Peeter Joot
peeter.joot@gmail.com

Singular Value Decomposition

We've been presented with the definition of the singular value decomposition or SVD , but not how to compute it.

Recall that the definition was

Definition 1.1: Singular value decomposition (SVD)

Given $M \in \mathbb{R}^{m \times n}$, we can find a representation of M

$$M = U \Sigma V^\mathrm{T}, \tag{1.1}$$

where U and V are orthogonal matrices such that $U^\mathrm{T} U = 1$, and $V^\mathrm{T} V = 1$, and

$$\Sigma = \begin{bmatrix} \sigma_1 & & & & & & \\ & \sigma_2 & & & & & \\ & & \ddots & & & & \\ & & & \sigma_r & & & \\ & & & & 0 & & \\ & & & & & \ddots & \\ & & & & & & 0 \end{bmatrix} \tag{1.2}$$

The values $\sigma_i, i \le \min(n, m)$ are called the singular values of M. The singular values are subject to the ordering

$$\sigma_1 \ge \sigma_2 \ge \cdots \ge 0 \tag{1.3}$$

If r is the rank of M, then the σ_r above is the minimum non-zero singular value (but the zeros are also called singular values).

Here I'll review some of the key ideas from the MIT OCW SVD lecture by Prof. Gilbert Strang [1]. This is largely to avoid having to rewatch this again in a few years as I just did.

The idea behind the SVD is to find an orthogonal basis that relates the image space of the transformation, as well as the basis for the vectors that the transformation applies to. That is a relation of the form

$$\mathbf{u}_i = M\mathbf{v}_j \tag{1.4}$$

Where $\mathbf{v}_j$ are orthonormal basis vectors, and $\mathbf{u}_i$ are orthonormal basis vectors for the image space. Let's suppose that such a set of vectors can be computed and that M has a representation of the desired form

$$M = U\Sigma V^* \tag{1.5}$$

where

$$U = \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_m \end{bmatrix}, \tag{1.6}$$

and

$$V = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n \end{bmatrix}. \tag{1.7}$$

By left or right multiplication of M with its (conjugate) transpose, we will see that the decomposed form of these products have a very simple form

$$M^*M = V\Sigma^*U^*U\Sigma V^* = V\Sigma^*\Sigma V^*, \tag{1.8a}$$

$$MM^* = U\Sigma V^*V\Sigma^*U^* = U\Sigma\Sigma^*U^*. \tag{1.8b}$$

Because Σ is diagonal, the products $\Sigma^*\Sigma$ and $\Sigma\Sigma^*$ are also both diagonal, and populated with the absolute squares of the singular values that we have presumed to exist. Because $\Sigma\Sigma^*$ is an $m \times m$ matrix, whereas $\Sigma^*\Sigma$ is an $n \times n$ matrix, so the numbers of zeros in each of these will differ, but each will have the structure

$$\Sigma^*\Sigma \sim \Sigma\Sigma^* \sim \begin{bmatrix} |\sigma_1|^2 & & & & & & \\ & |\sigma_2|^2 & & & & & \\ & & \ddots & & & & \\ & & & |\sigma_r|^2 & & & \\ & & & & 0 & & \\ & & & & & \ddots & \\ & & & & & & 0 \end{bmatrix} \tag{1.9}$$

This shows us one method of computing the singular value decomposition (for full rank systems), because we need only solve for the eigensystem of either MM^* or M^*M to find the singular values, and one of U or V. To form U we will be able to find r orthonormal eigenvectors of MM^* and then supplement that with a mutually orthonormal set of vectors from the Null space of M. We can form Σ by taking the square roots of the eigenvalues of MM^*. With both Σ and U computed, we can then compute V by inversion. We could similarly solve for Σ and V by computing the eigensystem of M^*M and similarly supplementing those eigenvectors with vectors from the null space, and then compute U by inversion.

Let's work the examples from the lecture to see how this works.

