Singular Value Decomposition

Nilanjan Chakraborty

April 14, 2019

In this note, we will study singular value decomposition (SVD). There are multiple ways in which a matrix can be decomposed, i.e., written as a product of simpler matrices. SVD is a popular matrix decomposition scheme that has applications across a wide variety of fields including robotics and machine learning. Let $\mathbf{T}$ be a $m \times n$ matrix with real entries, i.e., $\mathbf{T} \in \mathbb{R}^{m \times n}$. The matrix $\mathbf{T}$ can also be thought of as a linear operator from $\mathbb{R}^n$ to $\mathbb{R}^m$, i.e., $\mathbf{T} : \mathbb{R}^n \to \mathbb{R}^m$. Let the rank of $\mathbf{T}$ be r. The singular value decomposition of $\mathbf{T}$ is given by

$$\mathbf{T} = \mathbf{U}\mathbf{S}\mathbf{V}^{\mathbf{T}} \tag{1}$$

where $\mathbf{U} \in \mathbb{R}^{m \times m}$, $\mathbf{V} \in \mathbb{R}^{n \times n}$ are orthonormal matrices (i.e., $\mathbf{U}\mathbf{U}^T = \mathbf{U}^T\mathbf{U} = \mathbf{I}_{m \times m}$ and $\mathbf{V}\mathbf{V}^T = \mathbf{V}^T\mathbf{V} = \mathbf{I}_{n \times n}$) and $\mathbf{S}$ is a diagonal matrix with the first r diagonal entries being the singular values of T and the last $n - r$ diagonal entries being 0. Note that the columns of $\mathbf{U}$ form an orthonormal basis for $\mathbb{R}^m$ and the columns of $\mathbf{V}$ forms an orthonormal basis for $\mathbb{R}^n$.

Remark 0.1. In the context of our robotics class, one can think of $\mathbf{T}$ as the manipulator Jacobian, $\mathbf{J}$, or the grasp matrix, $\mathbf{G}$. For the manipulator Jacobian, in general, $m = 6$, and n is the degree-of-freedom of the robot. For the grasp matrix, in general, $m = 6$, and $n = kn_c$, where n_c is the number of contacts and $k = 1, 3, 4$ depending on whether the contact is frictionless contact, point contact with friction, or soft finger contact.

Definition 0.1. A square symmetric $n \times n$ matrix, $\mathbf{A}$, with real entries is called a positive semidefinite (psd) matrix iff $\mathbf{x}^T\mathbf{A}\mathbf{x} \geq 0$, $\forall \mathbf{x} \in \mathbb{R}^n$. If the inequality is strict, i.e., $\mathbf{x}^T\mathbf{A}\mathbf{x} > 0$, then $\mathbf{A}$ is called positive definite. Alternatively, $\mathbf{A}$ is psd iff all of its eigenvalues are non-negative.

Recall that the four fundamental subspaces associated with a matrix $\mathbf{T}$ are:

1. Range space of $\mathbf{T}$ ($\mathcal{R}(\mathbf{T}) \subseteq \mathbb{R}^m$)
2. Range space of $\mathbf{T}^T$ ($\mathcal{R}(\mathbf{T}^T) \subseteq \mathbb{R}^n$)
3. Null space of $\mathbf{T}$ ($\mathcal{N}(\mathbf{T}) \subset \mathbb{R}^n$)
4. Null space of $\mathbf{T}^T$ ($\mathcal{R}(\mathbf{T}) \subset \mathbb{R}^m$)

Let us define two matrices $\mathbf{L} = \mathbf{T}^T\mathbf{T} \in \mathbb{R}^{n \times n}$ and $\mathbf{M} = \mathbf{T}\mathbf{T}^T \in \mathbb{R}^{m \times m}$. Then, both $\mathbf{L}$ and $\mathbf{M}$ are real symmetric positive definite matrices. Furthermore, the four fundamental subspaces associated with the matrix $\mathbf{T}$ has equivalent characterization in terms of the matrices $\mathbf{L}$ and $\mathbf{M}$. Thus, all the subspaces can be thought of as subspaces connected to positive semidefinite matrices.

