remark 4.2, because  $A$  and  $A^*$  have the same singular values.
2. If  $m \geq n$  then  $A$  has a SVD  $A = U \begin{pmatrix} \Sigma \\ 0 \end{pmatrix} V^*$ , and  $A^*A = V \Sigma^2 V^*$  is a SVD of  $A^*A$ . Since  $\Sigma$  and  $\Sigma^2$  have the same number of nonzero diagonal elements,  $\text{rank}(A) = \text{rank}(A^*A)$ . Also,  $AA^* = U \begin{pmatrix} \Sigma^2 & 0 \\ 0 & 0 \end{pmatrix} U^*$  is a SVD of  $AA^*$ . As before,  $\text{rank}(A) = \text{rank}(AA^*)$  because  $\Sigma$  and  $\Sigma^2$  have the same number of nonzero diagonal elements.  
A similar argument applies when  $m < n$ .
3. Since  $A^*A$  is  $n \times n$ ,  $A^*A$  is nonsingular if and only if  $n = \text{rank}(A^*A) = \text{rank}(A)$ , where the second equality follows from item 2.
4. The proof is similar to that of item 3.

□

In item 3 above the matrix  $A$  has linearly independent columns, and in item 4 it has linearly independent rows. Below we give another name to such matrices.

**Definition 4.15 (Full Rank).** A matrix  $A \in \mathbb{C}^{m \times n}$  has full column rank if  $\text{rank}(A) = n$ , and full row rank if  $\text{rank}(A) = m$ .

A matrix  $A \in \mathbb{C}^{m \times n}$  has full rank, if  $A$  has full column rank or full row rank. A matrix that does not have full rank is rank deficient.

**Example.**

- A nonsingular matrix has full row rank and full column rank.
- A nonzero column vector has full column rank, and a nonzero row vector has full row rank.
- If  $A \in \mathbb{C}^{n \times n}$  is nonsingular then  $(A \ B)$  has full row rank for any matrix  $B \in \mathbb{C}^{n \times m}$ , and  $\begin{pmatrix} A \\ C \end{pmatrix}$  has full column rank, for any matrix  $C \in \mathbb{C}^{m \times n}$ .
- A singular square matrix is rank deficient.

□

Below we show that matrices with orthonormal columns also have full column rank. Recall from Definition 3.37 that  $A$  has orthonormal columns if  $A^*A = I$ .

**Fact 4.16** A matrix  $A \in \mathbb{C}^{m \times n}$  with orthonormal columns has  $\text{rank}(A) = n$ , and all singular values are equal to one.

**Proof.** Fact 4.14 implies  $\text{rank}(A) = \text{rank}(A^*A) = \text{rank}(I_n) = n$ . Thus  $A$  has full column rank, and we can write its SVD as  $A = U \begin{pmatrix} \Sigma \\ 0 \end{pmatrix} V^*$ . Then  $I_n = A^*A = V \Sigma^2 V^*$  implies  $\Sigma = I_n$ , so that all singular values of  $A$  are equal to one. □

## Exercises

- (i) Let  $A \in \mathbb{C}^{m \times n}$ . Show: If  $Q \in \mathbb{C}^{m \times m}$  and  $P \in \mathbb{C}^{n \times n}$  are unitary then  $\text{rank}(A) = \text{rank}(QAP)$ .
- (ii) What can you say about the rank of a nilpotent matrix, and the rank of an idempotent matrix?
- (iii) Let  $A \in \mathbb{C}^{m \times n}$ . Show: If  $\text{rank}(A) = n$  then  $\|(A^*A)^{-1}A^*\|_2 = 1/\sigma_n$ , and if  $\text{rank}(A) = m$  then  $\|(AA^*)^{-1}A\|_2 = 1/\sigma_m$ .
- (iv) Let  $A \in \mathbb{C}^{m \times n}$  with  $\text{rank}(A) = n$ . Show that  $A(A^*A)^{-1}A^*$  is idempotent and Hermitian, and  $\|A(A^*A)^{-1}A^*\|_2 = 1$ .
- (v) Let  $A \in \mathbb{C}^{m \times n}$  with  $\text{rank}(A) = m$ . Show that  $A^*(AA^*)^{-1}A$  is idempotent and Hermitian, and  $\|A^*(AA^*)^{-1}A\|_2 = 1$ .

