Andrey Nikishaev
Jul 2, 2019 · 4 min read · Listen

Linear Algebra. SVD and PCA

Want become an expert in Computer Vision & Object Detection?

Subscribe on New Practical Course

· · ·

SVD

$$\mathbf{X}_{mxn} = \mathbf{U}_{mxm}\mathbf{S}_{mxn}\mathbf{V}^{\top}_{nxn}$$

Geometric meaning

The main idea of SVD is that matrix A is basically a linear transformation. And SVD decompose it into 3 factors Rotation(U)*Scaling/Shear(Σ)*Rotation(V). This gives you interesting posibilities like use it for dimension reduction(that i will show in PCA example), or for point mathching in space

Nice interactive demo: https://www.geogebra.org/m/mrey8VJX

How to compute SVD

1. We have matrix A for which we want to compute SVD

$$A = \begin{bmatrix} 4 & 0 \\ 3 & -5 \end{bmatrix}$$

2. We need to compute A.T and gram(A) = A.T * A

Since $A^T = \begin{bmatrix} 4 & 3 \\ 0 & -5 \end{bmatrix}$ then, $A^TA = \begin{bmatrix} 4 & 3 \\ 0 & -5 \end{bmatrix}\begin{bmatrix} 4 & 0 \\ 3 & -5 \end{bmatrix}$

$$A^TA = \begin{bmatrix} 25 & -15 \\ -15 & 25 \end{bmatrix}$$

3. From gram(A) we can compute eigenvalues and singular values which will be real, cause gram(A) is NxN size.

$$A^TA - cI = \begin{bmatrix} 25 - c & -15 \\ -15 & 25 - c \end{bmatrix} \qquad |A^TA - cI| = (25 - c)(25 - c) - (-15)(-15) = 0$$

characteristic equation $\longrightarrow$ $c^2 - 50c + 400 = 0$

The quadratic equation gives two values. In decreasing order, these are $\longrightarrow$ $|40| > |10|$

eigenvalues $\longrightarrow$ $c_1 = 40 \quad c_2 = 10$

singular values $\longrightarrow$ $s_1 = \sqrt{40} = 6.3245 > s_2 = \sqrt{10} = 3.1622$

Note that singular values should be sorted from biggest to smallest
s1 ≥ s2 ≥ ⋯ ≥ sn ≥ 0

4. Construct **S** matrix

$$S = \begin{bmatrix} 6.3245 & 0 \\ 0 & 3.1622 \end{bmatrix}$$

5. Now we need find matrix **V**. For this we need to solve equation **gram(A) — cn*I** for each eigenvalue, where I — is identity matrix

for $c_1 = 40$

$$A^TA - cI = \begin{bmatrix} 25 - 40 & -15 \\ -15 & 25 - 40 \end{bmatrix} = \begin{bmatrix} -15 & -15 \\ -15 & -15 \end{bmatrix}$$

$$(A^TA - cI)\,X_1 = 0$$

$$\begin{bmatrix} -15 & -15 \\ -15 & -15 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$-15x_1 + -15x_2 = 0$$
$$-15x_1 + -15x_2 = 0$$

Solving for x_2 for either equation: $x_2 = -x_1$

$$X_1 = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ -x_1 \end{bmatrix}$$

Dividing by its length,

$$L = \sqrt{x_1^2 + x_2^2} = x_1\sqrt{2}$$

$$X_1 = \begin{bmatrix} x_1 / L \\ -x_1 / L \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{-1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} 0.7071 \\ -0.7071 \end{bmatrix}$$

for $c_2 = 10$

$$A^TA - cI = \begin{bmatrix} 25 - 10 & -15 \\ -15 & 25 - 10 \end{bmatrix} = \begin{bmatrix} 15 & -15 \\ -15 & 15 \end{bmatrix}$$

$$(A^TA - cI)\,X_2 = 0$$

$$\begin{bmatrix} 15 & -15 \\ -15 & 15 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$15x_1 + -15x_2 = 0$$
$$-15x_1 + 15x_2 = 0$$

Solving for x_2 for either equation: $x_2 = x_1$

$$X_2 = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_1 \end{bmatrix}$$

Dividing by its length,

$$L = \sqrt{x_1^2 + x_2^2} = x_1\sqrt{2}$$

$$X_2 = \begin{bmatrix} x_1 / L \\ x_1 / L \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} 0.7071 \\ 0.7071 \end{bmatrix}$$

$$V = \begin{bmatrix} X_1 & X_2 \end{bmatrix} = \begin{bmatrix} 0.7071 & 0.7071 \\ -0.7071 & 0.7071 \end{bmatrix}$$

