Calculating the Singular Value Decomposition

Suppose we have a matrix A of dimension $m \times n$ where $m > n$ and A has rank r.
We can find the singular value decomposition (SVD)

$$A = U\Sigma V^* = \sum_{i=1}^{n} \sigma_i \vec{u}_i \vec{v}_i^*$$

with the following steps.

(1) Find the eigenvalues λ_i of A^*A and order them such that $\lambda_1 \geq \cdots \geq \lambda_r > 0$ and $\lambda_{r+1} = \cdots = \lambda_n = 0$.

(2) Find the orthonormal eigenvectors of A^*A, so that

$$A^*A\vec{v}_i = \lambda_i \vec{v}_i, \qquad i = 1, \ldots, r$$

Note that the vectors must be orthonormal, that is $\langle \vec{v}_i, \vec{v}_i \rangle = 1$ and $\langle \vec{v}_i, v_j \rangle = 0$ for $i \neq j$.

(3) Let $\sigma_i = \sqrt{\lambda_i}$ and set

$$\vec{u}_i = \frac{A\vec{v}_i}{\sigma_i}, \qquad i = 1, \ldots, r$$

(4) If $r < n$ then we complete the U and V matrices by adding vectors $\vec{u}_{r+1}, \ldots, \vec{v}_m$ and $\vec{v}_{r+1}, \ldots, \vec{v}_n$ to create an orthonormal bases for $\mathbb{R}^m$ and $\mathbb{R}^n$.

1 SVD and Fundamental Subspaces

Define the matrix

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

a) Find the SVD of A.

Answer

First, compute $A^*A = \begin{bmatrix} 9 & -9 \\ -9 & 9 \end{bmatrix}$. The eigenvalues of A^*A are 18 and 0, with corresponding unit eigenvectors

$$\vec{v}_1 = \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}, \vec{v}_2 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}.$$

Therefore, A is rank 1 and has one singular vector $\sqrt{18} = 3\sqrt{2}$
We obtain

$$\vec{u}_1 = \frac{1}{\sigma_1} A\vec{v}_1 = \frac{1}{3\sqrt{2}} \begin{bmatrix} 1 & -1 \\ -2 & 2 \\ 2 & -2 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix} = \begin{bmatrix} 1/3 \\ -2/3 \\ 2/3 \end{bmatrix}$$

and A can be decomposed as

$$A = 3\sqrt{2} \begin{bmatrix} 1/3 \\ -2/3 \\ 2/3 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}$$

b) Find the rank of A.

Answer

A has 1 nonzero singular value. So A has rank 1.

c) Find a basis for the kernel (or nullspace) of A.

Answer

$$\ker(A) = \operatorname{span} \left\{ \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} \right\}$$

d) Find a basis for the range (or columnspace) of A.

Answer

$$\operatorname{range}(A) = \operatorname{span}\{\vec{u}_1\} = \operatorname{span} \left\{ \begin{bmatrix} 1/3 \\ -2/3 \\ 2/3 \end{bmatrix} \right\}$$

e) Repeat parts (a) - (d), but instead, create the SVD of $B = A^*$. What are the relationships between the answers for A and the answers for $B = A^*$?

Answer

$$B = A^* = \begin{bmatrix} 1 & -2 & 2 \\ -1 & 2 & -2 \end{bmatrix} \implies B^*B = AA^* = \begin{bmatrix} 2 & -4 & 4 \\ -4 & 8 & -8 \\ 4 & -8 & 8 \end{bmatrix}$$

$$\lambda = 18, 0, 0 \implies \text{Rank}(A^*) = 1$$

$$\vec{v}_1 = \begin{bmatrix} 1/3 \\ -2/3 \\ 2/3 \end{bmatrix}$$

Then we compute $\vec{u}_1$

$$\vec{u}_1 = \frac{B\vec{v}_1}{\sigma_1} = \frac{1}{3\sqrt{2}} \begin{bmatrix} 1 & -2 & 2 \\ -1 & 2 & -2 \end{bmatrix} \begin{bmatrix} 1/3 \\ -2/3 \\ 2/3 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}$$

At this point, we already know the rank is 1. The column space is also formed by the $\vec{u}$ vector

$$\text{range}(A^*) = \text{span}\left\{ \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix} \right\}$$

Notice how the Range of A^* is in fact orthogonal to the Nul (A).