- (vi) Nilpotent Matrices.  
 Let  $A \in \mathbb{C}^{n \times n}$  be nilpotent so that  $A^j = 0$  and  $A^{j-1} \neq 0$  for some  $j \geq 1$ . Let  $b \in \mathbb{C}^n$  with  $A^{j-1}b \neq 0$ . Show that  $K = (b \quad Ab \quad \dots \quad A^{j-1}b)$  has full column rank.
- (vii) In Fact 4.13 let  $B$  be a multiple of  $A_k$ , i.e.  $B = \alpha A_k$ . Determine  $\|A - B\|_2$ .

- Let  $A \in \mathbb{C}^{n \times n}$ . Show that there exists a unitary matrix  $Q$  such that  $A^* = QAQ$ .
- Polar Decomposition.  
 Show: If  $A \in \mathbb{C}^{m \times n}$  has  $\text{rank}(A) = n$  then there is a factorization  $A = PH$ , where  $P \in \mathbb{C}^{m \times n}$  has orthonormal columns, and  $H \in \mathbb{C}^{n \times n}$  is Hermitian positive definite.
- The polar factor  $P$  is the closest matrix with orthonormal columns in the two norm.  
 Let  $A \in \mathbb{C}^{n \times n}$  have a polar decomposition  $A = PH$ . Show that  $\|A - P\|_2 \leq \|A - Q\|_2$  for any unitary matrix  $Q$ .
- The distance of a matrix  $A$  from its the polar factor  $P$  is determined by how close the columns  $A$  are to being orthonormal.  
 Let  $A \in \mathbb{C}^{m \times n}$  with  $\text{rank}(A) = n$  have a polar decomposition  $A = PH$ . Show that

$$\frac{\|A^*A - I_n\|_2}{1 + \|A\|_2} \leq \|A - P\|_2 \leq \frac{\|A^*A - I_n\|_2}{1 + \sigma_n}.$$

- Let  $A \in \mathbb{C}^{n \times n}$  and  $\sigma > 0$ . Show:  $\sigma$  is a singular value of  $A$  if and only if the matrix

$$\begin{pmatrix} A & -\sigma I \\ -\sigma I & A^* \end{pmatrix}$$

is singular.

- Rank Revealing QR Factorization.  
 With an appropriate permutation of the columns, a QR factorization can almost reveal the smallest singular value of a full column rank matrix.  
 Let  $A \in \mathbb{C}^{m \times n}$  with  $\text{rank}(A) = n$  and smallest singular value  $\sigma_n$ . Let the corresponding singular vectors be  $Av = \sigma_n u$ , where  $\|v\|_2 = \|u\|_2 = 1$ . Choose a permutation  $P$  so that  $w = P^*v$  and  $|w_n| = \|w\|_\infty$ , and let  $AP = Q \begin{pmatrix} R \\ 0 \end{pmatrix}$  be a QR decomposition of  $AP$ . Show:  $|r_{nn}| \leq \sqrt{n}\sigma_n$ .

### 4.3 Singular Vectors

The singular vectors of a matrix  $A$  give information about the column spaces and null spaces of  $A$  and  $A^*$ .

The column space of a matrix  $A$  is the set of all right hand sides  $b$  for which the system  $Ax = b$  has a solution, and the null space of  $A$  determines whether these solutions are unique.

**Definition 4.17 (Column Space and Null Space).** If  $A \in \mathbb{C}^{m \times n}$  then the set

$$\mathcal{R}(A) = \{b \in \mathbb{C}^m : b = Ax \text{ for some } x \in \mathbb{C}^n\}$$

is the column space or range of  $A$ , and the set

$$\text{Ker}(A) = \{x \in \mathbb{C}^n : Ax = 0\}$$

is the kernel or null space of  $A$ .

**Example.**

- The column space of a  $m \times n$  zero matrix is the zero vector, and the null space is  $\mathbb{C}^n$ , i.e.  $\mathcal{R}(0_{m \times n}) = \{0_{m \times 1}\}$  and  $\text{Ker}(0_{m \times n}) = \mathbb{C}^n$ .
- The column space of a  $n \times n$  nonsingular complex matrix is  $\mathbb{C}^n$ , and the null space consists of the single vector  $0_{n \times 1}$ .
- $\text{Ker}(A) = \{0\}$  if and only if the columns of the matrix  $A$  are linearly independent.
- If  $A \in \mathbb{C}^{m \times n}$  then for all  $k \geq 1$

$$\mathcal{R}\begin{pmatrix} A & 0_{m \times k} \end{pmatrix} = \mathcal{R}(A), \quad \text{Ker}\begin{pmatrix} A \\ 0_{k \times n} \end{pmatrix} = \text{Ker}(A).$$

- If  $A \in \mathbb{C}^{n \times n}$  is nonsingular then for any  $B \in \mathbb{C}^{m \times p}$  and  $C \in \mathbb{C}^{p \times n}$

$$\mathcal{R}\begin{pmatrix} A & B \end{pmatrix} = \mathcal{R}(A), \quad \text{Ker}\begin{pmatrix} A \\ C \end{pmatrix} = \{0_{n \times 1}\},$$

□

The column and null spaces of  $A^*$  are also important, and we give them names that relate to the matrix  $A$ .