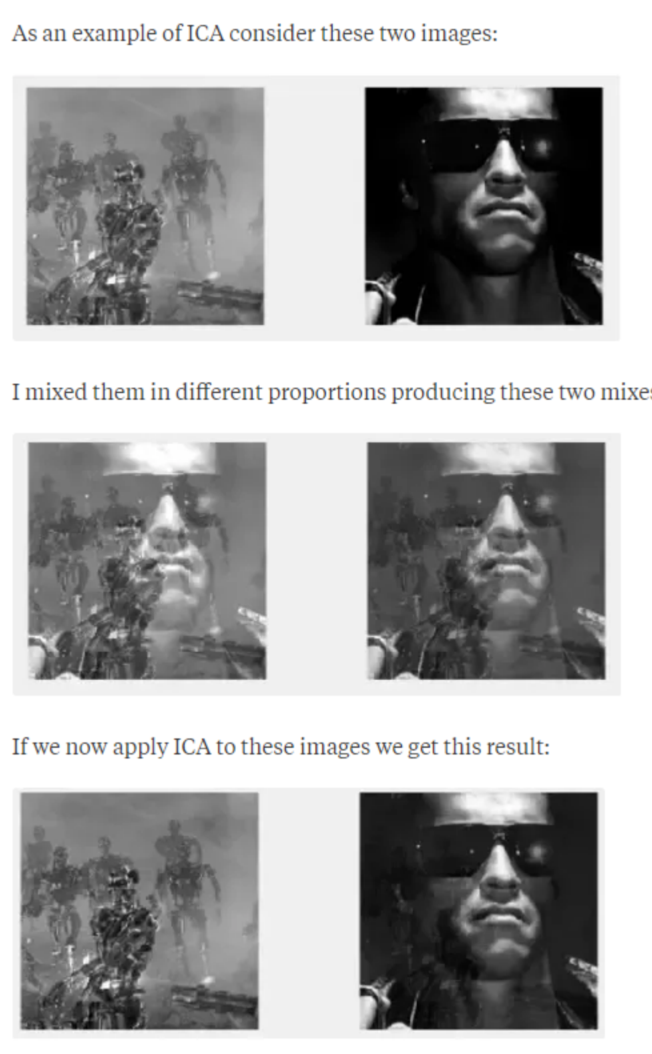
$$V^T = \begin{bmatrix} 0.7071 & -0.7071 \\ 0.7071 & 0.7071 \end{bmatrix}$$

6. Now we need to find matrix **U**. If determinant(diag(S)) != 0 we can get U from equation **U = A*V*inv(S)**. In other situation when determinant(diag(S)) != 0 and matrix is Singular (not invertible) we can use formula **un = (1/sn) * A * vn**, where vn is V[;,n] vector. U = [u1,..,un]

$$\begin{bmatrix} 0.4472 & 0.8944 \\ 0.8944 & -0.4472 \end{bmatrix}$$

PCA

First of all i want to say that PCA has a brother ICA, and the difference between them can be shown on this image example:



https://www.quora.com/What-is-the-difference-between-PCA-and-ICA/answer/Luis-Argerich

So idea of PCA is a reduction of space, while ICA separating features of that space.
But right now we will talk about PCA, which is simple but still quite usefull things(of course it not the only algo in area of dimension reduction)

How to compute PCA

So we have matrix **X**. First we will need center it by **X/std — mean** (by columns)

Then the covariance of X equal

$$\mathbf{C} = \mathbf{X}^{\top}\mathbf{X}/(n-1)$$

C matrix is symmetric so can be diagonalized

$$\mathbf{C} = \mathbf{V}\mathbf{L}\mathbf{V}^{\top},$$

where V — matrix of eigenvectors, L — diagonal matrix of eigenvalue.
The eigenvectors are called **principal axes** and projection of the data on them **principal components (XV)**

Now lets perform **SVD** over **X**

$$\mathbf{X} = \mathbf{U}\mathbf{S}\mathbf{V}^{\top},$$

Pasting this into formula of covariance of X will gives us this

$$\mathbf{C} = \mathbf{V}\mathbf{S}\mathbf{U}^{\top}\mathbf{U}\mathbf{S}\mathbf{V}^{\top}/(n-1) = \mathbf{V}\frac{\mathbf{S}^2}{n-1}\mathbf{V}^{\top}$$

where singular values are related to the eigenvalues of covariance matrix as

$$\lambda_i = s_i^2/(n-1)$$

Principal components are given by

$$\mathbf{T} = \mathbf{X}\mathbf{V} = \mathbf{U}\mathbf{S}\mathbf{V}^{\top}\mathbf{V} = \mathbf{U}\mathbf{S}$$

So:

1. Use SVD on A* and get U, V and S.
2. Set new N dimension to which you want to reduce your data. And get first N values from U and S.
Ur= U[:,:n], Sr = S[:n,:n]

3) Reduced(A, N) = Ur * Sr

Of course it's ony one way of computing PCA, but as i know almost all curent algorithms use SVD(different variations of it) for that, so it most popular.

Code

You can run all this code in Google Colab

Sources

- https://math.berkeley.edu/~hutching/teach/54-2017/svd-notes.pdf
- http://math.mit.edu/classes/18.095/2016IAP/lec2/SVD_Notes.pdf
- https://en.wikipedia.org/wiki/Singular_value_decomposition

Support

Become a Patron and support our community to make more interesting articles & tutorials

More cool things about Linear Algebra

Linear Algebra. Points matching with SVD in 3D space
Problem
towardsdatascience.com

Linear Algebra. Polynoms. Interpolation. Least squares
In this part of lectures we will talk about simplest "ml models" — polynoms.
towardsdatascience.com