1. $\mathcal{N}(\mathbf{T}) = \mathcal{N}(\mathbf{L})$. Let $\{\boldsymbol{\phi}_1, \boldsymbol{\phi}_2, \ldots, \boldsymbol{\phi}_{n-r}\}$, be an orthonormal basis for $\mathcal{N}(\mathbf{T})$. These are the orthonormal eigenvectors corresponding to the 0 eigenvalue of matrix $\mathbf{L}$.
2. $\mathcal{N}(\mathbf{T}^T) = \mathcal{N}(\mathbf{M})$. Let $\{\boldsymbol{\psi}_1, \boldsymbol{\psi}_2, \ldots, \boldsymbol{\psi}_{m-r}\}$, be an orthonormal basis for $\mathcal{N}(\mathbf{T}^T)$. These are the orthonormal eigenvectors corresponding to the 0 eigenvalue of matrix $\mathbf{M}$.
3. $\mathcal{R}(\mathbf{T}^T) = \mathcal{R}(\mathbf{L})$. Let $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_r\}$, be an orthonormal basis for $\mathcal{R}(\mathbf{T}^T)$. These are the orthonormal eigenvectors corresponding to the positive eigenvalue of matrix $\mathbf{L}$, arranged as $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r > 0$.
4. $\mathcal{R}(\mathbf{T}) = \mathcal{R}(\mathbf{M})$. Let $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_r\}$, be an orthonormal basis for $\mathcal{R}(\mathbf{T})$, where

$$\mathbf{u}_j = \frac{1}{s_j}\mathbf{T}\mathbf{v}_j, \quad j = 1, 2, \ldots, r. \tag{2}$$

   Here $s_j = \sqrt{\lambda_j}$ are called the singular values of $\mathbf{T}$. Thus, **singular values** of $\mathbf{T}$ are the same as the square root of positive eigenvalues of $\mathbf{L}$ or $\mathbf{T}^T\mathbf{T}$.

Lemma 0.1. *The relationship between the basis of* $\mathbf{T}$ *and* $\mathbf{T}^T$ *are as follows:*

$$\boxed{\mathbf{T}\mathbf{v}_j = s_j\mathbf{u}_j}, \qquad \boxed{\mathbf{T}^T\mathbf{u}_j = s_j\mathbf{v}_j}, \qquad j = 1, 2, \ldots, r. \tag{3}$$

Proof. From Equation (2), multiplying both sides by s_j gives $\mathbf{T}\mathbf{v}_j = s_j\mathbf{u}_j$. Therefore,

$$\begin{aligned} &\mathbf{T}^T\mathbf{T}\mathbf{v}_j = s_j\mathbf{T}^T\mathbf{u}_j \\ \Rightarrow &\mathbf{L}\mathbf{v}_j = s_j\mathbf{T}^T\mathbf{u}_j \qquad (\because \mathbf{T}^T\mathbf{T} = \mathbf{L}) \\ \Rightarrow &s_j^2\mathbf{v}_j = s_j\mathbf{T}^T\mathbf{u}_j \qquad (\because \mathbf{L}\mathbf{v}_j = \lambda_j\mathbf{v}_j \text{ and } s_j = \sqrt{\lambda_j}) \\ \Rightarrow &s_j\mathbf{v}_j = \mathbf{T}^T\mathbf{u}_j \end{aligned} \tag{4}$$

□

We will now define two matrices $\mathbf{U}$ and $\mathbf{V}$ as follows:

$$\mathbf{U} = [\mathbf{u}_1 \ \mathbf{u}_2 \ \ldots \ \mathbf{u}_r \ \boldsymbol{\psi}_1 \ \boldsymbol{\psi}_2 \ \ldots \ \boldsymbol{\psi}_{m-r}] \tag{5}$$

$$\mathbf{V} = [\mathbf{v}_1 \ \mathbf{v}_2 \ \ldots \ \mathbf{v}_r \ \boldsymbol{\phi}_1 \ \boldsymbol{\phi}_2 \ \ldots \ \boldsymbol{\phi}_{n-r}] \tag{6}$$