Two vectors that create a basis for the nullspace of A^* are

$$\begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ -1 \end{bmatrix}$$

Notice how the basis vectors for the nullspace of A^* are also orthogonal to the basis vectors of Col (A). If wanted to create an orthonormal basis for the nullspace of A^*, we could perform Gram-Schmidt to get

$$\vec{v}_2 = \frac{1}{\sqrt{5}} \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} \quad \vec{v}_3 = \frac{1}{\sqrt{45}} \begin{bmatrix} 2 \\ -4 \\ -5 \end{bmatrix}$$

Note, if we had just noticed

$$A^* = (U\Sigma V^*)^* = V\Sigma^* U^*$$

We could've skipped many steps for SVD calculation.

2 Understanding the SVD

We can compute the SVD for a wide matrix A with dimension m x n where $n > m$ using A^*A with the method described above. However, when doing so you may realize that A^*A is much larger than AA^* for such wide matrices. This makes it more efficient to find the eigenvalues for AA^*. In this question we will explore how to compute the SVD using AA^* instead of A^*A.

a) What are the dimensions of AA^* and A^*A.

Answer

dim(AA^*) = m x m
dim(A^*A) = n x n

b) Given that the $A = U\Sigma V^*$, find a symbolic expression for AA^*.

Answer

$$AA^* = U\Sigma V^* V \Sigma^* U^*$$

$$V^*V = I$$

$$AA^* = U\Sigma\Sigma^* U^*$$

c) Using the solution to the previous part explain how to find U and Σ from AA^*.

Answer

Knowing that AA^* is a symmetric matrix, we know that its normalized eigenvectors will be orthonormal.

From the properties of the SVD we know that U is an orthonormal matrix of dimension m x m and $\Sigma\Sigma^*$ is an m x m diagonal matrix with the entries on the diagonal being ${\sigma_i}^2$.

Using the above information we can see that we can calculate U by diagonalizing the symmetric matrix AA^*. By the spectral theorem for real symmetric matrices, we will get an orthonormal basis of eigenvectors. The square root of the corresponding eigenvalues of AA^* will give us the singular values σ_i (You can construct Σ by putting these on the diagonal of a zero matrix with the same dimensions as A) and the corresponding eigenvectors will form the U matrix.

d) Now that we have found the singular values σ_i and the corresponding vectors $\vec{u}_i$ in the matrix U, devise a way to find the corresponding vectors $\vec{v}_i$ in matrix V.

Answer

$$\vec{v}_i = \frac{V\Sigma^* U^* \vec{u}_i}{\sigma_i} = \frac{A^* \vec{u}_i}{\sigma_i}$$

e) Now we have a way to find the vectors $\vec{v}_i$ in matrix V, verify that they are orthonormal.

Answer

To verify that $\vec{v}_i$ in matrix V are orthonormal we show that:

a) $\vec{v}_i$ are orthogonal to one another
b) each $\vec{v}_i$ has norm 1.

