mr. sc. Vida Zadelj Martić *Institute of Geomatics, Faculty of Geodesy, University of Zagreb, Kačićeva 26, 10000 Zagreb, vzadelj@geof.hr*

THE SINGULAR VALUE DECOMPOSITION AND APPLICATIONS IN GEODESY

Abstract

The paper considers the singular value decomposition (SVD) of a general matrix. Some immediate applications, such as determining the spectral and Frobenius norm, rank and pseudoinverse of the matrix are described. Applications also include approximating the given matrix by a matrix of a lower rank. It is also shown how to use SVD for solving the homogeneous linear system and the least squares problem. The paper consists of three parts:

1.) The singular value decomposition,
2.) Some applications of the singular value decomposition,
3.) Applications in geodesy.

Keywords

Singular value decomposition
Unitary matrices
Frobenius norm
Spectral norm
Pseudoinverse
Homogeneous system of linear equations
Least squares problem
Rank

1. THE SINGULAR VALUE DECOMPOSITION

1.1 BASIC NOTIONS AND NOTATION

By $\square^{m\times n}$ ($\square^{m\times n}$) is denoted the set of $m \times n$ complex (real) matrices.

Let $A\in\square^{m\times n}$. Two most common vector subspaces associated with A are the range of A,

$$\Re(A) = \{Ax\colon x\in\square^{n}\}\subseteq\square^{m},$$

and the kernel (null-subspace) of A:

$$N(A) = \{x\colon Ax = 0\} \subseteq\square^{n}.$$

The range of A is spanned by the columns of A, so it is sometimes called the column space of A. If A then $\Re(A)$ and $N(A)$ are appropriately defined using the vector spaces $\square^{m}$ and $\square^{n}$.

The dimensions of $\Re(A)$ and $N(A)$ are the rank and the defect of A.

The spectral radius of a square matrix A,

$$spr(A) = max\{|\lambda|;\lambda\in\sigma(A)\},$$

is the largest distance of an eigenvalue of A to zero. Here $\sigma(A)$ is the spectrum of A which is the set of eigenvalues of A.

The trace of A is the sum of its diagonal elements:

$$tr(A) = \sum_{k=1}^{n} a_{kk}\,.$$

It can be shown that $tr(A)$ is the sum of the eigenvalues of A.

The diagonal part of $A=(a_{i,j})$ is the diagonal matrix:

$$diag(a_{11},\ldots,a_{nm}) = \begin{pmatrix} a_{11} & & \\ & \ddots & \\ & & a_{mn} \end{pmatrix}.$$

If A is not square then $diag(A)$ has additional zero rows or columns.

The spectral and the Frobenius norm of an $m \times n$ matrix are defined by

$$\|A\|_2 = \sqrt{spr(A^* A)}\ ,$$

and

$$\|A\|_F = \sqrt{tr(A^* A)} = \sqrt{\sum_{i,j}|a_{i,j}|^2}\ ,$$

respectively. Here A^* is the hermitian transpose of A. If A is real, then $A^* = A^\tau$ is just the transpose of A.

The matrix A is normal if $A^*A = AA^*$. The complex (real) matrix U is unitary (orthogonal) if $U^*U = UU^* = I$.

1.2 THE SINGULAR VALUE DECOMPOSITION

The following theorem serves as the definition of the singular value decomposition.

Theorem 1.

Let $A\in\square^{m\times n}$. Then there exist unitary matrices $U\in\square^{m\times m}$ and $V\in\square^{n\times n}$,

so that

$$U^* AV = \Sigma = diag(\sigma_1, \sigma_2,\ldots,\sigma_{\min\{m,n\}}),$$

where $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_{\min\{m,n\}} \geq 0$. The nonnegative numbers $\sigma_1, \sigma_2, \ldots, \sigma_{\min\{m,n\}}$ are the singular values of A and the columns of $U(V)$ are the left (right) singular vectors of A.

Matrix Σ is uniquely determined by the matrix A. However, the matrices U and V are not unique. If the singular values are multiple, then U and V can be post-multiplied by an arbitrary block-diagonal unitary matrix, whose diagonal blocks are of appropriate dimensions.

The singular value decomposition has hundreds of applications. It is an excellent tool in matrix theory. It is often used for solving different matrix problems which arise in science, economy, engineering, medicine, and even in human sciences. Data mining, web searching, image recognition, just to mention some contemporary problems which use SVD, often of large matrices. Its widespread use is enhanced by the fact that there exist excellent, efficient and accurate methods for computing it. There are several classes of methods, the most important are: one-sided Jacobi methods, divide and conquer (DC), differential qd (DQD) and QR methods.

We will first list several immediate consequences of this decomposition.

Let $A = U\Sigma V^*$ be the singular value decomposition of the matrix $A \in \mathbb{C}^{m\times n}$, so that:

$$\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_r > \sigma_{r+1} = \ldots = \sigma_p = 0, \quad p = \min\{m,n\}.$$

Let $U = [u_1,\ldots,u_m]$, $V = [v_1,\ldots,v_m]$ be the column representations (often called partitions) of U and V, respectively.

