Lecture 5 More on the SVD

Fang Zhu

October 21, 2023

1 Prerequisite

We first give a more detailed proof of *Theorem 5.2* in the book.

Theorem 1.1. range$(\boldsymbol{A}) =< \boldsymbol{u}_1, \cdots, \boldsymbol{u}_l >$, null$(\boldsymbol{A}) =< \boldsymbol{v}_{l+1}, \cdots, \boldsymbol{v}_n >$

Proof. Assume that $\boldsymbol{A} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^\star$ and $\boldsymbol{A}$ has l's non-zero singular values, then

$$\begin{aligned}
\text{range}(\boldsymbol{A}) &= \{\boldsymbol{A}\boldsymbol{x} : \boldsymbol{x} \in \mathbb{R}^n\} \\
&= \{\boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^\star\boldsymbol{x} : \boldsymbol{x} \in \mathbb{R}^n\} \\
&= \{\boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{y} : \boldsymbol{y} = \boldsymbol{V}^\star\boldsymbol{x}, \boldsymbol{x} \in \mathbb{R}^n\} \\
&= \{\boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{y} : \boldsymbol{y} \in \mathbb{R}^n\}.
\end{aligned}$$

Let $\boldsymbol{c}_j = \sigma_j \boldsymbol{y}_j, (j = 1, 2, \cdots, l)$, it is clear that $\boldsymbol{c}_j$ can be any vector in $\mathbb{R}^n$ since $\boldsymbol{y}$ is an arbitrary vector in $\mathbb{R}^n$. Hence, we can conclude that

$$\text{range}(\boldsymbol{A}) = \{\boldsymbol{U}\boldsymbol{c} : \boldsymbol{c} \in \mathbb{R}^n\} =< \boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_l > .$$

Further the null space of $\boldsymbol{A}$ is as follows:

$$\begin{aligned}
\text{null}(\boldsymbol{A}) &= \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{A}\boldsymbol{x} = 0\} \\
&= \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^\star\boldsymbol{x} = 0\} \\
&= \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{\Sigma}\boldsymbol{V}^\star\boldsymbol{x} = 0\}.
\end{aligned}$$

Note that any vector $\boldsymbol{x}$ can be expressed by $\boldsymbol{x} = \sum_j^n x_i \boldsymbol{v}_j$. Substituting this expression into $\boldsymbol{\Sigma}\boldsymbol{V}^\star\boldsymbol{x} = 0$ gives that

$$(\sigma_1 x_1, \sigma_2 x_2, \cdots, \sigma_l x_l, \sigma_{l+1} x_{l+1}, \cdots, \sigma_n x_n)^T = 0.$$

Since $\sigma_j > 0$ for $1 \leq j \leq l$, it follows that $c_j = 0$ for these index values. Moreover, $\sigma_j = 0$ for $l + 1 \leq j \leq n$ implies that the coefficient c_j for these j-values are arbitrary. We can conclude that

$$\text{null}(\boldsymbol{A}) = \{\boldsymbol{x} = \sum_{i=1}^{l} c_i \boldsymbol{v}_i, c_i \in \mathbb{R}\} =< \boldsymbol{v}_{l+1}, \boldsymbol{v}_{l+2}, \cdots, \boldsymbol{v}_n > .$$

$\square$

2 Solutions

2.1 Exercise 5.1

First we compute the singular values σ_i by finding the eigenvalues of

$$\boldsymbol{A}^\star \boldsymbol{A} = \begin{pmatrix} 1 & 2 \\ 2 & 8 \end{pmatrix}.$$

The characteristic polynomial of $\boldsymbol{A}^\star \boldsymbol{A}$ is

$$\det(\boldsymbol{A}^\star \boldsymbol{A} - \lambda \boldsymbol{I}) = (1-\lambda)(8-\lambda) - 4 = 8 + \lambda^2 - 9\lambda - 4 = \lambda^2 - 9\lambda + 4,$$

so the singular values are

$$\sigma_{max} = \sqrt{\frac{9+\sqrt{65}}{2}}, \quad \sigma_{min} = \sqrt{\frac{9-\sqrt{65}}{2}}.$$

