

## 1. SVD

Suppose we have a matrix  $A \in \mathbb{R}^{m \times n}$  with rank  $r$ .

We define the compact SVD as follows:

$$\underbrace{A}_{m \times n} = \underbrace{U_r}_{m \times r} \underbrace{\Sigma_r}_{r \times r} \underbrace{V_r^\top}_{r \times n}.$$

Here,  $\Sigma_r \in \mathbb{R}^{r \times r}$  is a diagonal matrix containing non-zero singular values of  $A$ .

$$\Sigma_r = \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \end{bmatrix},$$

with  $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r$ .

Next,  $U_r \in \mathbb{R}^{m \times r}$  is given by,

$$U_r = [\vec{u}_1, \vec{u}_2, \dots, \vec{u}_r],$$

where  $u_i$  is a left singular vector corresponding to non-zero singular value,  $\sigma_i$ , for  $i = 1, 2, \dots, r$ . The columns of  $U_r$  are orthonormal and together they span the column space of  $A$ .

Finally,  $V_r^\top \in \mathbb{R}^{r \times n}$  is given by,

$$V_r^\top = \begin{bmatrix} \vec{v}_1^\top \\ \vec{v}_2^\top \\ \vdots \\ \vec{v}_r^\top \end{bmatrix},$$

where  $\vec{v}_j$  is a right singular vector corresponding to non-zero singular value,  $\sigma_j$  for  $j = 1, 2, \dots, r$ . The rows of  $V_r^\top$  are orthonormal and span the row space of  $A$ . Equivalently the columns of  $V_r$  span the column space of  $A^\top$ .

The matrix  $A$  can be expressed as,

$$A = \sigma_1 \vec{u}_1 \vec{v}_1^\top + \sigma_2 \vec{u}_2 \vec{v}_2^\top + \dots + \sigma_r \vec{u}_r \vec{v}_r^\top.$$

Assume now that  $m \geq n$ .

Another type of SVD which might be more familiar is the full SVD of  $A$  which is defined as follows:

$$\underbrace{A}_{m \times n} = \underbrace{U}_{m \times m} \underbrace{\Sigma}_{m \times n} \underbrace{V^\top}_{n \times n}.$$

Here,  $\Sigma \in \mathbb{R}^{m \times n}$  has non-diagonal entries as zero. The diagonal entries of  $\Sigma$  contain the singular values and we can write  $\Sigma$  in terms of  $\Sigma_r$  as,

$$\Sigma = \left[ \begin{array}{c|c} \Sigma_r & 0_{r \times (n-r)} \\ \hline 0_{(m-r) \times r} & 0_{(m-r) \times (n-r)} \end{array} \right]$$

Next,  $U \in \mathbb{R}^{m \times m}$  is an orthogonal matrix.  $U$  can be expressed in terms of  $U_r$  as,

$$U = \underbrace{\begin{bmatrix} U_r & \vec{u}_{r+1} & \dots & \vec{u}_m \end{bmatrix}}_{m \times r \quad m \times (m-r)}$$

The columns  $\vec{u}_{r+1}, \vec{u}_{r+2}, \dots, \vec{u}_m$  are left singular vectors corresponding to singular value 0, and together span the nullspace of  $A^\top$ .

Finally,  $V^\top$  is an orthogonal matrix and can be expressed in terms of  $V_r^\top$  as,

$$V^\top = \left\{ \begin{array}{l} \begin{bmatrix} V_r^\top \\ \vec{v}_{r+1}^\top \\ \vdots \\ \vec{v}_n^\top \end{bmatrix} \end{array} \right\} \begin{array}{l} r \times n \\ (n-r) \times n \end{array}$$

The rows  $\vec{v}_{r+1}^\top, \vec{v}_{r+2}^\top, \dots, \vec{v}_n^\top$  when transposed are the right singular vectors corresponding to singular value of 0 and together span the nullspace of  $A$ .

(a) For this problem assume that  $m > n > r$ . Which of the following are True:

(a)  $UU^\top = I$

**Solution:** True.  $UU^\top = I_m$  because  $U$  is an orthogonal matrix.

(b)  $U^\top U = I$

**Solution:** True.  $U^\top U = I_m$  because  $U$  is an orthogonal matrix.

(c)  $V^\top V = I$

**Solution:** True.  $V^\top V = I_n$  because  $V$  is an orthogonal matrix.

(d)  $VV^\top = I$

**Solution:** True.  $VV^\top = I_n$  because  $V$  is an orthogonal matrix.

(e)  $U_r^\top U_r = I$

**Solution:** True.  $U_r^\top U_r = I_r$  because the columns of  $U_r$  are orthonormal.

(f)  $U_r U_r^\top = I$

**Solution:** False.  $U_r U_r^\top$  is a  $m \times m$ , matrix but has rank less than or equal to  $r$  (since  $U_r$  has rank  $r$  and product of matrices has rank less than or equal to minimum of individual ranks).

(g)  $V_r V_r^\top = I$

**Solution:** False.  $V_r V_r^\top$  is a  $n \times n$ , matrix but has rank less than or equal to  $r$  (since  $V_r$  has rank  $r$  and product of matrices has rank less than or equal to minimum of individual ranks).