Please note that $\mathbf{U} \in \mathbb{R}^{m\times m}$ is a orthonormal matrix and the columns of $\mathbf{U}$ form a basis for $\mathbb{R}^m$. Similarly, $\mathbf{V} \in \mathbb{R}^{n\times n}$ is a orthonormal matrix and the columns of $\mathbf{V}$ form a basis for $\mathbb{R}^n$. Let us define the diagonal matrix $\mathbf{S}$ as

$$\mathbf{S} = \begin{bmatrix} s_1 & & & & & & \\ & s_2 & & & & & \\ & & \ddots & & & & \\ & & & s_r & & & \\ & & & & 0 & & \\ & & & & & \ddots & \\ & & & & & & 0 \end{bmatrix} \tag{7}$$

We will now prove that the matrix $\mathbf{T}$ can be written in the form of Equation (1).

Lemma 0.2. *Let* $\mathbf{T}$ *be a* $\mathbb{R}^{m\times n}$ *matrix. Then* $\mathbf{T} = \mathbf{U}\mathbf{S}\mathbf{V}^{\mathbf{T}}$*, with* $\mathbf{U}$*,* $\mathbf{V}$*, and* $\mathbf{S}$ *as defined above.*

Proof.

$$
\begin{aligned}
\mathbf{T}\mathbf{V} &= \mathbf{T}[\mathbf{v}_1\ \mathbf{v}_2\ \ldots\ \mathbf{v}_r\ \boldsymbol{\phi}_1\ \boldsymbol{\phi}_2\ \ldots\ \boldsymbol{\phi}_{n-r}] \\
&= [\mathbf{T}\mathbf{v}_1\ \mathbf{T}\mathbf{v}_2\ \ldots\ \mathbf{T}\mathbf{v}_r\ \mathbf{T}\boldsymbol{\phi}_1\ \mathbf{T}\boldsymbol{\phi}_2\ \ldots\ \mathbf{T}\boldsymbol{\phi}_{n-r}] \\
&= [s_1\mathbf{u}_1\ s_2\mathbf{u}_2\ \ldots\ s_r\mathbf{u}_r\ \mathbf{0}\ \mathbf{0}\ \ldots\ \mathbf{0}] \\
&= [\mathbf{u}_1\ \mathbf{u}_2\ \ldots\ \mathbf{u}_r\ \boldsymbol{\psi}_1\ \boldsymbol{\psi}_2\ \ldots\ \boldsymbol{\psi}_{m-r}]
\begin{bmatrix}
s_1 & & & & & & \\
& s_2 & & & & & \\
& & \ddots & & & & \\
& & & s_r & & & \\
& & & & 0 & & \\
& & & & & \ddots & \\
& & & & & & 0
\end{bmatrix} \\
&= \mathbf{U}\mathbf{S} \\
\Rightarrow \mathbf{T}\mathbf{V}\mathbf{V}^T &= \mathbf{U}\mathbf{S}\mathbf{V}^{\mathbf{T}} \\
\Rightarrow \mathbf{T} &= \mathbf{U}\mathbf{S}\mathbf{V}^{\mathbf{T}} \qquad (\because \mathbf{V}\mathbf{V}^T = \mathbf{I})
\end{aligned}
$$