**Definition 4.18 (Row Space and Left Null Space).** Let  $A \in \mathbb{C}^{m \times n}$ . The set

$$\mathcal{R}(A^*) = \{d \in \mathbb{C}^n : d = A^*y \text{ for some } y \in \mathbb{C}^m\}$$

is the row space of  $A$ . The set

$$\text{Ker}(A^*) = \{y \in \mathbb{C}^m : A^*y = 0\}$$

is the left null space of  $A$ .

Note that all spaces of a matrix are defined by *column* vectors.

**Example 4.19** If  $A$  is Hermitian then  $\mathcal{R}(A^*) = \mathcal{R}(A)$  and  $\text{Ker}(A^*) = \text{Ker}(A)$ .

■

The singular vectors reproduce the four spaces associated with a matrix. Let  $A \in \mathbb{C}^{m \times n}$  with  $\text{rank}(A) = r$  and SVD

$$A = U \begin{pmatrix} \Sigma_r & 0 \\ 0 & 0 \end{pmatrix} V^*$$

where  $\Sigma_r$  is nonsingular, and

$$\begin{matrix} & r & n-r \\ r & \begin{pmatrix} \Sigma_r & 0 \\ 0 & 0 \end{pmatrix} \\ m-r & \end{matrix}, \quad U = \begin{pmatrix} U_r & U_{m-r} \end{pmatrix}, \quad V = \begin{pmatrix} V_r & V_{n-r} \end{pmatrix}.$$

**Fact 4.20 (Spaces of a Matrix and Singular Vectors)** Let  $A \in \mathbb{C}^{m \times n}$ .

1. The leading  $r$  left singular vectors represent the column space of  $A$ :  
If  $A \neq 0$  then  $\mathcal{R}(U_r) = \mathcal{R}(A)$ , otherwise  $\mathcal{R}(A) = \{0_{m \times 1}\}$ .
2. The trailing  $n - r$  right singular vectors represent the null space of  $A$ :  
If  $\text{rank}(A) = r < n$  then  $\mathcal{R}(V_{n-r}) = \text{Ker}(A)$ , otherwise  $\text{Ker}(A) = \{0_{n \times 1}\}$ .
3. The leading  $r$  right singular vectors represent the row space of  $A$ :  
If  $A \neq 0$  then  $\mathcal{R}(A^*) = \mathcal{R}(V_r)$ , otherwise  $\mathcal{R}(A^*) = \{0_{n \times 1}\}$ .
4. The trailing  $m - r$  left singular vectors represent the left null space of  $A$ :  
If  $r < m$  then  $\mathcal{R}(U_{m-r}) = \text{Ker}(A^*)$ , otherwise  $\text{Ker}(A^*) = \{0_{m \times 1}\}$ .

**Proof.** Although the statements may be intuitively obvious, they are proved rigorously in Section 6.1.  $\square$

The singular vectors help us to relate the spaces of  $A^*A$  and  $AA^*$  to those of the matrix  $A$ . Since  $A^*A$  and  $AA^*$  are Hermitian, we need to specify only two spaces, see Example 4.19.

**Fact 4.21 (Spaces of  $A^*A$  and  $AA^*$ )** Let  $A \in \mathbb{C}^{m \times n}$ .

1.  $\text{Ker}(A^*A) = \text{Ker}(A)$  and  $\mathcal{R}(A^*A) = \mathcal{R}(A^*)$ .
2.  $\mathcal{R}(AA^*) = \mathcal{R}(A)$  and  $\text{Ker}(AA^*) = \text{Ker}(A^*)$ .

