Singular Value Decomposition

Yash Rohan

March 23, 2023

Vector Representation

Conventionally, we represent the space using i, j, k for X, Y and Z axis respectively. Any point in the space can be represented by using a vector representation:

$$\vec{v} =< v_1, v_2, v_3 >= \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$$

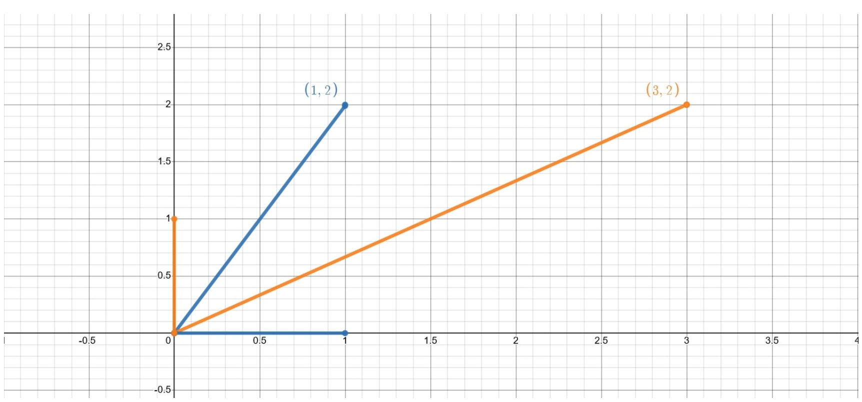
The i, j, k form the orthogonal basis of the 3-D space, meaning, any 3 dimensional vector can be represented using these vectors, which are also perpendicular, as well as linearly independent, to each other.

Linear Transformation

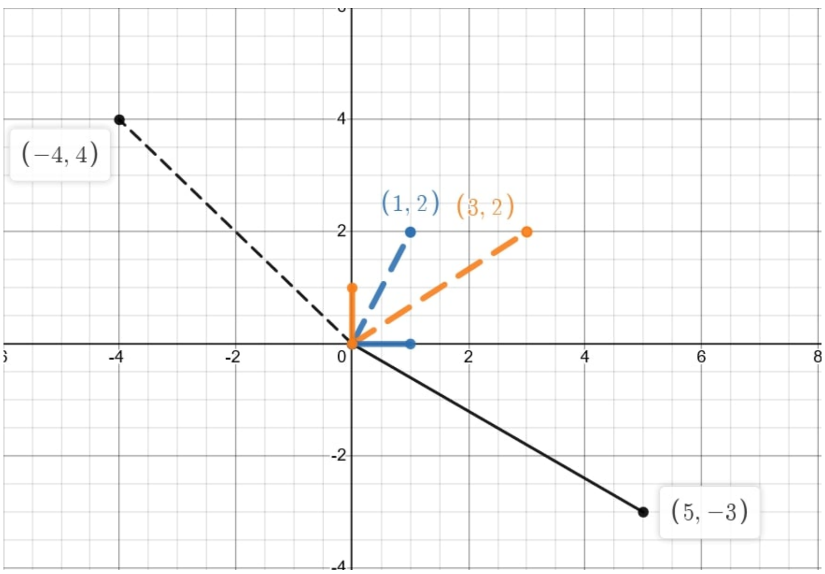
A matrix:

$$\begin{bmatrix} 1 & 3 \\ 2 & 2 \end{bmatrix}$$

means that the basis vectors, $\vec{v} =< 1, 0 > and < 0, 1 >$, changes to, $\vec{v'} =< 1, 2 > and < 3, 2 >$.