□

Computationally, SVD is a standard function provided by any decent Linear Algebra software including MATLAB. Once you have computed the SVD of a matrix, the last $(n-r)$ columns of $\mathbf{V}$ will give the orthonormal basis for the null space of $\mathbf{T}$, i.e., $\mathcal{N}(\mathbf{T})$. Thus an element, $\boldsymbol{\alpha}$ in the null space of $\mathbf{T}$ can be written as:

$$\boldsymbol{\alpha} = \sum_{j=1}^{n-r} w_j \boldsymbol{\phi}_j \tag{8}$$

For a given matrix, $\boldsymbol{\phi}_j$ are obtained from SVD and you can choose the weights w_j to generate any element in $\mathcal{N}(\mathbf{T})$. Also, please note that the first r columns of $\mathbf{U}$ gives the basis for $\mathcal{R}(\mathbf{T})$ and any vector in $\mathcal{R}(\mathbf{T})$ can be written analogous to Equation (8).

Representing/Computing an element of the null space of Jacobian or Grasp matrix

Recall that in general the manipulator Jacobian $\mathbf{J}$ is a $6\times n$ matrix, where n is the degree-of-freedom of the robot. For redundant robots $n > 6$. For an indoor mobile manipulator with a 7-DoF arm, $n = 10$. The grasp matrix $\mathbf{G}$ is a $6 \times kn_c$ matrix. For a 3-fingered grip with soft finger contact model, $k = 4$, $n_c = 3$. Thus the grasp matrix $\mathbf{G}$ is a 6×12 matrix. In the context of the discussion in this lecture, $m = 6$, $n = 10$ for the Jacobian of mobile manipulator and $m = 6$, $n = 12$ for the Grasp matrix for a 3-fingered soft contact grasp.

Remember that the Jacobian is dependent on the configuration of the robot. The null space of $\mathbf{J}$ represents the set of all joint velocities that causes self-motion of the manipulator without affecting the motion of the end effector. We have learned two ways about finding a vector in the null space. The first is by selecting any vector in $\mathbb{R}^n$ and then projecting it into $\mathcal{N}(\mathbf{J})$ using the equation

$$\dot{\mathbf{\Theta}}_n = (\mathbf{I} - \mathbf{J}^T(\mathbf{J}\mathbf{J}^T)^{-1}\mathbf{J})\dot{\mathbf{\Theta}}'$$

where $\dot{\mathbf{\Theta}}'$ is any vector in $\mathbb{R}^n$. The second is by explicitly computing a orthonormal basis for $\mathcal{N}(\mathbf{J})$ and then using Equation (8) with any arbitrary choice of the parameters w_j. Both the representations and ways of computing the null space vector are useful. For optimization problems however, the form in Equation (8) is more useful because it allows one to search over the weight space to get the best possible null space vector.

In a similar vein, the grasp matrix is dependent on the contact points and hence the contact normals. The null space of $\mathbf{G}$ represents the set of all internal (or squeezing) forces that does not cause any net wrench on the object. For the grasp matrix also, we can find a vector in the null space (or the squeezing force) by selecting any vector in $\mathbb{R}^n$ and then projecting it into $\mathcal{N}(\mathbf{G})$ using the equation

$$\dot{\boldsymbol{\beta}}_n = (\mathbf{I} - \mathbf{G}^T(\mathbf{G}\mathbf{J}^T)^{-1}\mathbf{G})\boldsymbol{\beta}'$$

where $\boldsymbol{\beta}'$ is any vector in $\mathbb{R}^n$. The second is by explicitly computing a orthonormal basis for $\mathcal{N}(\mathbf{G})$ and then using Equation (8) with any arbitrary choice of the parameters w_j. Both the representations and ways of computing the null space vector are useful. However, for grasping one has to be more careful since the magnitudes of the contact wrench at each contact (i.e., elements of $\boldsymbol{\beta}$ for each contact, where $\boldsymbol{\beta}$ is as defined in the lectures on grasping) should also satisfy the friction constraints. Therefore grasping force computation problems are naturally set up as optimization problems. For grasping force optimization problems however, the form in Equation (8) is more useful because it allows one to search over the weight space to optimize certain criterion like minimizing the squeezing force at each contact while ensuring that the object does not fall. This is particularly important for fragile objects.