**Proof.** Fact 4.14 implies that  $A^*A$  and  $AA^*$  have the same rank as  $A$ . Since  $A^*A$  has the same right singular vectors as  $A$ , Fact 4.20 implies  $\text{Ker}(A^*A) = \text{Ker}(A)$  and  $\mathcal{R}(A^*A) = \mathcal{R}(A^*)$ . Since  $AA^*$  has the same left singular vectors as  $A$ , Fact 4.20 implies  $\mathcal{R}(AA^*) = \mathcal{R}(A)$  and  $\text{Ker}(AA^*) = \text{Ker}(A^*)$ .  $\square$

In the special case when the rank of a matrix is equal to the number of rows, then the number of elements in the column space is as large as possible. When the rank of the matrix is equal to the number of columns then the number of elements in the null space is as small as possible.

**Fact 4.22 (Spaces of Full Rank Matrices)** Let  $A \in \mathbb{C}^{m \times n}$ . Then

1.  $\text{rank}(A) = m$  if and only if  $\mathcal{R}(A) = \mathbb{C}^m$ .
2.  $\text{rank}(A) = n$  if and only if  $\text{Ker}(A) = \{0\}$ .

**Proof.** Let  $A = U \begin{pmatrix} \Sigma_r & 0 \\ 0 & 0 \end{pmatrix} V^*$  be a SVD of  $A$ , where  $\Sigma_r$  is nonsingular.

1. From Fact 4.20 follows  $\mathcal{R}(A) = \mathcal{R}(U_r)$ . Hence  $r = m$  if and only if  $U_r = U$ , because  $U$  is nonsingular so that  $\mathcal{R}(U) = \mathbb{C}^m$ .
2. Fact 4.20 also implies  $r = n$  if and only if  $V_{n-r}$  is empty, which means  $\text{Ker}(A) = \{0\}$ .

□

If the matrix in a linear system has full rank, then existence or uniqueness of a solution is guaranteed.

**Fact 4.23 (Solutions of Full Rank Linear Systems)** Let  $A \in \mathbb{C}^{m \times n}$ .

1. If  $\text{rank}(A) = m$  then  $Ax = b$  has a solution  $x = A^*(AA^*)^{-1}b$  for every  $b \in \mathbb{C}^m$ .
2. If  $\text{rank}(A) = n$  and if  $b \in \mathcal{R}(A)$  then  $Ax = b$  has the unique solution  $x = (A^*A)^{-1}A^*b$ .

**Proof.**

1. Fact 4.22 implies that  $Ax = b$  has a solution for every  $b \in \mathbb{C}^m$ , and Fact 4.14 implies that  $AA^*$  is nonsingular. Clearly,  $x = A^*(AA^*)^{-1}b$  satisfies  $Ax = b$ .
2. Since  $b \in \mathcal{R}(A)$ ,  $Ax = b$  has a solution. Multiplying on the left by  $A^*$  gives  $A^*Ax = A^*b$ . According to Fact 4.14,  $A^*A$  is nonsingular, so that  $x = (A^*A)^{-1}A^*b$ .

Suppose  $Ax = b$  and  $Ay = b$  then  $A(x - y) = 0$ . Fact 4.22 implies that  $\text{Ker}(A) = \{0\}$ , so  $x = y$ , which proves uniqueness.

□

## Exercises

- (i) Fredholm's Alternatives.
  - (a) The first alternative implies that  $\mathcal{R}(A)$  and  $\text{Ker}(A^*)$  have only the zero vector in common. Assume  $b \neq 0$  and show:  
If  $Ax = b$  has a solution then  $b^*A \neq 0$ .  
In other words: If  $b \in \mathcal{R}(A)$  then  $b \notin \text{Ker}(A^*)$ .
  - (b) The second alternative implies that  $\text{Ker}(A)$  and  $\mathcal{R}(A^*)$  have only the zero vector in common. Assume  $x \neq 0$  and show:  
If  $Ax = 0$  then there is no  $y$  such that  $x = A^*y$ .  
In other words: If  $x \in \text{Ker}(A)$  then  $x \notin \mathcal{R}(A^*)$ ,

(ii) Normal Matrices.

If  $A \in \mathbb{C}^n$  is Hermitian then  $\mathcal{R}(A^*) = \mathcal{R}(A)$  and  $\text{Ker}(A^*) = \text{Ker}(A)$ . These equalities remain true for a larger class of matrices, the so-called “normal matrices”. A matrix  $A \in \mathbb{C}^n$  is *normal* if  $A^*A = AA^*$ .

Show: If  $A \in \mathbb{C}^{n \times n}$  is normal then  $\mathcal{R}(A^*) = \mathcal{R}(A)$  and  $\text{Ker}(A^*) = \text{Ker}(A)$ .