2.2 Exercise 5.2

Give a full SVD of $\boldsymbol{A}$ as $\boldsymbol{A} = \boldsymbol{U\Sigma V}^\star$ if $\mathrm{rank}(\boldsymbol{A}) = n$, then the claim is clearly true (sequence as $\{\boldsymbol{A}_1 = \boldsymbol{A}, \boldsymbol{A}_2 = \boldsymbol{A}, \cdots, \boldsymbol{A}_n = \boldsymbol{A}, \cdots\}$. Otherwise, if $\mathrm{rank}(\boldsymbol{A}) = r < n$, we can define a sequence of $\boldsymbol{A}$ as

$$\boldsymbol{A}_n = \boldsymbol{U} \begin{pmatrix} \sigma_1 & & & & & & \\ & \sigma_2 & & & & & \\ & & \ddots & & & & \\ & & & \sigma_r & & & \\ & & & & \frac{1}{n} & & \\ & & & & & \ddots & \\ & & & & & & \frac{1}{n} \end{pmatrix} \boldsymbol{V}^\star,$$

then we can easily verify that

$$\lim_{n\to\infty} \|\boldsymbol{A} - \boldsymbol{A}_n\|_2 = 0.$$

2.3 Exercise 5.3(a)

First we compute the singular values σ_i by finding the eigenvalues of

$$\boldsymbol{A}^\star \boldsymbol{A} = \begin{pmatrix} 104 & -72 \\ -72 & 146 \end{pmatrix}.$$

The characteristic polynomial of $\boldsymbol{A}^\star \boldsymbol{A}$ is

$$\det(\boldsymbol{A}^\star \boldsymbol{A} - \lambda \boldsymbol{I}) = (\lambda - 200)(\lambda - 50) = 0,$$

so the singular values are

$$\sigma_1 = \sqrt{\lambda_1} = 5\sqrt{2}, \quad \sigma_2 = \sqrt{\lambda_2} = 10\sqrt{2}.$$

For $\lambda_1 = 200$, we have

$$\boldsymbol{A}^\star \boldsymbol{A} - \lambda \boldsymbol{I} = \begin{pmatrix} -96 & -72 \\ -72 & -54 \end{pmatrix},$$

a unit vector in the kernel of the matrix is $\boldsymbol{v}_1 = (-3/5, 4/5)^T$. For $\lambda_2 = 50$, we have

$$\boldsymbol{A}^\star \boldsymbol{A} - \lambda \boldsymbol{I} = \begin{pmatrix} 54 & -72 \\ -72 & 96 \end{pmatrix},$$

a unit vector in the kernel of the matrix is $\boldsymbol{v}_2 = (4/5, 3/5)^T$. So at this point, we know that

$$\boldsymbol{A} = (\boldsymbol{u}_1, \boldsymbol{u}_2) \begin{pmatrix} -\frac{3}{5} & \frac{4}{5} \\ \frac{4}{5} & \frac{3}{5} \end{pmatrix}.$$

Further we can compute $\boldsymbol{u}_1, \boldsymbol{u}_2$ by $\sigma_i \boldsymbol{u}_i = \boldsymbol{A}\boldsymbol{v}_i$, which gives that

$$\boldsymbol{u}_1 = \begin{pmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{pmatrix}, \quad \boldsymbol{u}_2 = \begin{pmatrix} \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2}. \end{pmatrix}$$

So in its full glory the SVD is

$$\boldsymbol{A} = \begin{pmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \end{pmatrix} \begin{pmatrix} 10\sqrt{2} & 0 \\ 0 & 5\sqrt{2} \end{pmatrix} \begin{pmatrix} -\frac{3}{5} & \frac{4}{5} \\ \frac{4}{5} & \frac{3}{5} \end{pmatrix}.$$

2.4 Exercise 5.3(b)

Singular values:

$$\sigma_1 = 10\sqrt{2}, \quad \sigma_2 = 5\sqrt{2}$$

. Left singular vectors:

$$\boldsymbol{u}_1 = \begin{pmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{pmatrix}, \quad \boldsymbol{u}_2 = \begin{pmatrix} \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} \end{pmatrix}.$$

