The SVD-Fundamental Theorem of Linear Algebra

A. G. Akritas[1], **G. I. Malaschonok**[2], **P. S. Vigklas**[1]

[1]Department of Computer and Communication Engineering, University of Thessaly
GR-38221 Volos, Greece
{akritas, pviglas}@uth.gr

[2]Laboratory for Algebraic Computations, Tambov State University
Internatsionalnaya st. 33, 392622 Tambov, Russia
malaschonok@math-iu.tstu.ru

Received: 10.01.2006 **Revised:** 06.05.2006 **Published online:** 18.05.2006

Abstract. Given an $m \times n$ matrix A, with $m \geq n$, the four subspaces associated with it are shown in Fig. 1 (see [1]).

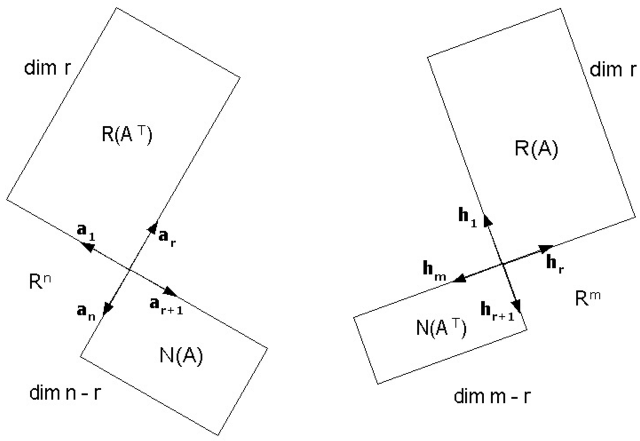

Fig. 1. The row spaces and the nullspaces of A and A^T; $\mathbf{a}_1$ through $\mathbf{a}_n$ and $\mathbf{h}_1$ through $\mathbf{h}_m$ are abbreviations of the *alignerframe* and *hangerframe* vectors respectively (see [2]).

The Fundamental Theorem of Linear Algebra tells us that $N(A)$ is the orthogonal complement of $R(A^T)$. These four subspaces tell the whole story of the Linear System $A\mathbf{x} = \mathbf{y}$. So, for example, the absence of $N(A^T)$ indicates that a solution always exists, whereas the absence of $N(A)$ indicates that this solution

is unique. Given the importance of these subspaces, computing bases for them is the gist of Linear Algebra. In "Classical" Linear Algebra, bases for these subspaces are computed using Gaussian Elimination; they are orthonormalized with the help of the Gram-Schmidt method. Continuing our previous work [3] and following Uhl's excellent approach [2] we use SVD analysis to compute orthonormal bases for the four subspaces associated with A, and give a 3D explanation. We then state and prove what we call the "SVD-Fundamental Theorem" of Linear Algebra, and apply it in solving systems of linear equations.

Keywords: fundamental theorem of linear algebra, singular values decomposition, pseudoinverse, orthonormal bases, systems of linear equations.

1 Introduction

1.1 Basic definitions

Let A be an $m \times n$ matrix with $m \geq n$. Then one form of the singular-value decomposition of A is

$$A = U_h \Sigma V_a^T \tag{1}$$

where U_h and V_a^T are orthonormal and Σ is diagonal. The indices a and h indicate matrices with aligner and hanger vectors respectively. That is, $U_h^T U_h = I_m$, $V_a V_a^T = I_n$, U_h is $m \times m$, V_a is $n \times n$ and

$$\Sigma = \begin{pmatrix} \sigma_1 & 0 & \cdots & 0 & 0 & \\ 0 & \sigma_2 & \cdots & 0 & 0 & \\ \vdots & \vdots & \ddots & \vdots & \vdots & \mathbf{0} \\ 0 & 0 & \cdots & \sigma_{r-1} & 0 & \\ 0 & 0 & \cdots & 0 & \sigma_r & \\ & & \mathbf{0} & & & \end{pmatrix} \tag{2}$$

is an $m \times n$ diagonal matrix (the same dimensions as A). In addition $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n \geq 0$. The σ_i's are called the *singular values* (or the *stretchers* [2]) of A and the number of the non-zero σ_i's is equal to the rank of A. The ratio $\frac{\sigma_1}{\sigma_n}$, if $\sigma_n \neq 0$ can be regarded as a condition number of the matrix A.