Then it is easy to show that the following holds:

(i) $\quad rank(A) = r$,

(ii) $\quad N(A) = span\{v_{r+1},\ldots,v_n\}$,

(iii) $\quad \Re(A) = span\{u_1,\ldots,u_r\}$,

(iv) $\quad A = \sum_{I=1}^{r} \sigma_i u_i v_i^* = U_r \Sigma_r V_r^*$, *such that,*

$\quad U_r = [u_1, \ldots, u_r] \in \mathbb{C}^{m\times r}$,

$\quad V_r = [v_1, \ldots, v_r] \in \mathbb{C}^{n\times r}$,

$\quad \Sigma_r = diag(\sigma_1, \ldots, \sigma_r) \in \mathbb{C}^{r\times r}$,

(v) $\quad \| A \|_F^2 = \sigma_1^2 + \ldots + \sigma_r^2$,

(vi) $\quad \| A \|_F = \sigma_1$.

It can also be shown that the distance (in spectral norm) of a square matrix A and the set of singular matrices of the same dimension, is equal to the smallest singular value of A. This follows from a more general result, the Ekhard, Young and Mirsky theorem.

Theorem 2.

Let $A \in \mathbb{C}^{m\times n}$ and let $A = U\Sigma V^*$ be the singular value decomposition of A so that:

$$\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_r > \sigma_{r+1} = \ldots = \sigma_p = 0, \quad p = \min\{m,n\},$$

holds. If $k < r = rank(A)$ and $C_k = \sum_{i=1}^{k} \sigma_i u_i v_i^*$ then:

$$\min_{rank(K)=k} \|A-K\|_2 = \|C-C_k\|_2 = \sigma_{k+1}.$$

Thus, the closest rank k approximation of A is obtained by using the first k singular values and vectors of A.

2. SOME APPLICATIONS OF THE SINGULAR VALUE DECOMPOSITION

Here, we will address several applications.

2.1 DETERMINING THE PSEUDOINVERSE OF A GENERAL MATRIX

Here, we will consider the Moore-Penrose pseudoinverse.

Let $A \in \mathbb{C}^{m\times n}$. The matrix $X \in \mathbb{C}^{n\times m}$ is pseudoinverse of $A \in \mathbb{C}^{m\times n}$, if the following four conditions are fulfilled:

(i) $\quad AXA = A$,

(ii) $\quad XAX = X$,

(iii) $\quad (AX)^* = AX$,

(iv) $\quad (XA)^* = XA$.

The conditions *(i)* - *(iv)* ensure that the matrix X is unique. It is usually denoted by $A^\dagger$.

Let $A = U\Sigma V^*$ be the singular value decomposition of A. Then the pseudoinverse $A^\dagger$ of the matrix A is given by the expression:

$$A^\dagger = V\Sigma^+ U^*,$$

where:

$$\Sigma^\dagger = diag(\sigma_1^+, \ldots, \sigma^+_{\min\{m,n\}}) \in \mathbb{C}^{n\times m},$$

and:

$$\sigma_i^+ = \left\{ \begin{matrix} \frac{1}{\sigma_i} & \sigma_i > 0 \\ 0 & \sigma_i = 0 \end{matrix} \right\}.$$

By using the SVD of A, it is easy to prove the following properties of $A^\dagger$.

1. $\quad (A^\dagger)^\dagger = A$,
2. $\quad (A^\tau)^\dagger = (A^\dagger)^\tau$,
3. $\quad (\overline{A})^\dagger = (\overline{A}^\dagger)$,
4. $\quad rank(A) = rank(A^\dagger) = rank(AA^\dagger) = rank(A^\dagger A)$,
5. $\quad$ If $A \in \mathbb{C}^{m\times n}$ with rank n, then $A^\dagger = (A^*A)^{-1}A^*$ *and* $A^\dagger A = I_n$,
6. $\quad$ If $A \in \mathbb{C}^{m\times n}$ with rank m, then $A^\dagger = A^*(AA^*)^{-1}$ *and* $AA^\dagger = I_m$.

In addition, one can easily prove that $AA^\dagger$ $(A^\dagger A)$ is orthogonal projector onto $\Re(A)$ $(\Re(A^*))$. This fact is important when solving the overdetermined system of the linear equations $Ax = b$, $A \in \mathbb{C}^{m\times n}$ in the least squares sense.

2.2 SOLVING A HOMOGENEOUS SYSTEM OF LINEAR EQUATIONS

Let $A \in \mathbb{R}^{m \times n}$ be the matrix of rank r. Then, solving the homogeneous system of linear equations, $Ax = 0$, reduces to determining the null-subspace of the matrix A. From item (ii) of immediate consequences of the SVD, one finds out that:

$$N(A) = span\{v_{r+1},\ldots,v_n\}.$$

Here $v_{r+1},\ldots,v_n$ are the last n-r columns of the matrix V from the SVD of A. Since V is unitary, these vectors form an orthonormal basis of the subspace $N(A)$. Since the SVD of A can be very accurately computed, this approach for solving the homogenous linear system is most commonly used.

