Singular Value Decomposition

Nirav bhatt

August 2023

In this section, singular value decomposition (known as SVD) will be studied. Before we start the SVD, let us discuss the backbone of the SVD algorithms, the eigenvalues and eigen vectors of the matrix. Given an $(n \times n)-$dimensional matrix, $\mathbf{A}$, the problem of eigenvalue–eigenvector is to find the solutions of the equation

$$\mathbf{A}\mathbf{x} = \lambda \mathbf{x} \tag{1}$$

where λ is a scalar, known as an eigenvalue of the matrix $\mathbf{A}$, and $\mathbf{x}$ is an $n-$dimensional vector of the matrix $\mathbf{A}$. Here, there are two unknowns, λ and $\mathbf{x}$. In a solution of a set of linear systems $\mathbf{A}\mathbf{x} = \mathbf{b}$, $\mathbf{x}$ is unknown. Note that Eq. (1) is a nonlinear equation due to "$\lambda\,\mathbf{x}$" term. However, if λ is known, then Eq. (1) becomes a linear equation. Then, the solution of eigenvalue–eigenvectors can be decomposed into two steps:

1. Find λ and
2. Find $\mathbf{x}$.

Next, we will find λ values. Let us re–write Eq. (1) in the following form:

$$\begin{aligned} \mathbf{A}\mathbf{x} - \lambda\mathbf{x} &= \mathbf{0} \\ (\mathbf{A} - \lambda\mathbf{I})\mathbf{x} &= \mathbf{0} \end{aligned} \tag{2}$$

In Eq. (2), the vector $\mathbf{x}$ is in the null space of the matrix $\mathbf{A} - \lambda\mathbf{I}$. Alternatively, we can say that we would like to choose value of λ such that $\mathbf{A} - \lambda\mathbf{I}$ has null space. Note that $\mathbf{x} = \mathbf{0}$ is a trivial solution of the problem and we are not interested in the trivial solution. Hence, the number λ is an eigenvalue of $\mathbf{A}$ if and only if $\mathbf{A} - \lambda\mathbf{I}$ is singular, i.e.,

$$\det(\mathbf{A} - \lambda\mathbf{I}) = 0 \tag{3}$$

Eq. (3) is the characteristic equation.

Example 1. *Consider a matrix* $\mathbf{A}$ *as follows*

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

Then,

$$\det\left(\mathbf{A}-\lambda\mathbf{I}\right)=0$$
$$\det\begin{bmatrix}2-\lambda & 1\\ 1 & 2-\lambda\end{bmatrix}=0$$
$$\lambda^2-4\,\lambda+3=0$$

The solution is $\lambda=3,\,1$. *Let us compute the eigen vectors corresponding to these eigenvalues*

1. $\lambda=1$. *We have to find the null space of the following matrix:*

$$\begin{bmatrix}1 & 1\\ 1 & 1\end{bmatrix}\begin{bmatrix}x_{11}\\ x_{21}\end{bmatrix}=\boldsymbol{0}\rightarrow x_{11}=-x_{21} \tag{4}$$

Then, the eigen vector corresponding to $\lambda=1$ *is* $\mathbf{x}_1=\begin{bmatrix}1\\ -1\end{bmatrix}$

2. $\lambda=3$. *Similarly, the eigen vector for the second eigenvalue is* $\mathbf{x}_2=\begin{bmatrix}1\\ 1\end{bmatrix}$

If $\mathbf{A}$ is a real symmetric matrix of $n\times n-$dimensional matrix with the eigenvalues $\lambda_1,\,\lambda_2,\,\ldots,\lambda_n$ and the corresponding orthonormal independent eigenvectors $\mathbf{x}_1,\,\mathbf{x}_2,\,\ldots,\mathbf{x}_n$. Then,

$$\mathbf{A}\mathbf{x}_i=\lambda_i\mathbf{x}_i,\quad i=1,\,2,\,\ldots,n \tag{5}$$

By arranging all the eigenvectors in the following form

$$\begin{aligned}\mathbf{A}\begin{bmatrix}\mathbf{x}_1 & \mathbf{x}_2 & \ldots & \mathbf{x}_n\end{bmatrix} &= \begin{bmatrix}\mathbf{x}_1 & \mathbf{x}_2 & \ldots & \mathbf{x}_n\end{bmatrix}\,diag([\lambda_1,\,\lambda_2,\,\ldots,\lambda_n])\\ \mathbf{A}\mathbf{X} &= \mathbf{X}\boldsymbol{\Lambda}\\ \mathbf{A} &= \mathbf{X}\boldsymbol{\Lambda}\mathbf{X}^{-1}\\ \mathbf{A} &= \mathbf{X}\boldsymbol{\Lambda}\mathbf{X}^{\mathrm{T}}\quad(\because\mathbf{A}\text{ is symmetric})\\ \mathbf{A} &= \sum_{i=1}^{n}\lambda_i\mathbf{x}_i\mathbf{x}_i^{\mathrm{T}}\end{aligned} \tag{6}$$

Note that the eigenvalues need not be distinct. It can be zero. Eq. (6) is called *Eigen or Spectral decomposition.* Next, we will define two types of matrices: (i) Positive semi-definite matrices, and (ii) Negative semi-definite matrices.