Orthogonality:
To show orthogonality we must show that any two vectors $\vec{v}_i = \frac{A^*\vec{u}_i}{\sigma_i}$ and $\vec{v}_j = \frac{A^*\vec{u}_j}{\sigma_j}$ with $i \neq j$ have an inner product of zero.

$$\begin{align}
\langle \vec{v}_i, \vec{v}_j \rangle &= \langle \frac{A^*\vec{u}_i}{\sigma_i}, \frac{A^*\vec{u}_j}{\sigma_j} \rangle \tag{1}\\
&= \frac{1}{\sigma_i\sigma_j} \langle AA^*\vec{u}_i, \vec{u}_j \rangle \tag{2}\\
&= \frac{1}{\sigma_i\sigma_j} \langle \sigma_i^2\vec{u}_i, \vec{u}_j \rangle \tag{3}\\
&= \frac{\sigma_i^2}{\sigma_i\sigma_j} \langle \vec{u}_i, \vec{u}_j \rangle \tag{4}\\
&= 0 \tag{5}
\end{align}$$

This is because we know that $\vec{u}_i$ and $\vec{u}_j$ are orthonormal as they are eigenvectors of a symmetric matrix AA^*.
Thus for all $i \neq j$

$$\langle \vec{v}_i, \vec{v}_j \rangle = 0$$

Norm of 1: Follow the steps above with $i = j$ to see

$$\begin{align}
\langle \vec{v}_i, \vec{v}_i \rangle &= \langle \frac{A^*\vec{u}_i}{\sigma_i}, \frac{A^*\vec{u}_i}{\sigma_i} \rangle \tag{6}\\
&= \frac{1}{\sigma_i^2} \langle AA^*\vec{u}_i, \vec{u}_i \rangle \tag{7}\\
&= \frac{1}{\sigma_i^2} \langle \sigma_i^2\vec{u}_i, \vec{u}_i \rangle \tag{8}\\
&= \langle \vec{u}_i, \vec{u}_i \rangle \tag{9}\\
&= 1 \tag{10}
\end{align}$$

f) Now that we have found $\vec{v}_i$ you may notice that we only have $m < n$ vectors of dimension n. This is not enough for a basis. How would you complete the m vectors to form an orthonormal basis?

Answer

Gram Schmidt.

Just add in the standard basis for n-dimensional space, and orthonormalize. We know that this collection of $n + m$ vectors spans the whole space, and so after orthonormalization, we will have a collection of orthonormal vectors that spans the whole space. Along the way, some vectors will be found to be linearly dependent on those that came before — that is fine, discard these. At the end, we will have n orthonormal vectors, the first set of which are the original $\vec{v}_i$.

g) Using the previous parts of this question and what you learned from lecture write out a procedure on how to find the SVD for any matrix.

Answer

We calculate the SVD of matrix A as follows.

a) Pick A^*A or AA^* — whichever one is smaller.

b) (i) If using A^*A, find the eigenvalues λ_i of A^*A and order them, so that $\lambda_1 \geq \cdots \geq \lambda_r > 0$ and $\lambda_{r+1} = \cdots = \lambda_n = 0$.

If using AA^*, find its eigenvalues $\lambda_1, \ldots, \lambda_m$ and order them the same way.

(ii) If using A^*A, find orthonormal eigenvectors $\vec{v}_i$ such that

$$A^*A\vec{v}_i = \lambda_i\vec{v}_i, \quad i = 1, \ldots, r$$

If using AA^*, find orthonormal eigenvectors $\vec{u}_i$ such that

$$AA^*\vec{u}_i = \lambda_i\vec{u}_i, \quad i = 1, \ldots, r$$

(iii) Set $\sigma_i = \sqrt{\lambda_i}$.

If using A^*A, obtain $\vec{u}_i$ from $\vec{u}_i = \frac{1}{\sigma_i}A\vec{v}_i, \quad i = 1, \ldots, r$.

If using AA^*, obtain $\vec{v}_i$ from $\vec{v}_i = \frac{1}{\sigma_i}A^*\vec{u}_i, \quad i = 1, \ldots, r$.

c) If you want to completely construct the U or V matrix, complete the basis (or columns of the appropriate matrix) using Gram-Schmidt to get a full orthonormal matrix.

The full matrix form of SVD is taken to better understand the matrix A in terms of the 3 nice matrices U, Σ, V. Often in practice, we do not completely construct the U and V matrices. After all, in many applications, we don't need all the vectors.