Singular Value Decomposition and Applications

F. de Sá, M. Velloso

August 2022

1 Introduction

2 Matrices, vectors and products

We begin by introducing some notation and interpretations of the multiplication operations among matrices with matrices and matrices with and vectors. These will be essential for interpreting Singular Value Decomposition intuitively.

Throughout this text, vectors $\mathbf{a} \in \mathbb{R}^m$ are represented by $m \times 1$ column matrices $\begin{bmatrix} a^1 & a^2 & \dots & a^m \end{bmatrix}^\mathsf{T} \in \mathbb{R}^{m \times 1}$. Vectors (and matrices) are bold-typed while its components (entries) are not.

First we highlight two complementary ways to see matrices and their products. We can interpret a matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$ as an array of n column vectors $\mathbf{a}_i \in \mathbb{R}^{m \times 1}$, with $i = 1, \dots, n$:

$$\mathbf{A} = \begin{bmatrix} | & | & & | \\ \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \\ | & | & & | \end{bmatrix}, \tag{1}$$

or as a column array of m transpose vectors $\boldsymbol{\alpha}_i^\mathsf{T} = \begin{bmatrix} \alpha_i^1, & \dots & \alpha_i^n \end{bmatrix} \in \mathbb{R}^{1 \times n}$, $i = 1, \dots, m$

$$\mathbf{A} = \begin{bmatrix} \boldsymbol{\alpha}_1^\mathsf{T} \\ \boldsymbol{\alpha}_2^\mathsf{T} \\ \vdots \\ \boldsymbol{\alpha}_m^\mathsf{T} \end{bmatrix}. \tag{2}$$

find out how to insert horizontal bars to represent the complete rows

The transpose vectors $\boldsymbol{\alpha}_i^\mathsf{T}$ are the ith row of the matrix $\mathbf{A}$.

add a simple ecxample displaying the $\mathbf{a}$ and $\boldsymbol{\alpha}$ vectors

The product of a matrix with a vector, say $\mathbf{Ax}$, can be seen as the $m \times 1$

matrix whose ith row is the scalar product of the $\boldsymbol{\alpha_i}$ vector with $\mathbf{x}$ vector, *i.e.*

$$\mathbf{A}\mathbf{x} = \begin{bmatrix} \boldsymbol{\alpha}_1^\mathsf{T} \\ \boldsymbol{\alpha}_2^\mathsf{T} \\ \vdots \\ \boldsymbol{\alpha}_m^\mathsf{T} \end{bmatrix} \mathbf{x} = \begin{bmatrix} \boldsymbol{\alpha}_1^\mathsf{T}\mathbf{x} \\ \boldsymbol{\alpha}_2^\mathsf{T}\mathbf{x} \\ \vdots \\ \boldsymbol{\alpha}_m^\mathsf{T}\mathbf{x} \end{bmatrix}. \tag{3}$$

Alternatively, we can think of this product as a linear combination of the column vectors of $\mathbf{A}$, with the coefficients being the components x^i of the vector $\mathbf{x}$:

$$\mathbf{A} = \begin{bmatrix} | & | & & | \\ \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \\ | & | & & | \end{bmatrix} \begin{bmatrix} x^1 \\ x^2 \\ \vdots \\ x^n \end{bmatrix} = \sum_i \mathbf{a}_i x^i \tag{4}$$

add a simple example

As for the products of matrices $\mathbf{A} \in \mathbb{R}^{m \times n}$ and $\mathbf{B} \in \mathbb{R}^{n \times k}$, we can see it as

$$\begin{aligned}\mathbf{A}\mathbf{B} &= \begin{bmatrix} \boldsymbol{\alpha}_1^\mathsf{T} \\ \boldsymbol{\alpha}_2^\mathsf{T} \\ \vdots \\ \boldsymbol{\alpha}_m^\mathsf{T} \end{bmatrix} \begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 & \ldots & \mathbf{b}_k \end{bmatrix} \\ &= \begin{bmatrix} \boldsymbol{\alpha}_1^\mathsf{T}\mathbf{b}_1 & \boldsymbol{\alpha}_1^\mathsf{T}\mathbf{b}_2 & \ldots & \boldsymbol{\alpha}_1^\mathsf{T}\mathbf{b}_k \\ \boldsymbol{\alpha}_2^\mathsf{T}\mathbf{b}_1 & \boldsymbol{\alpha}_2^\mathsf{T}\mathbf{b}_2 & \ldots & \boldsymbol{\alpha}_2^\mathsf{T}\mathbf{b}_k \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{\alpha}_m^\mathsf{T}\mathbf{b}_m & \boldsymbol{\alpha}_1^\mathsf{T}\mathbf{b}_2 & \ldots & \boldsymbol{\alpha}_m^\mathsf{T}\mathbf{b}_k \end{bmatrix}\end{aligned} \tag{5}$$

In this form it's clear why $\mathbf{A}^\mathsf{T}\mathbf{A}$, which equals

$$\mathbf{A}^\mathsf{T}\mathbf{A} = \begin{bmatrix} \mathbf{a}_1^\mathsf{T}\mathbf{a}_1 & \mathbf{a}_1^\mathsf{T}\mathbf{a}_2 & \ldots & \mathbf{a}_1^\mathsf{T}\mathbf{a}_k \\ \mathbf{a}_2^\mathsf{T}\mathbf{a}_1 & \mathbf{a}_2^\mathsf{T}\mathbf{a}_2 & \ldots & \mathbf{a}_2^\mathsf{T}\mathbf{a}_k \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{a}_m^\mathsf{T}\mathbf{a}_m & \mathbf{a}_1^\mathsf{T}\mathbf{a}_2 & \ldots & \mathbf{a}_m^\mathsf{T}\mathbf{a}_k \end{bmatrix}, \tag{6}$$

is called the *self-covariance matrix*. The diagonal entries are $\|\mathbf{a}_i\|^2$, while the off-diagonal ones are the projections of the ith and jth column vectors of $\mathbf{A}$. In a matrix where the columns are highly correlated, $\mathbf{A}^\mathsf{T}\mathbf{A}$ is a dense matrix, whereas diagonal terms dominate if the columns are less correlated..