It is easily verified that the singular-value decomposition can be also written as

$$A = U_h \Sigma V_a^T = \sum_{i=1}^{n} \sigma_i \mathbf{h}_i \mathbf{a}_i^T. \tag{3}$$

The matrix $\mathbf{h}_i \mathbf{a}_i^T$ is the *outer product* of the i-th row of U_h with the corresponding row of V_a. Note that each of these matrices can be stored using only $m + n$ locations rather than mn locations.

1.2 The SVD-fundamental theorem of linear algebra

We now prove constructively (showing at the same time how SVD analysis is done) that,

$$A\mathbf{a}_i = \sigma_i \mathbf{h}_i \ , \tag{4}$$

which we call the SVD-Fundamental Theorem of Linear Algebra:

Theorem 1. *For an m by n matrix, with $m \geq n$, $A\colon R^n \to R^m$ there exists an orthonormal basis $\{\mathbf{a}_1, ..., \mathbf{a}_n\}$ of R^n, obtained from the Spectral Theorem*[1] [2], *as the eigenvectors of $A^T A$. Define*

(i) $\sigma_i = \|A\mathbf{a}_i\|, \quad i = 1, 2, \ldots, r,$
the nonzero stretchers, where $\|\|$ is the Euclidean norm, and

(ii) $\mathbf{h}_i = \dfrac{1}{\sigma_i} A\mathbf{a}_i, \quad i = 1, 2, \ldots, r,$
where $r \leq n$.

Proof. Then we clearly have

$$\sigma_i \mathbf{h}_i = \sigma_i \frac{1}{\sigma_i} A\mathbf{a}_i = A\mathbf{a}_i$$

or

$$A\mathbf{a}_i = \sigma_i \mathbf{h}_i, \quad i = 1, 2, \ldots, r.$$

□

[1] The Spectral Theorem states that if A is a square symmetric matrix (i.e. $A = A^T$), then there is an orthonormal basis of the column space $R(A)$ consisting of unit eigenvectors of A.

Note, that from $A\mathbf{a}_i = \sigma_i\mathbf{h}_i$ it follows that $AV_a = U_h\Sigma$ and multiplying both sides by V_a^T we obtain (1).

We will use this theorem in solving systems of linear equations. Also note that if $\sigma_i = 0$ then $A\mathbf{a}_i = 0$ and the corresponding aligner vector $\mathbf{a}_i$ belongs to the *nullspace* of A, $N(A)$. If on the other hand $\sigma_i \neq 0$, then $A\mathbf{a}_i = \sigma_i\mathbf{h}_i$, the corresponding aligner vector $\mathbf{a}_i$ belongs to the row space of A, $R(A^T)$, whereas the corresponding hanger vector $\mathbf{h}_i$ belongs to the column space of A, $R(A)$. All of the above mentioned vectors form an orthonormal base of the subspace in which they belong.

So far, we have obtained orthonormal bases for all subspaces associated with A except for the nullspace of A^T, $N(A^T)$. One way to obtain an orthonormal basis for $N(A^T)$ is through the application of Spectral Theorem on AA^T, instead of A^TA, completing thus the SVD analysis of A.

1.3 Geometric interpretation

The figure below (Fig. 2) represents the operation of an $m \times n$ matrix A on a vector $\mathbf{x} \in R^n$ (the left space). That is, it shows schematically what happens when an arbitrary vector from A's domain (the space corresponding dimensionally to A's row dimension n) is mapped by A onto the range space (the space corresponding dimensionally to A's column dimension m). Hence Fig. 2 shows what happens to

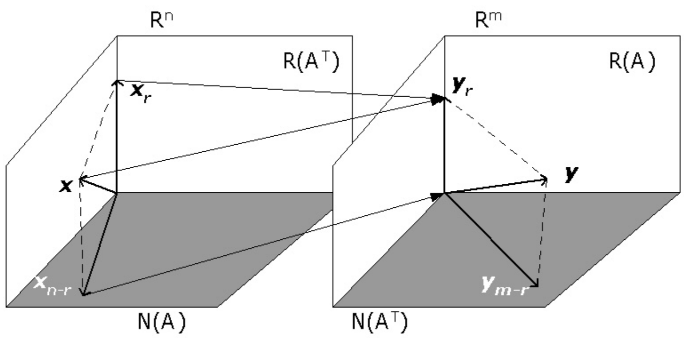

Fig. 2. The domain of A is the space on the left-hand side, whereas its range is the space on the right-hand side. Matrix A operates on a vector $\mathbf{x}$ with both a row-space and a nullspace component, $\mathbf{x}_r$ and $\mathbf{x}_{n-r}$ respectively, mapping it on the space on the right hand side.

