Singular Value Decomposition

Justin Wyss-Gallifent

October 25, 2021

8.1 Introduction

Factoring a matrix means writing the matrix as a product of other matrices. For example if we have a matrix A and we manage to write it as $A = BC$ for some B and C then we've factored it into a product of two other matrices.

There are many ways to factor a matrix and many of them are extremely useful. For example if a matrix A is diagonalizable then we can write the matrix as $A = PDP^{-1}$. This is useful because the entries in D are the eigenvalues and the columns of P are the eigenvectors.

Another really useful factorization of a matrix is the *singular value decomposition* which is a way of factoring a matrix which is used in areas like data compression, matrix approximation, pseudoinverses, signal analysis, handwriting and facial recognition, the list goes on.

In this chapter we define the *singular value decomposition* and see what mathematical properties it has.

8.2 Definitions

Definition 8.2.0.1. For any $m \times n$ real matrix A the *singular value decomposition* (SVD) of A is a factorization of A as:

$$A = U \Sigma V^T$$

where:

- U is an $m \times m$ orthogonal matrix.

- Σ is an $m \times n$ rectangular diagonal matrix with the diagonal entries $\sigma_1, ..., \sigma_{\min(m,n)}$ all nonnegative and in nonincreasing order. These last requirements aren't strictly necessary for the definition but they are commonly applied and the discussion is easier if we assume it.

- V is an $n \times n$ orthogonal matrix.

Just to be sure everything's clear, here are some auxilliary definitions:

Definition 8.2.0.2. An $n \times n$ matrix U is *orthogonal* if its columns are unit vectors all perpendicular to one another. This is equivalent to saying that the columns form an orthonormal basis of $\mathbb{R}^n$ and is equivalent to having $UU^T = U^T U = I$.

Definition 8.2.0.3. An $m \times n$ *rectangular diagonal matrix* (which may or may not be square) is a matrix in which the only nonzero entries can be in the $(1,1)$,

$(2, 2)$, ... positions. If there happen to be more rows than columns then the additional rows are all zeros and likewise if there are more columns than rows.

8.3   Constructing the SVD

In applications we'll use technology to find the SVD but it's worth stepping through some of the details to fully understand what's going on in the background.

Several observations will allow us to see how the SVD may be easily constructed.

8.3.1   Observation 1

First observe that if $k \leq \min(m, n)$ of the σ_i are nonzero then the product $U\Sigma$ (and hence $U\Sigma V^T$) depends only on the first k columns of U and the product ΣV^T (and hence $U\Sigma V^T$) depends only on the first k columns on V. These follow from the definition of matrix multiplication.

Second observe that piggybacking off this we have:

$$A = U\Sigma V^T \iff U^T A = U^T U\Sigma V^T \iff U^T A = \Sigma V^T \iff A^T U = V\Sigma^T$$

$$\iff \forall 1 \leq i \leq k,\ A^T \bar{u}_i = \sigma_i \bar{v}_i$$

and that:

$$A = U\Sigma V^T \iff AV = U\Sigma V^T V \iff AV = U\Sigma$$

$$\iff \forall 1 \leq i \leq k,\ A\bar{v}_i = \sigma_i \bar{u}_i$$

Consequently if we can find some $k \leq \min(m, n)$ along with orthonormal $\bar{u}_1, ..., \bar{u}_k$ and orthonormal $\bar{v}_1, ..., \bar{v}_k$ satisfying $A^T \bar{u}_i = \sigma_i \bar{v}_i$ and $A\bar{v}_i = \sigma_i \bar{u}_i$ then we can assign Σ and fill out U and V as to be orthogonal and we are done.

8.3.2   Observation 2

Let's look at the two matrices AA^T and $A^T A$.

First, note that the matrix AA^T is $m \times m$ and is symmetric (this is obvious) and by the Spectral Theorem (proof omitted) has m eigenvalues (counting multiplicity) and m orthogonal unit eigenvectors. The same is true for $A^T A$ with n instead of m.

Second, note that the eigenvalues for AA^T and for $A^T A$ are all nonnegative. To see this note that if $(\lambda, \bar{v})$ is an eigenpair for $A^T A$ with $\bar{v}$ a unit vector then

$$||A\bar{v}||^2 = (A\bar{v})^T A\bar{v} = \bar{v}^T A^T A\bar{v} = \bar{v}^T \lambda \bar{v} = ||\bar{v}||^2 \lambda = \lambda$$

A similar argument holds for AA^T. Thus the eigenvalues can be written as squares of positive real numbers.