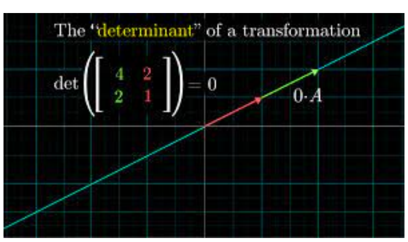


It means that, if in the original grid, there was a vector $\vec{a} =< 5, -3 >= \begin{bmatrix} 5 \\ -3 \end{bmatrix}$ it will be transformed to $(5)\begin{bmatrix} 1 \\ 2 \end{bmatrix} + (-3)\begin{bmatrix} 3 \\ 2 \end{bmatrix}$, that is $\begin{bmatrix} -4 \\ 4 \end{bmatrix}$. This is the point (5,-3) with respect to the new basis vectors.

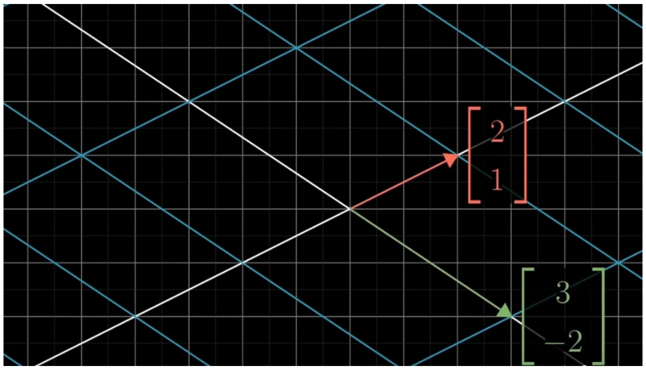


Determinant

Determinant is the factor by which a Linear Transformation scales any grid square.(changes the area in 2D and changes the volumes in 3D)
Negative determinant implies that the orientation of space has reversed, and the absolute value of it gives us the scaling factor.
Determinant 0 implies that the space is "squished" into a line, or in some

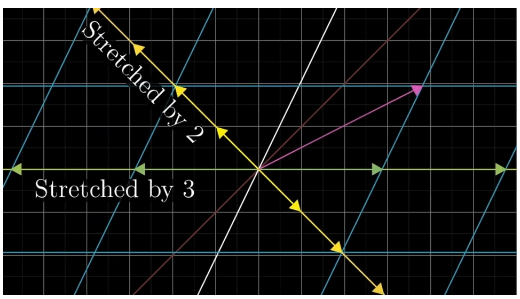


extreme cases, into a point.
From the diagram below, we can understand that the matrix $\begin{bmatrix} 3 & 2 \\ 1 & -2 \end{bmatrix}$ scales the space by a value its determinant, which is 7.

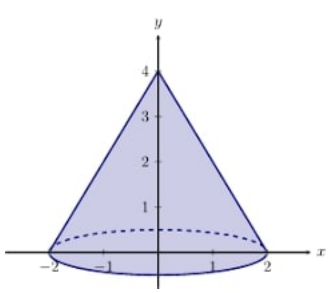


Matrix as Dimension Changer

Characteristic Equation



In order to find the Eigen vectors, we use the characteristic equation $(A - (\lambda)I)u = 0$, or, $Au = \lambda u$. The λ(eigen value) here represents the scale by which the Eigen vector is stretched. In 3D the line of Eigen vector is called the axis of the body.