$\mathbf{x}$ from the left space as A transforms it to the range, the right space. In short, this figure represents the fundamental theorem of linear algebra (forward problem).

Both spaces are made up of two orthogonal subspaces; R^n (the space on the left-hand side) comprises the row and nullspace, $R(A^T)$ and $N(A)$, respectively. The R^n space is spanned by the aligner vectors $\{\mathbf{a}_i\}$, whereas the hanger vectors $\{\mathbf{h}_i\}$, span the column space of A.

The inverse problem is depicted in Fig. 3.

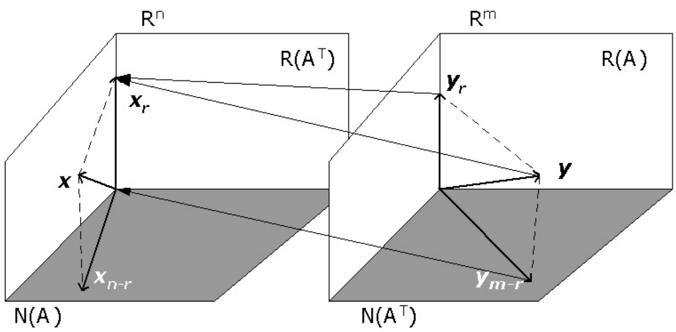

Fig. 3. The inverse problem.

We conclude this section by pointing out two facts regarding the linear system $A\mathbf{x} = \mathbf{y}$:

1. The absence of $N(A^T)$ means that $\mathbf{y} \in R(A)$, and the system $A\mathbf{x} = \mathbf{y}$ has *always*, at least, one solution. On the other hand the presence of $N(A^T)$ indicates that a solution will exist (in the classical sense) only if $\mathbf{y} \in R(A)$, which has to be investigated.

2. The absence of $N(A)$ means that if a solution exists, it is unique, whereas the presence of $N(A)$ indicates an infinity of solutions. These points will be examined in the next section.

2 Classical methods for computing bases for the four subspaces associated with matrix A

Let's us briefly review the traditional algorithms of linear algebra used for computing bases for the four subspaces. Remember that for the orthonormalization of the bases of the following spaces the Gram-Schmidt procedure is used.

- Algorithm to determine a basis for $R(A)$, the *column space* of matrix A, i.e the space spanned by the columns of A:
  1. Find the reduced echelon matrix, R_E, for the input matrix A.
  2. Identify the pivot columns of R_E.
  3. Identify the corresponding columns in the input matrix A.
  4. A basis for $R(A)$ is the set of pivot column vectors (of A) from step 3.
- Algorithm to determine a basis for $R(A^T)$, the *row space* of matrix A, i.e the space spanned by the rows of A:
  1. Find the reduced echelon matrix, R_E, for the input matrix A.
  2. Identify non-zero rows of R_E.
  3. A basis for $R(A^T)$ is the set of row vectors (of R_E) from step 2.
- Algorithm to determine a basis for $N(A)$, the *right nullspace* of matrix A, i.e the space spanned by solutions to $A\mathbf{x} = \mathbf{0}$:
  1. Find the reduced echelon matrix, R_E, for the input matrix A.
  2. Identify free variables and pivot variables of R_E.
  3. Set one free variable to 0, other free variables to 1, and solve for pivot variables.
  4. Repeat step 3 for each free variable.
  5. A basis for $N(A)$ is the set of special solution vectors from each step 3.
- Algorithm to determine a basis for $N(A^T)$, the *left nullspace* of matrix A, i.e the space spanned by solutions to $\mathbf{x}^T A = A^T\mathbf{x} = \mathbf{0}$:
  1. Find the transpose A^T.
  2. Perform the nullspace algorithm on A^T.