(h)  $V_r^\top V_r = I$

**Solution:** True.  $V_r^\top V_r = I_r$  because the columns of  $V_r$  are orthonormal.

(b) Going from the full SVD to compact SVD. Find the compact SVD of  $A$  which has the full SVD:

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

**Solution:** The compact SVD of  $A$  is given by:

$$A = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \begin{bmatrix} 2 \end{bmatrix} \begin{bmatrix} 1 & 0 \end{bmatrix}.$$

(c) Going from compact SVD to full SVD: Find the full SVD of  $A$  which has the compact SVD:

$$A = \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

**Solution:** Observe that in this case, the full SVD of  $A$  has  $\Sigma$  and  $V^\top$  as those in the compact SVD but  $U \in \mathbb{R}^{3 \times 3}$ . Thus we need to find a unit-norm column  $\vec{u}_3$  orthogonal to columns of  $U_r$ . We can use a system of linear equations to solve this. That is we want  $u_3 = [x, y, z]$  so we must have

- $[1/\sqrt{2}, 1/\sqrt{2}, 0]^\top \vec{u}_3 = 0$
- $[0, 0, 1]^\top \vec{u}_3 = 0$
- $\|\vec{u}_3\|_2 = 1$

Check that  $\vec{u}_3 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}$  satisfies our requirements.

Thus the full SVD of  $A$  is given by:

$$A = \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Alternatively we can use Gram-Schmidt procedure to find  $\vec{u}_3$ . This has the added advantage of being useful when we want to find the full SVD when more than one singular vector is missing.

(d) **(Optional)** For a given matrix  $A \in \mathbb{R}^{m \times n}$  with  $\text{rank}(A) = r = \min\{m, n\}$ . Prove the rank nullity theorem, i.e.,  $n = r + \dim(\mathcal{N}(A))$

**Solution:** Since rank is  $r$ , there exists  $r$  linearly independent columns of  $A$  which without loss of generality we assume are the first  $r$  columns which form a matrix  $A_1$ . So we can write  $A = \begin{bmatrix} A_1 & A_2 \end{bmatrix}$  where  $A_2 \in \mathbb{R}^{m \times (n-r)}$  which can all be written as linear combinations of columns of  $A_1$  and hence  $A_2 = A_1 B$  for some  $B \in \mathbb{R}^{r \times (n-r)}$ . Hence  $A = \begin{bmatrix} A_1 & A_1 B \end{bmatrix}$ . If  $\vec{x} = \begin{bmatrix} \vec{x}_1 \\ \vec{x}_2 \end{bmatrix} \in \mathcal{N}(A)$ , then  $\vec{0} = A\vec{x} = \begin{bmatrix} A_1 & A_2 \end{bmatrix} \vec{x} = A_1 \vec{x}_1 + A_2 \vec{x}_2 = A_1 (\vec{x}_1 + B \vec{x}_2)$ . As columns of  $A_1$  are linearly independent, this implies  $(\vec{x}_1 + B \vec{x}_2) = \vec{0}$ . Equivalently,

$$\vec{x} = \begin{bmatrix} -B \\ I_{n-r} \end{bmatrix} \vec{x}_2 = C \vec{x}_2$$

Thus every vector in the null space of  $A$  can be expressed as above. The matrix  $C$  above has dimensions  $n \times (n-r)$  so its rank can be at most  $n-r$ . To see that the rank is exactly  $n-r$ , follows from noticing that the columns of the  $C$  are linearly independent thanks to the  $I_{n-r}$ .

## 2. SVD Part 2

Consider  $A$  to be the  $4 \times 3$  matrix

$$A = \begin{bmatrix} \vec{a}_1 & \vec{a}_2 & \vec{a}_3 \end{bmatrix} \quad (1)$$

where  $\vec{a}_i$  for  $i \in \{1, 2, 3\}$  form a set of *orthogonal* vectors satisfying  $\|\vec{a}_1\|_2 = 3$ ,  $\|\vec{a}_2\|_2 = 2$ ,  $\|\vec{a}_3\|_2 = 1$ .

- (a) What is the SVD of  $A$ ? Express it as  $A = U\Sigma V^\top$ , with  $\Sigma$  the diagonal matrix of singular values ordered in decreasing fashion, and explicitly describe  $U$  and  $V$ .

**Solution:** The SVD of  $A = U\Sigma V^\top$ . Due to the orthogonality of the  $\vec{a}_i$  we have that

$$A^\top A = V\Sigma^2 V^\top = \begin{bmatrix} 9 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (2)$$

Thus  $V = I$  and  $\Sigma = \text{diag}(3, 2, 1)$ . Finally we have that  $U = A\Sigma^{-1}$  which becomes

$$U = \begin{bmatrix} \frac{\vec{a}_1}{3} & \frac{\vec{a}_2}{2} & \vec{a}_3 \end{bmatrix} \quad (3)$$

- (b) What is the dimension of the null space,  $\dim(\mathcal{N}(A))$ ?