Example 1.1: Full rank 2x2 matrix

In the lecture the SVD decomposition is computed for

$$M = \begin{bmatrix} 4 & 4 \\ 3 & -3 \end{bmatrix} \tag{1.10}$$

For this we have

$$MM^* = \begin{bmatrix} 32 & 0 \\ 0 & 18 \end{bmatrix} \tag{1.11a}$$

$$M^*M = \begin{bmatrix} 25 & 7 \\ 7 & 25 \end{bmatrix} \tag{1.11b}$$

The first is already diagonalized so $U = I$, and the singular values are found by inspection $\{\sqrt{32}, \sqrt{18}\}$, or

$$\Sigma = \begin{bmatrix} \sqrt{32} & 0 \\ 0 & \sqrt{18} \end{bmatrix} \tag{1.12}$$

Because the system is full rank we can solve for V by inversion

$$\Sigma^{-1} U^* M = V^*, \tag{1.13}$$

or

$$V = M^* U \left(\Sigma^{-1}\right)^*. \tag{1.14}$$

In this case that is

$$V = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \tag{1.15}$$

We could alternately compute V directly by diagonalizing M^*M. We see that the eigenvectors are $(1, \pm 1)/\sqrt{2}$, with respective eigenvalues $\{32, 18\}$.

This gives us eq. (1.12) and eq. (1.15). Again, because the system is full rank, we can compute U by inversion. That is

$$U = MV\Sigma^{-1}. \tag{1.16}$$

Carrying out this calculation recovers $U = I$ as expected. Looks like I used a different matrix than Prof. Strang used in his lecture (alternate signs on the 3's). He had some trouble that arrived from independent calculation of the respective eigenspaces. Calculating one from the other avoids that trouble since there are different signed eigenvalues that can be chosen, and we are looking for specific mappings between the eigenspaces that satisfy the $\mathbf{u}_i = M\mathbf{v}_j$ constraints encoded by the relationship $M = U\Sigma V^*$.

Let's work a non-full rank example, as in the lecture.

Example 1.2: 2x2 matrix without full rank

How about

$$M = \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix}. \tag{1.17}$$

For this we have

$$M^* M = \begin{bmatrix} 5 & 5 \\ 5 & 5 \end{bmatrix} \tag{1.18a}$$

$$M M^* = \begin{bmatrix} 2 & 4 \\ 4 & 8 \end{bmatrix} \tag{1.18b}$$

For which the non-zero eigenvalue is 10 and the corresponding eigenvalue is

$$\mathbf{v} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}. \tag{1.19}$$

This gives us

$$\Sigma = \begin{bmatrix} \sqrt{10} & 0 \\ 0 & 0 \end{bmatrix} \tag{1.20}$$

Since we require V to be orthonormal, there is only one choice (up to a sign) for the vector from the null space.

Let's try

$$V = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \tag{1.21}$$

We find that MM^* has eigenvalues $\{10, 0\}$ as expected. The eigenvector for the non-zero eigenvalue is found to be

$$\mathbf{u} = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 \\ 2 \end{bmatrix}. \tag{1.22}$$

It's easy to expand this to an orthonormal basis, but do we have to pick a specific sign relative to the choice that we've made for V?

Let's try

$$U = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 & 2 \\ 2 & -1 \end{bmatrix}. \tag{1.23}$$

Multiplying out $U \Sigma V^*$ we have

$$
\begin{aligned}
\frac{1}{\sqrt{5}}\begin{bmatrix}1 & 2\\ 2 & -1\end{bmatrix}\begin{bmatrix}\sqrt{10} & 0\\ 0 & 0\end{bmatrix}\frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix} &= \begin{bmatrix}1 & 2\\ 2 & -1\end{bmatrix}\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix}\\
&= \begin{bmatrix}1 & 0\\ 2 & 0\end{bmatrix}\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix}\\
&= \begin{bmatrix}1 & 1\\ 2 & 2\end{bmatrix}.
\end{aligned} \tag{1.24}
$$

It appears that this works, although we haven't demonstrated why that should be, and we could have gotten lucky with signs. There's some theoretical work to do here, but let's leave that for another day (or rely on software to do this computational task).

Bibliography

[1] Gilbert Strang. *Singular Value Decomposition*, 2014. URL http://ocw.mit.edu/courses/mathematics/18-06-linear-algebra-spring-2010/video-lectures/lecture-29-singular-value-decomposition/. [Online; accessed 30-Sept-2014]. 1