Third, note that we have a bijection between the eigenpairs of AA^T and $A^T A$ corresponding to nonzero eigenvalues and using unit eigenvectors. This bijection is given by:

$$\phi : \text{Eigenpairs of } AA^T \to \text{Eigenpairs of } A^T A$$

$$(\sigma^2, \bar{u}) \mapsto \left(\sigma^2, \frac{1}{\sigma} A^T \bar{u} \right)$$

with inverse:

$$\psi : \text{Eigenpairs of } A^T A \to \text{Eigenpairs of } AA^T$$

$$(\sigma^2, \bar{v}) \mapsto \left(\sigma^2, \frac{1}{\sigma} A \bar{v} \right)$$

Straightforward calculations show that both ϕ and ψ actually map eigenpairs to eigenpairs, that they are inverses, and that both map unit eigenvectors to unit eigenvectors.

Definition 8.3.2.1. The eigenvectors of AA^T are called the *left singular vectors of A* and The eigenvectors of AA^T are called the *right singular vectors of A* .

8.3.3   Observation 3

We now know that the for the j eigenvalues which are positive (nonzero) and common to AA^T and $A^T A$ that the eigenpairs may be grouped.

$$(\sigma_1^2, \bar{u}_1, \bar{v}_1), ..., (\sigma_k^2, \bar{u}_j, \bar{v}_j)$$

such that for each triplet we have:

$$\bar{v}_i = \tfrac{1}{\sigma_i} A^T \bar{u}_i \text{ and } \bar{u}_i = \tfrac{1}{\sigma_i} A \bar{v}_i$$

which is equivalent to:

$$A^T \bar{u}_i = \sigma_i \bar{v}_i \text{ and } A \bar{v}_i = \sigma_i \bar{u}_i$$

This is precisely what was needed to satisfy $A = U \Sigma V^T$ with $k = j$.

8.3.4   Construction in Brief

The practical result here is that we can construct the SVD as follows:

1. Find the positive (nonzero) eigenvalues for AA^T and A^TA. These eigenvalues will be shared (we can check either). Denote these σ_i^2 for $1 \leq i \leq \min(m, n)$.

2. For each eigenvalue σ_i^2, find a unit eigenvector $\bar{u}_i \in \mathbb{R}^m$ of AA^T and a unit eigenvector $\bar{v}_i \in \mathbb{R}^n$ of A^TA satisfying $A^T\bar{u}_i = \sigma_i\bar{v}_i$ and $A\bar{v}_i = \sigma_i\bar{u}_i$. Multiplicity must be taken into account here; that is, if σ_i^2 is repeated then we must choose multiple $\bar{u}_i$ to span the eigenspace. We can if we wish just find the $\bar{u}_i$ and assign the $\bar{v}_i$.

3. Construct the matrix Σ using the σ_i and construct the matrices U and V using the $\bar{u}_i$ and $\bar{v}_i$ respectively in the corresponding order as the σ_i. Fill out U and V to be orthogonal if necessary. We can do this using remaining unit eigenvectors (corresponding to the eigenvalue 0) of AA^T or A^TA respectively.

Again - in practice we will let technology do this but it's worth stepping through one example:

Example 8.1. For example if

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

Since A is 2×3 we know U will be 2×2, Σ will be 2×3, and V will be 3×3.

We calculate:

$$AA^T = \begin{bmatrix} 5 & 2 \\ 2 & 2 \end{bmatrix}$$

which has eigenvalues $\{6, 1\}$ and we calculate:

$$A^TA = \begin{bmatrix} 1 & 2 & 0 \\ 2 & 5 & -1 \\ 0 & -1 & 1 \end{bmatrix}$$

which has eigenvalues $\{6, 1, 0\}$

The shared eigenvalues are $\{\sigma_1^2, \sigma_2^2\} = \{6, 1\}$ and so the singular values are $\{\sigma_1, \sigma_2\} = \{\sqrt{6}, \sqrt{1}\} = \{2.4495, 1\}$.

This tells us that

$$\Sigma = \begin{bmatrix} 2.4495 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

To find U and V we examine the common eigenvalues.