**Solution:** From part (a) all of the singular values of the  $A$  are non-zero. So the dimension of the null space is 0. Alternatively, all the columns of  $A$  are orthogonal – so no (non-zero) linear combination of them can equal zero.

- (c) What is the rank of  $A$ ,  $\text{rank}(A)$ ? Provide an orthonormal basis for the range of  $A$ .

**Solution:** The rank of  $A$  is simply the number of non-zero singular values. So  $\text{rank}(A) = 3$ . The columns of  $U$  (defined above) provide an orthonormal basis for the range of  $A$ .

- (d) Let  $I_3$  denote the  $3 \times 3$  identity matrix. Consider the matrix  $\tilde{A} = \begin{bmatrix} A \\ I_3 \end{bmatrix} \in \mathbb{R}^{7 \times 3}$ . What are the singular values of  $\tilde{A}$  (in terms of the singular values of  $A$ )?

**Solution:** We have that  $\tilde{A}^\top \tilde{A} = A^\top A + I_3 = V(\Sigma^2 + I_3)V^\top$ . Hence if we denote  $\sigma_i$  as the singular values of  $A$  then the singular values of  $\tilde{A}$  are  $\tilde{\sigma}_i = \sqrt{\sigma_i^2 + 1}$  which are  $\sqrt{10}, \sqrt{5}, \sqrt{2}$ .

- (e) (Optional) Find an SVD of the matrix  $\tilde{A}$ .

**Solution:** The SVD of  $\tilde{A} = \tilde{U}\tilde{\Sigma}\tilde{V}^\top$  has  $\tilde{\Sigma} = \text{diag}(\tilde{\sigma}_1, \dots, \tilde{\sigma}_m)$ ; with  $\tilde{\sigma}_i = \sqrt{\sigma_i^2 + 1}$  where  $\sigma_i$  are the singular values of  $A$ .

The eigenvectors of  $\tilde{A}\tilde{A}^\top$  form the columns of matrix  $\tilde{U}$ , while the eigenvectors of  $\tilde{A}^\top \tilde{A}$  form the columns of  $\tilde{V}$ . We can see this by writing out  $\tilde{A}$  in terms of the SVD of  $A$  and the identity matrix  $I$ .

Since  $\tilde{A}^\top \tilde{A} = A^\top A + I$ , we can choose  $\tilde{V} = V$ .

The condition for  $(\tilde{u}, \tilde{v})$  to be a pair of left- and right-singular vectors of  $\tilde{A}$  are that both vectors must be unit-norm, and

$$\tilde{A}\tilde{v} = \tilde{\sigma}\tilde{u}, \quad \tilde{A}^\top \tilde{u} = \tilde{\sigma}\tilde{v}.$$

We have seen that we can choose  $\tilde{v} = v$  to be an eigenvector of  $A^\top A$  (that is, a right singular vector of  $A$ ).

Further, decomposing  $\tilde{u} = \begin{bmatrix} \tilde{u}^1 \\ \tilde{u}^2 \end{bmatrix}$ , with  $\tilde{u}^1 \in \mathbb{R}^4$  and  $\tilde{u}^2 \in \mathbb{R}^3$ , we obtain

$$Av = \tilde{\sigma}\tilde{u}^1, \quad v = \tilde{\sigma}\tilde{u}^2, \quad A^\top \tilde{u}^1 + \tilde{u}^2 = \tilde{\sigma}v.$$

Solving for the second equation:  $\tilde{u}^2 = v/\tilde{\sigma}$ , we obtain from the third  $A^\top \tilde{u}^1 = (\tilde{\sigma} - 1/\tilde{\sigma})v$ . Multiplying by  $A$ , and with the first equation, we then obtain

$$AA^\top \tilde{u}^1 = \tilde{\sigma}(\tilde{\sigma} - 1/\tilde{\sigma})\tilde{u}^1 = \sigma^2 \tilde{u}^1.$$

This shows that we can set  $\tilde{u}^1$  to be proportional to a left singular vector  $u$  of  $A$ , and  $\tilde{u}^2 = v/\tilde{\sigma}$  proportional to  $v$ . We have

$$\tilde{u} = \begin{bmatrix} \alpha u \\ \frac{1}{\tilde{\sigma}} v \end{bmatrix},$$

where  $\alpha$  must be chosen so that the above has unit Euclidean norm, that is:

$$\alpha = \frac{\sqrt{\tilde{\sigma}^2 - 1}}{\tilde{\sigma}} = \frac{\sigma}{\sqrt{\sigma^2 + 1}}.$$

We have obtained that a generic pair of left- and right singular vectors  $(\tilde{u}, \tilde{v})$  of  $\tilde{A}$  corresponding to the singular value  $\sqrt{\sigma^2 + 1}$ , can be constructed from a generic pair of left- and right singular vectors  $(u, v)$  of  $A$  corresponding to the singular value  $\sigma$ , with the choice

$$\tilde{u} = \begin{bmatrix} \frac{\sigma}{\sqrt{\sigma^2 + 1}} u \\ \frac{1}{\sqrt{\sigma^2 + 1}} v \end{bmatrix}, \quad \tilde{v} = v.$$