The bases obtained above are *not* orthonormal. This is where the Gram-Schmidt method comes in. Namely, any set of linearly independent vectors $\mathbf{a}_1, \ldots, \mathbf{a}_n$ can be converted into a set of orthonormal vectors $\mathbf{q}_1, \ldots, \mathbf{q}_n$ as follows: We first set

$\mathbf{a}'_1 = \mathbf{a}_1$, and then each $\mathbf{a}'_i$ is made orthogonal to the preceding $\mathbf{a}'_1, \dots, \mathbf{a}'_{i-1}$ by subtracting of the projections[2] of $\mathbf{a}_i$ in the directions of $\mathbf{a}'_1, \dots, \mathbf{a}'_{i-1}$:

$$\mathbf{a}'_i = \mathbf{a}_i - \sum_{j=1}^{i-1} \frac{\mathbf{a}'^T_j \mathbf{a}_i}{\mathbf{a}'^T_j \mathbf{a}'_j} \mathbf{a}'_j.$$

The i vectors $\mathbf{a}'_i$ span the same subspace as the $\mathbf{a}_i$. The vectors $\mathbf{q}_i = \mathbf{a}'_i / ||\mathbf{a}'_i||$ are orthonormal.

The problem with the Gram-Schmidt method is that before we can use it, we have to know that the given spanning set is linearly independent. Such a problem does not exist using SVD.

3 Solving systems of linear equations with SVD analysis

3.1 SVD and the pseudoinverse

The SVD analysis of a matrix A gives an easy and uniform way of computing its inverse A^{-1}, whenever it exists, or its pseudoinverse, A^+, otherwise. Namely if $A = U_h \Sigma V_a^T$, then the pseudoinverse of A is:

$$A^+ = V_a \Sigma^+ U_h^T \tag{6}$$

The singular values $\sigma_1, \dots, \sigma_r$ are located on the diagonal of the $m \times n$ matrix Σ, and the reciprocals of the singular values, $\frac{1}{\sigma_1}, \dots, \frac{1}{\sigma_r}$, are on the diagonal of the $n \times m$ matrix Σ^+. The pseudoinverse of A^+ is $A^{++} = A$. Whenever A^{-1} exists, then $A^+ = A^{-1}$.

Note that by making some simple calculations we see that $AA^+ = U_h U_h^T$, and in this way

$$AA^+A = A$$

Similarly, we can get $A^+A = V_a V_a^T$ and

$$A^+AA^+ = A^+$$

[2]It is well known that the projection of one vector $\mathbf{x}$ onto another $\mathbf{y}$ is given by the formula:

$$proj_{\mathbf{y}}\mathbf{x} = \frac{\mathbf{y} \cdot \mathbf{x}}{\mathbf{y} \cdot \mathbf{y}} \mathbf{y}. \tag{5}$$

It also follows that A^+A and AA^+ are symmetric, a unique property of the pseudoinverse.

If the rank of the $m \times n$ matrix A is $n < m$, then $\Sigma^T\Sigma$ is a $n \times n$ matrix of rank n and

$$A^TA = V_a\Sigma^TU_h^TU_h\Sigma V_a^T = V_a\Sigma^T\Sigma V_a^T$$

is invertible, so

$$(A^TA)^{-1}A^T = V_a(\Sigma^T\Sigma)^{-1}V_a^T(V_a\Sigma U_h^T)$$

and $(\Sigma^T\Sigma)^{-1}\Sigma = \Sigma^+$. Finally, we have

$$A^+ = V_a\Sigma^+U_h^T = (A^TA)^{-1}A^T$$

If $m = n$, then $\Sigma^+ = \Sigma^{-1}$ and so $A^+ = V_a\Sigma^{-1}U_h^T = A^{-1}$.

3.2 SVD and linear systems

From the previous discussion we see that, to solve

$$A\mathbf{x} = \mathbf{y} \tag{7}$$

we have to execute the following two steps:

Step 1. We first do an SVD analysis of the matrix $A{:}R^n \rightarrow R^m$, obtaining thus orthonormal bases for $R(A^T)$, R^n and R^m. Here, we could have $m \gtrless n$. The dimensions of the subspaces, $N(A)$ in R^n, and $R(A)$, $N(A^T)$ in R^m, are determined from the number r of the nonzero singular values, $\sigma_1, \ldots, \sigma_r$, where $r \leq min(m,n)$, which indicates the rank of the matrix A.