For the eigenvalue $\sigma_1^2 = 6$ we find a unit eigenvector of AA^T:

$$\bar{u}_1 = \begin{bmatrix} -0.8944 \\ -0.4472 \end{bmatrix}$$

and we find a unit eigenvector of $A^T A$:

$$\bar{v}_1 = \begin{bmatrix} -0.3651 \\ -0.9129 \\ 0.1826 \end{bmatrix}$$

Observe that $A^T \bar{u}_1 = \sigma_1 \bar{v}_1$ and $A\bar{v}_1 = \sigma_1 \bar{u}_1$.

For the eigenvalue $\sigma_2^2 = 1$ we find a unit eigenvector of AA^T:

$$\bar{u}_2 = \begin{bmatrix} 0.4472 \\ -0.8944 \end{bmatrix}$$

and we find a unit eigenvector of $A^T A$:

$$\bar{v}_2 = \begin{bmatrix} -0.4472 \\ 0 \\ -0.8944 \end{bmatrix}$$

Observe that $A^T \bar{u}_2 = \sigma_2 \bar{v}_2$ and $A\bar{v}_2 = \sigma_2 \bar{u}_2$.

So far then we have:

$$\begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & -1 \end{bmatrix} = \underbrace{\begin{bmatrix} -0.8944 & -0.4472 \\ -0.4472 & 0.8944 \end{bmatrix}}_{U} \underbrace{\begin{bmatrix} 2.4495 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}}_{\Sigma} \underbrace{\begin{bmatrix} -0.3651 & -0.4472 & ? \\ -0.9129 & 0 & ? \\ 0.1826 & -0.8944 & ? \end{bmatrix}^T}_{V^T}$$

We need to fill in V. We can do this by observing that $A^T A$ also had the eigenpair:

$$\left(0, \begin{bmatrix} 0.8165 \\ -0.4082 \\ -0.4082 \end{bmatrix} \right)$$

Thus we have:

$$\begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & -1 \end{bmatrix} = \underbrace{\begin{bmatrix} -0.8944 & -0.4472 \\ -0.4472 & 0.8944 \end{bmatrix}}_{U} \underbrace{\begin{bmatrix} 2.4495 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}}_{\Sigma} \underbrace{\begin{bmatrix} -0.3651 & -0.4472 & 0.8165 \\ -0.9129 & 0 & -0.4082 \\ 0.1826 & -0.8944 & -0.4082 \end{bmatrix}^T}_{V^T}$$

Example 8.2. For example if

$$A = \begin{bmatrix} 1 & 3 \\ 0 & 2 \\ 1 & 1 \\ -1 & 2 \end{bmatrix}$$

Since A is 4×2 we know U will be 4×4, Σ will be 4×2, and V will be 2×2.

We calculate:

$$AA^T = \begin{bmatrix} 10 & 6 & 4 & 5 \\ 6 & 4 & 2 & 4 \\ 4 & 2 & 2 & 1 \\ 5 & 4 & 1 & 5 \end{bmatrix}$$

which has eigenvalues $\{18.2621, 2.7379, 0\}$ where the 0 has multiplicity 2 and we calculate:

$$A^T A = \begin{bmatrix} 3 & 2 \\ 2 & 18 \end{bmatrix}$$

which has eigenvalues $\{18.2621, 2.7379\}$.

The shared eigenvalues are $\{\sigma_1^2, \sigma_2^2\} = \{18.2621, 2.7379\}$. and so the singular values are $\{\sqrt{18.2621}, \sqrt{2.7379}\} = \{4.2734, 1.6547\}$.

This tells us that:

$$\Sigma = \begin{bmatrix} 4.2734 & 0 \\ 0 & 1.6547 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

To find U and V we examine the common eigenvalues.

For the eigenvalue $\sigma_1^2 = 18.2621$ we find a unit eigenvector of AA^T:

$$\bar{u}_1 = \begin{bmatrix} -0.7265 \\ -0.4640 \\ -0.2624 \\ -0.4336 \end{bmatrix}$$

7

and we find a unit eigenvector of $A^T A$:

$$\bar{v}_1 = \begin{bmatrix} -0.1299 \\ -0.9915 \end{bmatrix}$$

Observe that $A^T \bar{u}_1 = \sigma_1 \bar{v}_1$ and $A\bar{v}_1 = \sigma_1 \bar{u}_1$.