Right singular vectors:

$$\boldsymbol{v}_1 = \begin{pmatrix} \frac{-3}{5} \\ \frac{4}{5} \end{pmatrix}, \quad \boldsymbol{v}_2 = \begin{pmatrix} \frac{4}{5} \\ -\frac{3}{5} \end{pmatrix}.$$

2.5 Exercise 5.3(c)

We can compute the norms via the definition in *Lecture 3 Norms*.

$$\begin{aligned} \|\boldsymbol{A}\|_1 &= \max_{1 \le j \le 2} \|\boldsymbol{a}_j\| = 16, \\ \|\boldsymbol{A}\|_2 &= \|\boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^\star\|_2 = \|\boldsymbol{\Sigma}\|_2 = 10\sqrt{2}, \\ \|\boldsymbol{A}\|_F &= \|\boldsymbol{\Sigma}\|_F = \sqrt{50 + 200} = 5\sqrt{10}. \end{aligned}$$

2.6 Exercise 5.3(d)

We can compute $\boldsymbol{A}^{-1}$ via SVD as follows:

$$\begin{aligned}
\boldsymbol{A}^{-1} &= (\boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^\star)^{-1} = \boldsymbol{V}\boldsymbol{\Sigma}^{-1}\boldsymbol{U}^\star \\
&= \begin{pmatrix} -\frac{3}{5} & \frac{4}{5} \\ \frac{4}{5} & \frac{3}{5} \end{pmatrix} \begin{pmatrix} \frac{1}{10\sqrt{2}} & 0 \\ 0 & \frac{1}{5\sqrt{2}} \end{pmatrix} \begin{pmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \end{pmatrix} \\
&= \begin{pmatrix} -\frac{1}{20} & -\frac{11}{100} \\ \frac{1}{10} & -\frac{1}{50} \end{pmatrix}.
\end{aligned}$$

2.7 Exercise 5.3(e)

The characteristic polynomial of A is

$$\det(\boldsymbol{A} - \lambda \boldsymbol{I}) = \lambda^2 - 3\lambda + 100 = 0,$$

we can get that

$$\lambda_1 = \frac{3 + i\sqrt{391}}{2}, \quad \lambda_2 = \frac{3 - i\sqrt{391}}{2}.$$

2.8 Exercise 5.3(f)

We can verify the claim as follows:

$$\begin{aligned}
\det(\boldsymbol{A}) &= -2 \times 5 + 10 \times 11 = 100 = \lambda_1 \cdot \lambda_2 \\
|\det(\boldsymbol{A})| &= 100 = 10\sqrt{2} \times 5\sqrt{2} = \sigma_1 \cdot \sigma_2.
\end{aligned}$$

2.9 Exercise 5.3(g)

It is clear that the ellipsoid is a rotation of another ellipsoid $\mathcal{E}$ whose equation are

$$\frac{x^2}{\sigma_1^2} + \frac{y^2}{\sigma_2^2} = 1$$

and hence we can compute the area of $\mathcal{E}$ as the area of the original ellipsoid. We first let

$$y = f(x) = \sigma_2 \sqrt{1 - \frac{x^2}{\sigma_1^2}},$$

then

$$\begin{aligned}
\text{Area}(\mathcal{E}) &= 4\int_0^1 f(x)dx \\
&= 4\int_0^1 \sigma_2\sqrt{1-\frac{x^2}{\sigma_1^2}}dx \\
&= 4\int_0^{\frac{\pi}{2}} \sigma_2\sigma_1\cos^2\theta d\theta, \quad (x=\sigma_1\sin\theta) \\
&= 4\int_0^{\frac{\pi}{2}} \frac{\sigma_1\sigma_2}{2}(1+\cos(2\theta))d\theta \\
&= 4\times\frac{\sigma_1\sigma_2\pi}{4} = \sigma_1\sigma_2\pi.
\end{aligned}$$

Thus the area of the original ellipsoid is $\sigma_1\sigma_2\pi = 100\pi$.