Definition 2. *A matrix* $\mathbf{A}$ *is said to be positive semi-definite if* $\mathbf{z}^{\mathrm{T}}\mathbf{A}\mathbf{z}\geq 0$ *for all non-zero vectors* $\mathbf{z}$. *Similarly, if* $\mathbf{z}^{\mathrm{T}}\mathbf{A}\mathbf{z}\leq 0$ *then, it is said to be negative semi-definite matrix.*

A given matrix, $\mathbf{A}$ is positive semi–definite iff the eigenvalues $\lambda_i\geq 0$.
If $\mathbf{A}$ is of dimension $m\times n$ with $m<n$. Then, $\mathbf{A}$ cannot be decomposed into the eigenvalues and eigenvectors as Eq. (6). Although eigen decomposition

cannot be applied to $\mathbf{A}$, the singular value decomposition can be used to decompose $\mathbf{A}$. The singular value decomposition of $\mathbf{A}$ having $m \times n-$dimension and rank r is given as:

$$\mathbf{A} = \mathbf{U}\Sigma\mathbf{V}^{\mathrm{T}} \tag{7}$$

where the columns of $\mathbf{U}$ $(m \times m)$ are eigenvectors of the matrix $\mathbf{A}\mathbf{A}^{\mathrm{T}}$, the columns of $\mathbf{V}$ $(n \times n)$ are eigenvectors of the matrix $\mathbf{A}^{\mathrm{T}}\mathbf{A}$, and $\mathbf{\Sigma}$ is $m \times n$-dimensional matrix having r singular values ($\sqrt{\lambda_i}$ as diagonal entries.

$$\begin{aligned}
\mathbf{A} &= \begin{bmatrix} | & | & & | \\ \mathbf{u}_1 & \mathbf{u}_2 & \dots & \mathbf{u}_m \\ | & | & & | \end{bmatrix} \left[\begin{matrix} \sqrt{\lambda_1} & & & & & \\ & \sqrt{\lambda_2} & & & 0 & \\ & & \ddots & & & \\ & & & \sqrt{\lambda_r} & & \\ 0 & & & & 0 & \\ & & & & & \ddots \\ & & & & & & 0 \end{matrix} \middle| \; 0 \right] \begin{bmatrix} — & \mathbf{v}_1^{\mathrm{T}} & — \\ — & \mathbf{v}_2^{\mathrm{T}} & — \\ & \vdots & \\ — & \mathbf{v}_m^{\mathrm{T}} & — \\ — & \mathbf{v}_{m+1}^{\mathrm{T}} & — \\ — & \mathbf{v}_{m+2}^{\mathrm{T}} & — \\ & \vdots & \\ — & \mathbf{v}_n^{\mathrm{T}} & — \end{bmatrix} \\
\mathbf{A} &= [\mathbf{U}_r \, \mathbf{U}_{m-r}] \begin{bmatrix} \mathbf{\Sigma}_{r \times r} & \mathbf{0}_{r \times n-r} \\ \mathbf{0}_{m-r \times r} & \mathbf{0}_{m-r \times n-r} \end{bmatrix} \begin{bmatrix} \mathbf{V}_r^{\mathrm{T}} \\ \mathbf{V}_{n-r}^{\mathrm{T}} \end{bmatrix} \\
\mathbf{A} &= \sqrt{\lambda_1}\mathbf{u}_1\mathbf{v}_1^{\mathrm{T}} + \sqrt{\lambda_2}\mathbf{u}_2\mathbf{v}_2^{\mathrm{T}} + \ldots + \sqrt{\lambda_r}\mathbf{u}_r\mathbf{v}_r^{\mathrm{T}}
\end{aligned} \tag{8}$$

Eq. (8) can be interpreted factorization of $\mathbf{A}$ by summation of r one-rank factors. There are four spaces associated with different two types of eigenvalues: (i) $\lambda_i > 0$, $i = 1, 2, \ldots, r$, and (ii) zero, $i = r + 1, \ldots, morn$. All four fundamental spaces are as follows:

1. The first r columns of $\mathbf{U}$ span the subspace spanned by the column space of $\mathbf{A}$, i.e., $\mathcal{C}(\mathbf{U}_r) \sim \mathcal{C}(\mathbf{A})$
2. The last $(m-r)$ columns of $\mathbf{U}$ are a basis for the left null space of $\mathbf{A}$, i.e., $\mathcal{C}(\mathbf{U}_{m-r}) \sim \mathcal{N}(\mathbf{A}^{\mathrm{T}})$
3. The first r columns of $\mathbf{V}$ are a basis for the row space of $\mathbf{A}$ (or column space of $\mathbf{A}^{\mathrm{T}}$, i.e., $\mathcal{C}(\mathbf{V}_r) \sim \mathcal{C}(\mathbf{A}^{\mathrm{T}})$.
4. The last $(n-r)$ columns of $\mathbf{V}$ are a basis for the null space of $\mathbf{A}$, $\mathcal{C}(\mathbf{V}_{n-r}) \sim \mathcal{N}(\mathbf{A})$.