SINGULAR VALUE DECOMPOSITION

Singular Value Decomposition (SVD) is the factorization of any $m \times n$ matrix A as $A = U\Sigma V^T$ where U is a $m \times m$ matrix made up of columns of mutually perpendicular unit vectors, Σ is a diagonal matrix of the same shape as A and V^T is a $n \times n$ matrix made up of rows of mutually perpendicular unit vectors.

1   Example

Let A be a 2×3 matrix as follows,

$$A = \begin{pmatrix} 3 & 0 & 0 \\ -8 & 0 & 3 \end{pmatrix}$$

To calculate unit vectors of U, eigenvalues and eigenvectors of AA^T are calculated,

$$AA^T = \begin{pmatrix} 3 & 0 & 0 \\ -8 & 0 & 3 \end{pmatrix} \begin{pmatrix} 3 & -8 \\ 0 & 0 \\ 0 & 3 \end{pmatrix}$$

$$= \begin{pmatrix} 9 & -24 \\ -24 & 73 \end{pmatrix}$$

$AA^T - \lambda I$ must be singular if eigenvectors are non-zero. Hence,

$$\begin{vmatrix} 9 - \lambda & -24 \\ -24 & 73 - \lambda \end{vmatrix} = 0$$

$$(9 - \lambda)(73 - \lambda) = 576$$

This gives $\lambda = 81, 1$ and $\hat{u} = \begin{pmatrix} \frac{1}{\sqrt{10}} \\ \frac{-3}{\sqrt{10}} \end{pmatrix}, \begin{pmatrix} \frac{3}{\sqrt{10}} \\ \frac{1}{\sqrt{10}} \end{pmatrix}$ respectively.

At this step, decomposition is as follows:

$$A = (\hat{u}_1, \hat{u}_2) \begin{pmatrix} \sqrt{\lambda_1} & 0 & 0 \\ 0 & \sqrt{\lambda_2} & 0 \end{pmatrix} V^T$$

$$A = \begin{pmatrix} \frac{1}{\sqrt{10}} & \frac{3}{\sqrt{10}} \\ \frac{-3}{\sqrt{10}} & \frac{1}{\sqrt{10}} \end{pmatrix} \begin{pmatrix} 9 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} V^T$$

To continue and calculate V, the relation $A\hat{v}_i = \sqrt{\lambda_i}\hat{u}_i$ is used. It is clear that V is a 3×3 matrix.

$$A\hat{v}_1 = \sqrt{\lambda_1}\hat{u}_1$$

$$\begin{pmatrix} 3 & 0 & 0 \\ -8 & 0 & 3 \end{pmatrix} \hat{v}_1 = \sqrt{81} \begin{pmatrix} \frac{1}{\sqrt{10}} \\ \frac{-3}{\sqrt{10}} \end{pmatrix}$$

$$\hat{v}_1 = \begin{pmatrix} \frac{3}{\sqrt{10}} \\ 0 \\ \frac{-1}{\sqrt{10}} \end{pmatrix}$$

Similarly,

$$A\hat{v}_2 = \sqrt{\lambda_2}\hat{u}_2$$

$$\begin{pmatrix} 3 & 0 & 0 \\ -8 & 0 & 3 \end{pmatrix} \hat{v}_2 = \sqrt{1} \begin{pmatrix} \frac{3}{\sqrt{10}} \\ \frac{1}{\sqrt{10}} \end{pmatrix}$$

$$\hat{v}_2 = \begin{pmatrix} \frac{1}{\sqrt{10}} \\ 0 \\ \frac{3}{\sqrt{10}} \end{pmatrix}$$

Notice that $\hat{v}_1.\hat{v}_2$ is zero. The choices for $\hat{v}_3$ which must be perpendicular to both $\hat{v}_1$ and $\hat{v}_2$ are $(0, \pm 1, 0)^T$ and both are admissible in the decomposition as follows,

$$A = \begin{pmatrix} \frac{1}{\sqrt{10}} & \frac{3}{\sqrt{10}} \\ \frac{-3}{\sqrt{10}} & \frac{1}{\sqrt{10}} \end{pmatrix} \begin{pmatrix} 9 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} \hat{v}_1 & \hat{v}_2 & \hat{v}_3 \end{pmatrix}^T$$

$$A = \begin{pmatrix} \frac{1}{\sqrt{10}} & \frac{3}{\sqrt{10}} \\ \frac{-3}{\sqrt{10}} & \frac{1}{\sqrt{10}} \end{pmatrix} \begin{pmatrix} 9 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} \frac{3}{\sqrt{10}} & \frac{1}{\sqrt{10}} & 0 \\ 0 & 0 & \pm 1 \\ \frac{-1}{\sqrt{10}} & \frac{3}{\sqrt{10}} & 0 \end{pmatrix}^T$$

$$\begin{pmatrix} 3 & 0 & 0 \\ -8 & 0 & 3 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{10}} & \frac{3}{\sqrt{10}} \\ \frac{-3}{\sqrt{10}} & \frac{1}{\sqrt{10}} \end{pmatrix} \begin{pmatrix} 9 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} \frac{3}{\sqrt{10}} & 0 & \frac{-1}{\sqrt{10}} \\ \frac{1}{\sqrt{10}} & 0 & \frac{3}{\sqrt{10}} \\ 0 & \pm 1 & 0 \end{pmatrix}$$