2.3 SOLVING THE LINEAR LEAST SQUARES PROBLEM

Let us consider an overdetermined system:

$$\sum_{J=1}^{n} A_{ij}\beta_j = y_i; \qquad i = 1,\ldots,m.$$

of m linear equations in n unknown coefficients $\beta_1, \beta_2,\ldots, \beta_n$, with $m > n$. In matrix form it reads:

$$A\beta = y,$$

where

$$A = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ A_{m1} & A_{m2} & & A_{mn} \end{bmatrix} \in \mathbb{R}^{m \times n}, \quad \beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix} \in \mathbb{R}^n, \quad y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix} \in \mathbb{R}^m.$$

Such a system, has usually no solution. Therefore, it is reformulated into the least squares problem,

$$min_\beta \| y - A\beta \|^2.$$

Suppose that there exists $n \times m$, matrix S, so that AS is an orthogonal projection onto $\Re(A)$. In that case, the solution is given by:

$$\beta = Sy,$$

because,

$$A\beta = A(Sy) = (AS)\, y,$$

is the orthogonal projection of y onto $\Re(A)$.

If $A = U\Sigma V^\tau$ is the singular value decomposition of A, then the pseudoinverse $A^\dagger$ of A, given by $A^\dagger = U\Sigma^+ V^\tau$ has the property that $AA^\dagger$ is orthogonal projector on $\Re(A)$. Indeed we have:

$$AA^\dagger = U\Sigma V^\tau V\Sigma^+ U^\tau = UPU^\tau,$$

where the square matrix P is obtained from Σ, by replacing the non-zero diagonal elements with ones. Matrices A and Σ have the same rank, and $AA^\dagger = UPU^\tau$ is the orthogonal projector onto $\Re(A)$. Thus, $A^\dagger$ is just the wanted matrix S. The final conclusion is that unknown $\beta \in \mathbb{R}^n$ is given by the expression:

$$\beta = A^\dagger y = V\Sigma^+ U^\tau y.$$

Such β is the solution of the least squares problem and among all solutions, it has the minimum Euclidean norm.

3. APPLICATIONS IN GEODESY

The singular value decomposition and the pseudoinverse are taught in the courses: Analysis and processing of geodetic measurements, and Special algorithms of geodetic measurement processing at the Faculty of Geodesy, University of Zagreb.

Using any web browser, one can find many applications of SVD in Geodesy. Such as the following presentation:

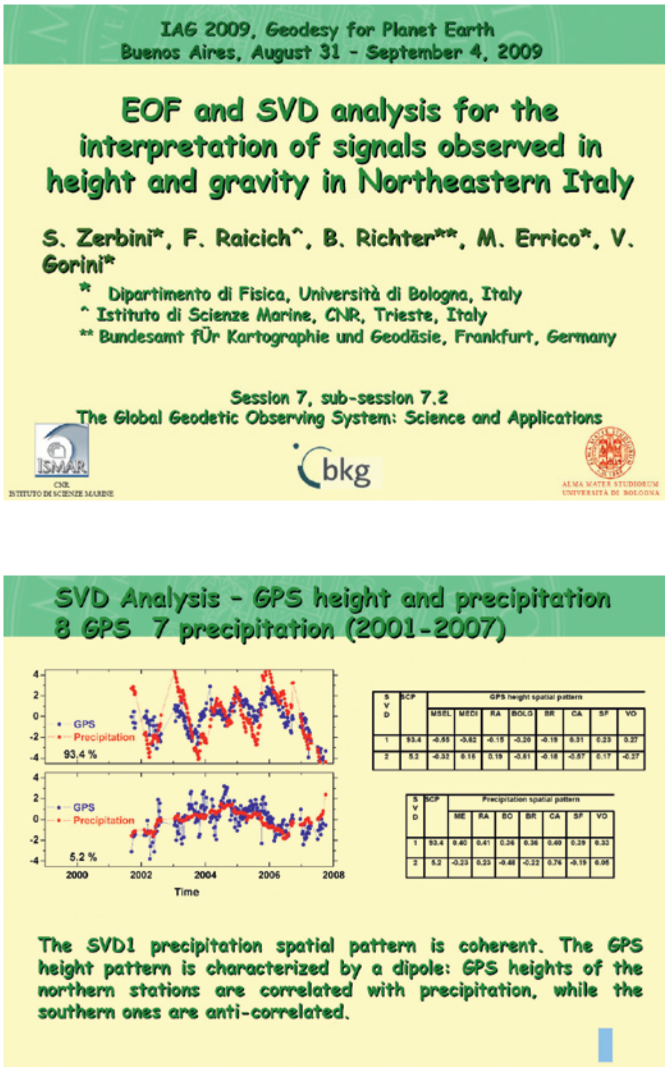

FIGURE 1. http://www.iag2009.com.ar/presentations

REFERENCES

- G.H. Golub and C.F. VanLoan (1989): Matrix Computations, The Johns Hopkins University Press.
- V. Zadelj-Martić (1999): Parallel Algorithms for Butterfly Matrices, M.Sc. Thesis (Croatian), University of Zagreb.
- URL-1: http://www.iag2009.com.ar/presentations.
- URL-2: http://en.wikipedia.org/wiki/Singular_value_decomposition.