2.10 Exercise 5.4

Let $\boldsymbol{A}$ be an $m\times m$ complex matrix with singular values $\sigma_1, \sigma_2, \cdots, \sigma_m$, and let $\boldsymbol{u}_i$ and $\boldsymbol{v}_i$ be the corresponding left and right singular vectors. If we define

$$\boldsymbol{\beta}_i = \begin{pmatrix} a\boldsymbol{v}_i \\ b\boldsymbol{u}_i \end{pmatrix} \in \mathbb{C}^{2m},$$

and the matrix

$$\boldsymbol{B} = \begin{pmatrix} \boldsymbol{0} & \boldsymbol{A}^\star \\ \boldsymbol{A} & \boldsymbol{0} \end{pmatrix},$$

then we claim that the eigenvalues of $\boldsymbol{B}$ are $\pm\sigma_1, \pm\sigma_2, \ldots, \pm\sigma_m$. Consider the product:

$$\boldsymbol{B}\boldsymbol{\beta}_i = \begin{pmatrix} b\sigma_i\boldsymbol{v}_i \\ a\sigma_i\boldsymbol{u}_i \end{pmatrix}$$

Given that $\boldsymbol{\beta}_i$ is an eigenvector of $\boldsymbol{B}$ with eigenvalue λ, we deduce the following systems:

For $\lambda = \sigma_i$:

$$\begin{cases} a = b \\ b = a \end{cases}$$

And for $\lambda = -\sigma_i$:

$$\begin{cases} a = -b \\ b = -a \end{cases}$$

From these systems, it is evident that $a = b$ and $a = -b$, respectively.

To construct the eigen decomposition of $\boldsymbol{B}$, we define the matrix:

$$\boldsymbol{S} = \begin{pmatrix} a\boldsymbol{v}_1, & a\boldsymbol{v}_2, & \cdots, & a\boldsymbol{v}_m, & a\boldsymbol{v}_1, & a\boldsymbol{v}_2, & \cdots, & a\boldsymbol{v}_m \\ a\boldsymbol{u}_1, & a\boldsymbol{u}_2, & \cdots, & a\boldsymbol{u}_m, & -a\boldsymbol{u}_1, & -a\boldsymbol{u}_2, & \cdots, & -a\boldsymbol{u}_m \end{pmatrix}$$

as well as

$$\boldsymbol{\Lambda} = \begin{pmatrix} \sigma_1 & & & & & & & \\ & \sigma_2 & & & & & & \\ & & \ddots & & & & & \\ & & & \sigma_m & & & & \\ & & & & -\sigma_1 & & & \\ & & & & & -\sigma_2 & & \\ & & & & & & \ddots & \\ & & & & & & & -\sigma_m \end{pmatrix},$$

then there is

$$\boldsymbol{B}\boldsymbol{S} = \boldsymbol{S}\boldsymbol{\Lambda},$$

But at this time $\boldsymbol{S}$ is not necessarily invertible. Next, we make $\boldsymbol{S}$ an orthogonal matrix by taking an appropriate a, so that $\boldsymbol{S}$ is invertible. It is not difficult to calculate, assuming that $\boldsymbol{S}$ is orthogonal, that is, $\boldsymbol{S}^\star\boldsymbol{S} = \boldsymbol{S}\boldsymbol{S}^\star = \boldsymbol{I}$, then

$$a^2 + a^2 = 1,$$

We can get $a = \pm\frac{1}{\sqrt{2}}$, without loss of generality, we make $a = \frac{1}{\sqrt{2}}$, then S is an orthogonal matrix, so it is invertible, then the eigendecomposition of B is

$$\boldsymbol{B} = \boldsymbol{S}^{-1}\boldsymbol{\Lambda}\boldsymbol{S}.$$

Remark. *In fact, this proof is not perfect. We did not detail how to find the eigenvalue of $\boldsymbol{B}$, that is, the root of the characteristic polynomial corresponding to $\boldsymbol{B}$ has and only $\pm\sigma_i$, which needs to be improved later. If readers have any good ideas, welcome to contact me.*