From the SVD analysis we have that the vectors $\{\mathbf{a}_1, ..., \mathbf{a}_r\}$ constitute an orthonormal basis of $R(A^T)$, whereas the vectors $\{\mathbf{a}_{r+1}, ..., \mathbf{a}_n\}$ an orthonormal basis of $N(A)$.

Likewise, the vectors $\{\mathbf{h}_1, ..., \mathbf{h}_r\}$ constitute an orthonormal basis of $R(A)$, whereas the vectors $\{\mathbf{h}_{r+1}, ..., \mathbf{h}_m\}$ an orthonormal basis of $N(A^T)$, (see Fig. 1).

Step 2. We perform the test

$$proj_{R(A)}\mathbf{y} \stackrel{?}{=} \sum_{k=1}^{r}(\mathbf{h}_k \cdot \mathbf{y})\mathbf{h}_k$$

to determine where $\mathbf{y}$ lies. To wit, we examine whether $\mathbf{y}$ coincides with its projection in $R(A)$, which is spanned by $\{\mathbf{h}_1, ..., \mathbf{h}_r\}$, (5). We distinguish two general cases:

Case A. $\mathbf{y} \in R(A)$: in this case the system has (at least) one solution. Since we now have $\mathbf{y} = proj_{R(A)}\mathbf{y}$, it follows that:

$$\mathbf{y} = proj_{R(A)}\mathbf{y} = \sum_{k=1}^{r} (\mathbf{h}_k \cdot \mathbf{y}) \frac{A.\mathbf{a}_k}{\sigma_k}$$

due to (4). Furthermore, due to the linearity of matrix operations, the last equation becomes

$$\mathbf{y} = A \sum_{k=1}^{r} \frac{(\mathbf{h}_k \cdot \mathbf{y})}{\sigma_k} \mathbf{a}_k$$

and comparing with (7) we obtain:

$$\mathbf{x} = \sum_{k=1}^{r} \frac{(\mathbf{h}_k \cdot \mathbf{y})}{\sigma_k} \mathbf{a}_k \tag{8}$$

as the solution of the linear system (7). But, the number of solutions of (7) depends on the existence or not of the nullspace $N(A)$, so we further examine two subcases:

A.1 If $N(A)$, the nullspace of A, does *not* exist, then the linear system (7) has one solution. We prove this by contradiction: Let's suppose that there is another analogous solution of (7), say, $\mathbf{x}^*$. This means that now both $A\mathbf{x} = \mathbf{y}$ and $A\mathbf{x}^* = \mathbf{y}$ hold true. Subtracting, we have

$$A(\mathbf{x} - \mathbf{x}^*) = \mathbf{y} - \mathbf{y} = \mathbf{0}.$$

Because there is no $N(A)$, it follows that only $A\mathbf{0} = \mathbf{0}$ is possible. This fact tells us that $\mathbf{x} - \mathbf{x}^* = \mathbf{0}$ or $\mathbf{x} = \mathbf{x}^*$.

A.2 If $N(A)$, the nullspace of A, *does exist*, then the linear system has infinite solutions, given by

$$\mathbf{x}_{inf} = \mathbf{x} + \sum_{k=r+1}^{n} \mu_k \mathbf{a}_k \tag{9}$$

where the vectors $\mathbf{a}_k$, $k = r+1,\ldots,n$ are an orthonormal basis (span) of $N(A)$. Since $N(A)$ exists, it follows that there are points $\mathbf{x}_N \in N(A)$ such that $A\mathbf{x}_N = \mathbf{0}$. Each of these points $\mathbf{x}_N$ can be expressed as a linear combination of the basis vectors, i.e.

$$\mathbf{x}_N = \sum_{k=r+1}^{n} \mu_k \mathbf{a}_k \tag{10}$$

where $\mathbf{a}_k$ are the base vectors of $N(A)$. Therefore, for each of these points we have

$$A(\mathbf{x}+\mathbf{x}_N) = A\mathbf{x} + A\mathbf{x}_N = \mathbf{y} + \mathbf{0} = \mathbf{y} \tag{11}$$

or $\mathbf{x}_{inf} = \mathbf{x} + \mathbf{x}_N$ or formula (9).