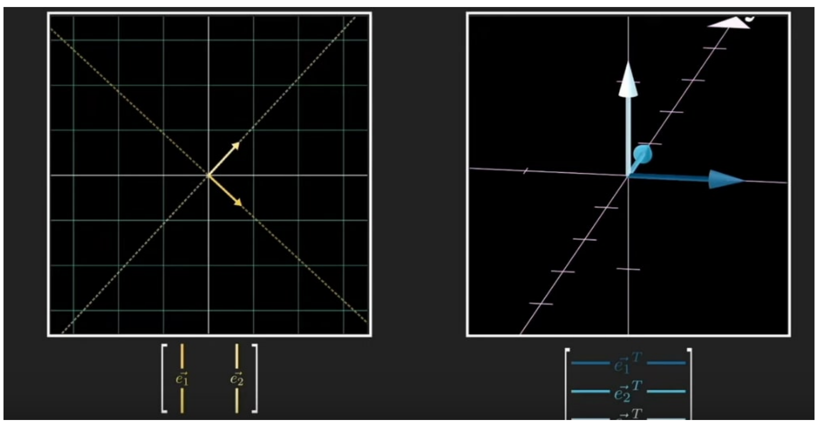
Eigen Value Decomposition

Let there be a Matrix A, and $u_1, u_2, \ldots, u_n$ be the eigen vectors of it. Let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be corresponding eigen values.
AU=Uλ (by definition of eigen vectors). If U^{-1} exists,then we can write:
A=Uλ U^{-1} [eigen value decomposition]

Symmetric Matrix

Symmetric matrices have perpendicular Eigen Vectors. if we package these normalized eigen vectors in matrix, we get an orthogonal matrix, which in terms of linear transformation, implies rotation. And the transpose of this matrix rotates the eigen vectors to standard bases.

Most of the vectors in the nature are not symmetric, however, we can con-



vert any matrix to a symmetric, simply by multiplying it with it's own transpose.

$$AA^T = symmetric.A^T A = symmetric$$

If A_{mxn} is a rectangular matrix, then:

$$AA^T = mm$$

The Eigen Vectors of S_L are called the left Singular Vector and is denoted by

$$S_L = AA^T;$$

$$A^T A = nXn$$

The Eigen Vectors of S_R are called the Right Singular Vectors.

$$S_R = A^T A$$

S_R and S_L are called PSD (positive semi-definite) matrices, meaning that they have non negative eigen values.

$$\lambda_i > 0$$

S_R and S_L have the same eigen values.

$$S_L - \lambda_1, \lambda_2 \ldots, \lambda_m$$

$$S_R - \lambda_1, \lambda_2, \ldots, \lambda_m, \ldots, \lambda_n$$

and therefore,

$$\lambda_1 = \lambda_1, \lambda_2 = \lambda_2 \ldots, \lambda_m = \lambda_m;$$

$$\lambda_{m+1} = \ldots = \lambda_n = 0$$

$$\lambda_1 > \lambda_2 > \ldots > \lambda_m$$

The singular values is defined as the nonnegative square roots of the eigen values of S_L or S_R, that is:

$$\sqrt{\lambda_1} = \sigma_1, \sqrt{\lambda_2} = \sigma_2, \ldots, \sqrt{\lambda_m} = \sigma_m$$

And $\begin{bmatrix} \sigma_1 & \sigma_2 & \ldots & \sigma_n \\ \vdots & & & \\ \sigma_{m-n} & \ldots & \ldots & \sigma m \end{bmatrix}$ is the singular value matrix of the matrix A.

Singular Value Decomposition

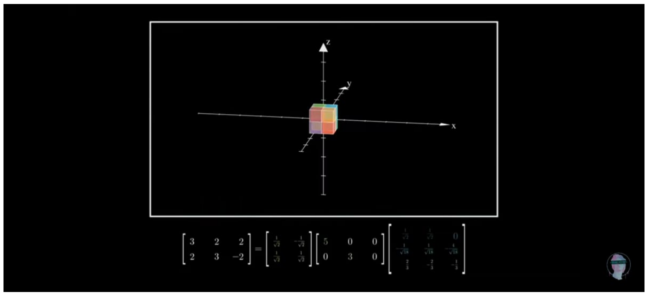
Any matrix A can be expressed, or decomposed, as the product of three very special types of matrices that are generated through the original matrix itself. It has the form:

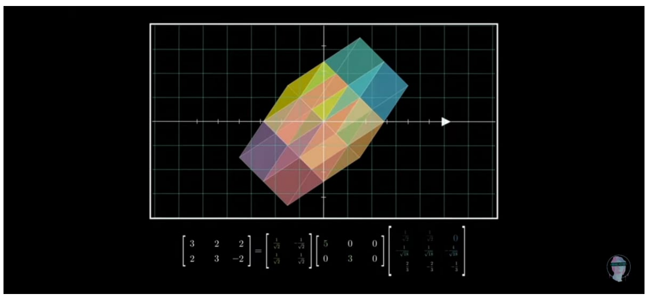
$$A = U\Sigma V^T$$

Where:
Σ is a Rectangular Diagonal Matrix, made up of Singular values of matrix A, arranged is descending order.
U is an orthogonal matrix, made up of normalized eigen values of S_L, arranged in descending order.
V consists of normalized Right Singular vectors of matrix A, also arranged is descending order.

Visualization of SVD

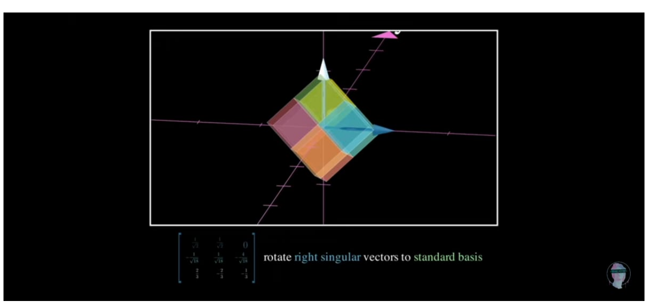
A matrix A $\begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{bmatrix}$ depicts a linear transformation.
For this transformation, we have





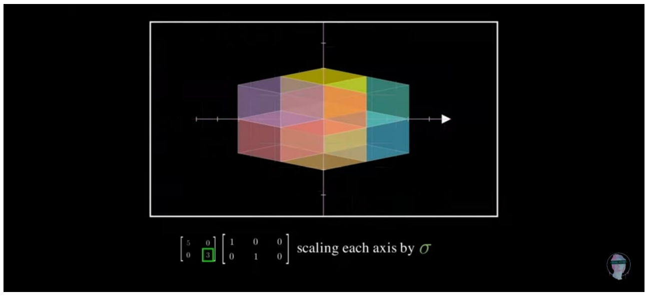
$$U=\begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix};\ \Sigma=\begin{bmatrix} 5 & 0 & 0 \\ 0 & 3 & 0 \end{bmatrix};\ V^T=\begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{-1}{\sqrt{18}} & \frac{1}{\sqrt{18}} & \frac{-4}{\sqrt{18}} \\ \frac{2}{3} & \frac{-2}{3} & \frac{-1}{3} \end{bmatrix}$$

It means that, sequentially, first V^T, the Σ then U, will produce the same Linear Transformation as A would.
V^T would rotate the right singular vector to Standard Basis.(Biggest Singular Value lies on X axis, then the Y axis and so on.)

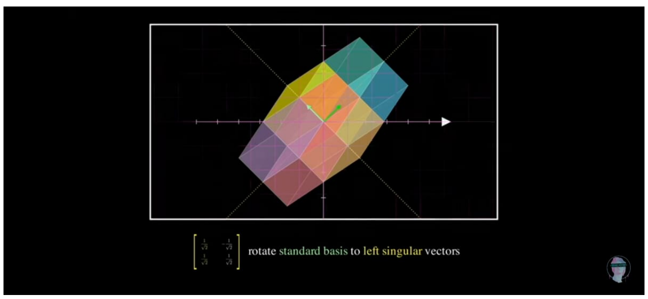


Σ is a square diagonal matrix composed with a dimension eraser (removes 3^{rd}

dimension). And it stretches the X and Y axis by the corresponding σ values.



U rotates the Standard Basis to the Left Singular Vector of A.

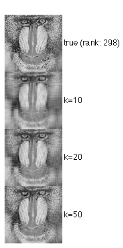


Therefore, in this visualization, we clearly see that, this sequence of linear transformations is same as the cumulative transformation brought about by A.