For the eigenvalue $\sigma_1^2 = 2.7379$ we find a unit eigenvector of AA^T:

$$\bar{u}_2 = \begin{bmatrix} 0.3637 \\ -0.1571 \\ 0.5207 \\ -0.7563 \end{bmatrix}$$

and we find a unit eigenvector of $A^T A$:

$$\bar{v}_2 = \begin{bmatrix} 0.9915 \\ -0.1299 \end{bmatrix}$$

Observe that $A^T \bar{u}_2 = \sigma_2 \bar{v}_2$ and $A\bar{v}_2 = \sigma_2 \bar{u}_2$.

At this point we only have two out of four necessary columns for U. The final two columns can be filled in using eigenvectors for the eigenvalue 0, by the Spectral Theorem there are two and they form an orthonormal set along with $\bar{u}_1$ and $\bar{u}_2$.

We have:
$$\bar{u}_3 = \begin{bmatrix} -0.5284 \\ 0.1276 \\ 0.7957 \\ 0.2672 \end{bmatrix} \quad \text{and} \quad \bar{u}_4 = \begin{bmatrix} 0.2465 \\ -0.8624 \\ 0.1641 \\ 0.4106 \end{bmatrix}$$

Note that we could have reversed the assignment of these, it does not matter.

Thus we have:

$$\begin{bmatrix} 1 & 3 \\ 0 & 2 \\ 1 & 1 \\ -1 & 2 \end{bmatrix} = \underbrace{\begin{bmatrix} -0.7265 & 0.3637 & -0.5284 & 0.2465 \\ -0.4640 & -0.1571 & 0.1276 & -0.8624 \\ -0.2624 & 0.5207 & 0.7957 & 0.1641 \\ -0.4336 & -0.7563 & 0.2672 & 0.4106 \end{bmatrix}}_{U}$$

$$\underbrace{\begin{bmatrix} 4.2734 & 0 \\ 0 & 1.6547 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}}_{\Sigma} \underbrace{\begin{bmatrix} -0.1299 & 0.9915 \\ -0.9915 & -0.1299 \end{bmatrix}^T}_{V^T}$$

8.4 Matlab

Calculations of the Singular Value Decomposition can be easily done in Matlab:

```
>> A=[
1 2 0
3 1 4
1 1 -1
0 1 2
]
A =
     1     2     0
     3     1     4
     1     1    -1
     0     1     2
>> [U,S,V] = svd(A)
U =
   -0.2079    0.7246   -0.3584   -0.5507
   -0.9171   -0.1139    0.3788   -0.0501
   -0.0103    0.6615    0.2666    0.7009
   -0.3400   -0.1560   -0.8106    0.4506
S =
    5.5200         0         0
         0    2.5538         0
         0         0    1.4170
         0         0         0
V =
   -0.5380    0.4090    0.7371
   -0.3049    0.7208   -0.6225
   -0.7859   -0.5596   -0.2631
```

8.5 Exercises

Exercise 8.1. Find the singular value decomposition of each of the following matrices. First do this by computing both AA^T and $A^T A$, finding the eigenvalue/eigenvector pairs of each, finding the corresponding singular values, and putting the results together. Then check your answer via technology.

(a) $A = \begin{bmatrix} 1.0 & 2.0 & -3.0 \\ 0 & 1.0 & 1.0 \\ 1.0 & 2.0 & 5.0 \\ -1.0 & 0 & 2.0 \end{bmatrix}$

(b) $A = \begin{bmatrix} -1.0 & 0 & 2.0 & 2.0 & 2.0 \\ 0 & 2.0 & 3.0 & 0 & 1.0 \\ 1.0 & 2.0 & -2.0 & 1.0 & 2.0 \end{bmatrix}$

(c) $A = \begin{bmatrix} 1.0 & 2.0 & -3.0 \\ 0 & 1.0 & 1.0 \\ 1.0 & 2.0 & 5.0 \\ -1.0 & 0 & 2.0 \end{bmatrix}$

(d) $A = \begin{bmatrix} 0.1 & 0.2 & 0.9 & 0.3 \\ 0.9 & 0.2 & 0 & 0.2 \\ 0.2 & 0.2 & 0.3 & 0.1 \\ 0 & 0.3 & 0.7 & 0.6 \end{bmatrix}$