Case B. $\mathbf{y} \notin R(A)$: in this case the system does not have a solution in the classical sense. In this case we can only compute the best approximate solution given also by the same formula (8). When we say "the best approximate solution", we mean a vector $\mathbf{y}_r$ in $R(A)$, that is closest to $\mathbf{y}$. In other words, the vector $(\mathbf{y} - \mathbf{y}_r)$ has to be perpendicular to each one of the vectors spanning $R(A)$ (see Fig. 3). Indeed for each ξ with $1 \le \xi \le r$ we have[3]

$$\begin{aligned}(\mathbf{y}-\mathbf{y}_r)\cdot\mathbf{h}_\xi &= (\mathbf{y}-A\mathbf{x}_r)\cdot\mathbf{h}_\xi = \Big(\mathbf{y}-A\Big(\sum_{k=1}^{r}\frac{(\mathbf{h}_k\cdot\mathbf{y})}{\sigma_k}\mathbf{a}_k\Big)\Big)\cdot\mathbf{h}_\xi \\ &= \Big(\mathbf{y}-\sum_{k=1}^{r}\frac{(\mathbf{h}_k\cdot\mathbf{y})}{\sigma_k}\underbrace{A\mathbf{a}_k}_{=\sigma_k\mathbf{h}_k}\Big)\cdot\mathbf{h}_\xi = \Big(\mathbf{y}-\sum_{k=1}^{r}(\mathbf{h}_k\cdot\mathbf{y})\mathbf{h}_k\Big)\cdot\mathbf{h}_\xi \\ &= \mathbf{y}\cdot\mathbf{h}_\xi - \sum_{k=1}^{r}(\mathbf{h}_k\cdot\mathbf{y})(\mathbf{h}_k\cdot\mathbf{h}_\xi) = \mathbf{y}\cdot\mathbf{h}_\xi - \mathbf{h}_\xi\cdot\mathbf{y} = 0\end{aligned}$$

This tell us that, indeed, $\mathbf{y} - A\mathbf{x}_r$ is perpendicular to $R(A)$, and the solution we obtain from (8) is the best approximate solution.

[3]Keep in mind that $\mathbf{h}_k$ and $\mathbf{h}_\xi$ are orthonormal, i.e.

$$\mathbf{h}_k\cdot\mathbf{h}_\xi = \begin{cases}0 & \text{if } k \neq \xi \\ 1 & \text{if } k = \xi\end{cases}.$$

Concluding we see that in both cases, A. and B., the solution to (7) is given by (8). However, note that

$$
\begin{aligned}
\mathbf{x} &= \sum_{k=1}^{r} \frac{(\mathbf{h}_k \cdot \mathbf{y})}{\sigma_k}\mathbf{a}_k \\
&= \left\{ \frac{\mathbf{h}_1 \cdot \mathbf{y}}{\sigma_1}, \dots, \frac{\mathbf{h}_r \cdot \mathbf{y}}{\sigma_r} \right\} \cdot \underbrace{\{\mathbf{a}_1, \dots, \mathbf{a}_r\}}_{V_a^T} \\
&= V_a \cdot \begin{pmatrix} \frac{1}{\sigma_1} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \frac{1}{\sigma_r} \end{pmatrix} \cdot \{\mathbf{h}_1 \cdot \mathbf{y}, \dots, \mathbf{h}_r \cdot \mathbf{y}\} \\
&= V_a \Sigma^+ \{\mathbf{h}_1 \cdot \mathbf{y}, \dots, \mathbf{h}_r \cdot \mathbf{y}\} \\
&= V_a \Sigma^+ \underbrace{\{\mathbf{h}_1, \dots, \mathbf{h}_r\}}_{U_h^T} \mathbf{y} = V_a \Sigma^+ U_h^T \mathbf{y} = A^+ \mathbf{y}
\end{aligned}
\tag{12}
$$

So, we conclude that in all cases $\mathbf{x} = A^+\mathbf{y}$ holds true.