We can also see the product of matrices $\mathbf{A}\mathbf{B}$ in the list of columns $\times$ column rows format.

add a simple example first

$$\mathbf{A}\mathbf{B} = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \ldots & \mathbf{a}_m \end{bmatrix} \begin{bmatrix} \boldsymbol{\beta}_1^\mathsf{T} \\ \boldsymbol{\beta}_2^\mathsf{T} \\ \vdots \\ \boldsymbol{\beta}_n^\mathsf{T} \end{bmatrix} = \mathbf{a}_1\boldsymbol{\beta}_1^\mathsf{T} + \mathbf{a}_2\boldsymbol{\beta}_2^\mathsf{T} + \cdots + \mathbf{a}_n\boldsymbol{\beta}_n^\mathsf{T} \tag{7}$$

3 Unitary Matrices

Besides the interpretations presented above, we also need to understand unitary (or othogonal) matrices to interpret SVD. Unitary matrices are special matrices $\mathbf{U} = \begin{bmatrix} \mathbf{u}_1 & \ldots & \mathbf{u}_n \end{bmatrix} \in \mathbb{R}^{mn}$ in which the column vectors $\mathbf{u}_i \in \mathbb{R}^m$ are all orthonormal, *i.e.* $\mathbf{u}_i^\intercal \mathbf{u}_j = \delta_{ij}$. This allows us to employ them as basis vectors, and expand any $\mathbf{x} \in \mathbb{R}^m$ as

$$\mathbf{x} = \sum_{i=1} x^i \mathbf{u}_i$$

> I think the column vectors form a basis only when $m = n$, in which case the basis is complete. Check this.

Unitary matrices are quite easy to invert. Their inverse $\mathbf{U}^{-1}$ is simply their transpose $\mathbf{U}^\intercal$. We can prove this claim by checking $U^\intercal U = 1$ in the column array of row vectors × array of column vectors interpretation

$$\mathbf{U}^\intercal \mathbf{U} = \begin{bmatrix} \mathbf{u}_1^\intercal \\ \mathbf{u}_2^\intercal \\ \vdots \\ \mathbf{u}_m^\intercal \end{bmatrix} \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_1 & \ldots & \mathbf{u}_m \end{bmatrix} = \begin{bmatrix} \mathbf{u}_1^\intercal \mathbf{u}_1 & \mathbf{u}_1^\intercal \mathbf{u}_2 & \ldots & \mathbf{u}_1^\intercal \mathbf{u}_k \\ \mathbf{u}_2^\intercal \mathbf{u}_1 & \mathbf{u}_2^\intercal \mathbf{u}_2 & \ldots & \mathbf{u}_2^\intercal \mathbf{u}_k \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{u}_m^\intercal \mathbf{u}_m & \mathbf{u}_1^\intercal \mathbf{u}_2 & \ldots & \mathbf{u}_m^\intercal \mathbf{u}_k \end{bmatrix}, \quad (8)$$

or, a little more subtly, we can prove $\mathbf{U}\mathbf{U}^\intercal = 1$ in the array of columns × column array of rows interpretation

$$\mathbf{U}\mathbf{U}^\intercal = \mathbf{u}_1 \mathbf{u}_1^\intercal + \mathbf{u}_2 \mathbf{u}_2^\intercal + \cdots + \mathbf{u}_m \mathbf{u}_m^\intercal \tag{9}$$

Recalling that $\mathbf{x}$ can be expanded in terms of the columns of $\mathbf{U}$, we have

$$\begin{aligned} \mathbf{U}\mathbf{U}^\intercal \mathbf{x} &= \mathbf{u}_1 \mathbf{u}_1^\intercal \mathbf{x} + \mathbf{u}_2 \mathbf{u}_2^\intercal \mathbf{x} + \cdots + \mathbf{u}_m \mathbf{u}_m^\intercal \mathbf{x} \\ &= \mathbf{u}_1 \mathbf{x}^1 + \mathbf{u}_2 \mathbf{x}^2 + \cdots + \mathbf{u}_m \mathbf{x}^m \\ &= \mathbf{x} \end{aligned} \tag{10}$$

since $\mathbf{x}$ is arbitrary, we conclude $\mathbf{U}\mathbf{U}^\intercal = 1$.

Unitary matrices have this name because they are analogous to unitary complex numbers in some sense. Unitary complex numbers have modulus 1 and thus preserve the modulus of any complex number upon multiplication. Similarly, unitary matrices preserve the norm, or more generally, the inner product, of any vector. To see this, let $\mathbf{y}_1 = \mathbf{U}^\intercal \mathbf{x}_1$ and $\mathbf{y}_2 = \mathbf{U}^\intercal \mathbf{x}_1$. It's easy to see that $\mathbf{y}_2^\intercal \mathbf{y}_1 = \mathbf{x}_2^\intercal \mathbf{U}\mathbf{U}^\intercal \mathbf{x}_1 = \mathbf{x}_2^\intercal \mathbf{x}_1$.

4 Singular Value Decomposition

Now we turn to our main goal: to make sens of the Singular Values Decomposition. Simply stated, the decomposition consists of expressing **any** matrix as the product of insightful matrices

$$\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^\intercal \tag{11}$$

5 Applications

5.1 Solving linear systems