2 Analysis

SVD can be visualized as follows:

$$A_{m \times n} = \begin{pmatrix} | & | & .. & | \\ \hat{u_1} & \hat{u_2} & .. & \hat{u_m} \\ | & | & .. & | \end{pmatrix}_{m \times m} \begin{pmatrix} \sqrt{\lambda_1} & 0 & .. & 0 & 0 & .. & 0 \\ 0 & \sqrt{\lambda_2} & .. & 0 & 0 & .. & 0 \\ 0 & 0 & .. & 0 & 0 & .. & 0 \\ 0 & 0 & .. & \sqrt{\lambda_r} & 0 & .. & 0 \\ 0 & 0 & .. & 0 & 0 & .. & 0 \\ . & . & .. & . & . & .. & . \\ . & . & .. & . & . & .. & . \\ 0 & 0 & .. & 0 & 0 & .. & 0 \end{pmatrix}_{m \times n} \begin{pmatrix} | & | & .. & | \\ \hat{v_1} & \hat{v_2} & .. & \hat{v_n} \\ | & | & .. & | \end{pmatrix}_{n \times n}^{T}$$

$u_i \perp u_j \; \forall \; i \neq j$

$v_i \perp v_j \; \forall \; i \neq j$

$\sqrt{\lambda_1} \geq \sqrt{\lambda_2} \geq .. \geq \sqrt{\lambda_r} > 0$

$r = rank(A)$

The product of the last two matrices can be simplified as follows,

$$A_{m \times n} = \begin{pmatrix} | & | & .. & | \\ \hat{u_1} & \hat{u_2} & .. & \hat{u_m} \\ | & | & .. & | \end{pmatrix}_{m \times m} \begin{pmatrix} \sqrt{\lambda_1}\hat{v_1}^T \\ \sqrt{\lambda_2}\hat{v_2}^T \\ . \\ . \\ \sqrt{\lambda_r}\hat{v_r}^T \\ 0 \\ . \\ . \\ 0 \end{pmatrix}_{m \times n}$$

leading to the representation of A as a sum of r rank-1 matrices as follows:

$$A = \sum_{i=1}^{r} \sqrt{\lambda_i} \; \hat{u_i} \; \hat{v_i}^T$$

Multiplying by $\hat{v_j}$ on both sides,

$$A \hat{v_j} = \sum_{i=1}^{r} \sqrt{\lambda_i} \; \hat{u_i} \; \hat{v_i}^T \; \hat{v_j}$$

Since all the $\hat{v_i}$s are mutually perpendicular to each other,

$$A\ \hat{v}_j = \sqrt{\lambda_j}\hat{u}_j$$

Note the similarity of this equation to the regular eigenvalue equation. The constraint of the same unit vector on both sides of the equation was let go in the return of universal applicability on any matrix whatsoever.

3   Application: Matrix Compression

Subjecting the matrix in example above to the rank-1 breakdown equation,

$$A = \sum_{i=1}^{r} \sqrt{\lambda_i}\ \hat{u}_i\ \hat{v}_i^{T}$$

$$\begin{pmatrix} 3 & 0 & 0 \\ -8 & 0 & 3 \end{pmatrix} = \begin{pmatrix} 2.7 & 0 & -0.9 \\ -8.1 & 0 & 2.7 \end{pmatrix} + \begin{pmatrix} 0.3 & 0 & 0.9 \\ 0.1 & 0 & 0.3 \end{pmatrix}$$

Note that the first matrix on the right hand side is a resonable approximation of the matrix on the left hand side and the other matrix just represents minor tweaks.

$$\begin{pmatrix} 3 & 0 & 0 \\ -8 & 0 & 3 \end{pmatrix} \approx \begin{pmatrix} 2.7 & 0 & -0.9 \\ -8.1 & 0 & 2.7 \end{pmatrix}$$

$$\approx \sqrt{81} \begin{pmatrix} \frac{1}{\sqrt{10}} \\ \frac{-3}{\sqrt{10}} \end{pmatrix} \begin{pmatrix} \frac{3}{\sqrt{10}} & 0 & \frac{-1}{\sqrt{10}} \end{pmatrix}$$

The left hand side representation requires to store mn entries while the right hand side only requires to store $k(m+n)$ entries. So if $k << m, n$, the right hand side is highly compressed. The right hand side is also a good approximation if k is of the order of the rank of the matrix. So for low-rank highly redundant matrices (example digital images), SVD can serve as a reasonable compression algorithm.

4   Appendix

$$A = U\Sigma V^{T}$$
$$AA^{T} = U\Sigma V^{T}V\Sigma^{T}U^{T}$$

Since $V^{T}V = I$ and Σ is diagonal.

$$AA^T = U\Sigma^2 U^T$$

$$AA^T = \sum_{i=1}^{r} \lambda_i \ \hat{u}_i \ \hat{u}_i{}^T$$

Multiplying both sides by $\hat{u}_j$ and using the fact that all u_is are mutually perpendicular to each other.

$$AA^T \ \hat{u}_j = \sum_{i=1}^{r} \lambda_i \ \hat{u}_i \ \hat{u}_i{}^T \ \hat{u}_j$$

$$AA^T \hat{u}_j = \lambda_j \ \hat{u}_j$$

It is clear that $\hat{u}$s are eigenvectors of AA^T. Similarly, it can be seen that $\hat{v}$s are eigenvectors of $A^T A$.