Applications of SVD

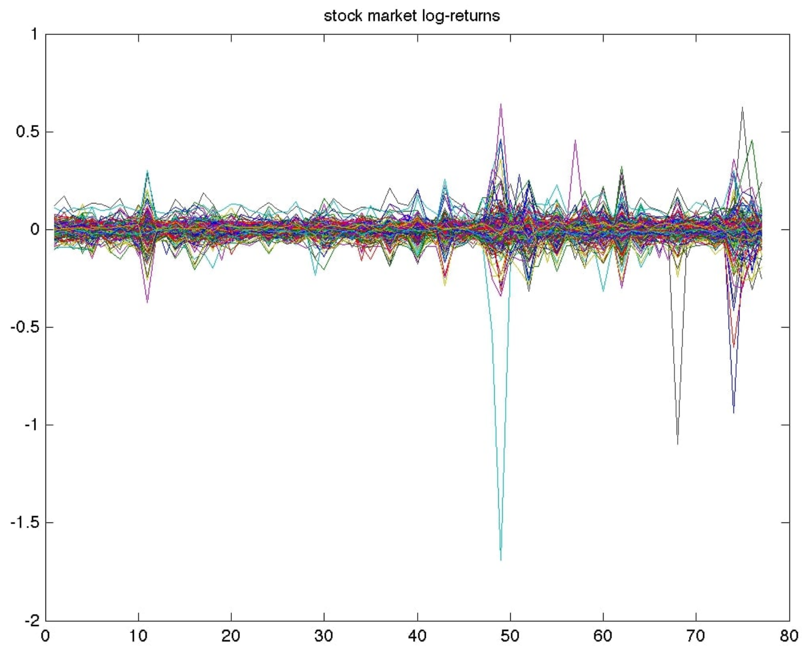
1 Image Compression

Images are a set of pixels. These pixels are arranged in a matrix of order nXm.By using Low Rank Approximation,we can reduce the rank of the pixel matrix, and hence, compress it(by sacrificing information).

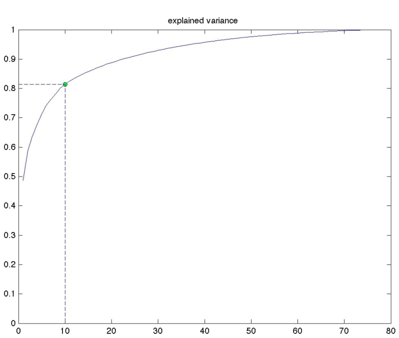


(k is the rank of the matrix)

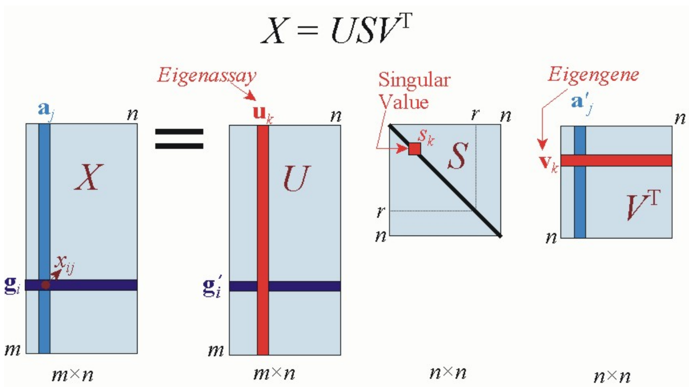
2 Market Data Analysis



The above figure represents the time series of a stock market, shown as a collection of time-series. We note that the log-returns hover around a mean which appears to be close to zero.
We can form the SVD of the matrix of log-returns, and plot the explained variance. We see that the first 10 singular values explain more than 0.8 of the data's variance.



3 Gene Expression Analysis



There are several more uses of SVD especially in PCA(principle component analysis) and Machine Learning which are beyond the scope of this paper.