3.3 Example

Here, we present an example showing in detail the process described above for solving a linear system using SVD. The following data depicts U.S. imports of telecommunication equipment from Greece (in thousands of dollars) from 2000 to 2004

Table 1. U.S. imports from Greece (in thousands of dollars)

Year	2000	2001	2002	2003	2004
US$	6076	4741	2502	3083	6787

We want to fit a polynomial of degree 2 to these points. To wit, the polynomial $f(x) = c_1 + c_2x + c_3x^2$ should be such that $f(x_i)$ should be as close as possible to y_i, where the points $\{x_i, y_i\}$, $i = 1, \dots, 5$ represent, respectively, the year (x_i) and the value of imported telecommunication equipment (y_i). Our linear system of equations $A\mathbf{c} = \mathbf{y}$, where $A : \mathbb{R}^3 \to \mathbb{R}^5$ is

$$\begin{pmatrix} 1 & 2000 & 4000000 \\ 1 & 2001 & 4004001 \\ 1 & 2002 & 4008004 \\ 1 & 2003 & 4012009 \\ 1 & 2004 & 4016016 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} 6076 \\ 4741 \\ 2502 \\ 3083 \\ 6787 \end{pmatrix}.$$

Performing the SVD analysis of the matrix $A = U_h \Sigma V_a^T$ we obtain the following matrices:

$$U_h = \begin{pmatrix} -0.44632 & -0.632771 & -0.534896 & -0.0257609 & -0.337079 \\ -0.446766 & -0.317017 & 0.267074 & 0.280919 & 0.741388 \\ -0.447213 & -0.000947734 & 0.534522 & -0.688191 & -0.201691 \\ -0.44766 & 0.315438 & 0.267448 & 0.636669 & -0.472466 \\ -0.448107 & 0.632139 & -0.534149 & -0.203636 & 0.269848 \end{pmatrix},$$

$$\Sigma = \begin{pmatrix} 8.96218 \cdot 10^6 & 0 & 0 \\ 0 & 3.16227 & 0 \\ 0 & 0 & 9.33546 \cdot 10^{-7} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

$$V_a^T = \begin{pmatrix} -2.495 \cdot 10^{-7} & -0.0004995 & -1 \\ -0.000999001 & -0.999999 & 0.0004995 \\ -1. & 0.000999001 & -2.49501 \cdot 10^{-7} \end{pmatrix}.$$

The information on orthonormal bases obtained from the SVD decomposition of matrix A (see 3.2) is presented below:

- Base of $R(A)$: $\{\mathbf{h}_1, \mathbf{h}_2, \mathbf{h}_3\}$

$$= \left\{ \begin{pmatrix} -0.44632 \\ -0.446766 \\ -0.447213 \\ -0.44766 \\ -0.448107 \end{pmatrix}, \begin{pmatrix} -0.632771 \\ -0.317017 \\ -0.000947734 \\ 0.315438 \\ 0.632139 \end{pmatrix}, \begin{pmatrix} -0.534896 \\ 0.267074 \\ 0.534522 \\ 0.267448 \\ -0.534149 \end{pmatrix} \right\}.$$

- Base of $R(A^T)$: $\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\}$

$$= \left\{ \begin{array}{c} \left(-2.495 \cdot 10^{-7} \ -0.0004995 \ -1\right), \\ \left(-0.000999001 \ -0.999999 \ 0.0004995\right), \\ \left(-1. \ 0.000999001 \ -2.49501 \cdot 10^{-7}\right) \end{array} \right\}.$$

- Base of $N(A)$: $\nexists$.
- Base of $N(A^T)$: $\{\mathbf{h}_4, \mathbf{h}_5\}$

$$= \left\{ \begin{pmatrix} -0.0257609 \\ 0.280919 \\ -0.688191 \\ 0.636669 \\ -0.203636 \end{pmatrix}, \begin{pmatrix} -0.337079 \\ 0.741388 \\ -0.201691 \\ -0.472466 \\ 0.269848 \end{pmatrix} \right\}.$$

We now compute the projection of y on R(A)

$$proj_{R(A)}\mathbf{y} = \sum_{k=1}^{r} (\mathbf{h}_k \cdot \mathbf{y})\mathbf{h}_k = (5334.01, 5219.07, 3058.74, 4257.25, 5898.8).$$

We observe that

$$proj_{R(A)}\mathbf{y} \neq \mathbf{y} = (6076,\ 4741,\ 2502,\ 3083,\ 6787)$$

which means (see Case B.) that $\mathbf{y} \notin R(A)$. Finally, the solution to our system is given by (12):

$$\mathbf{c} = (3.69257 \cdot 10^9,\ -3.68885 \cdot 10^6,\ 921.286).$$

In Fig. 4 the data points and the polynomial fitted to them are shown.

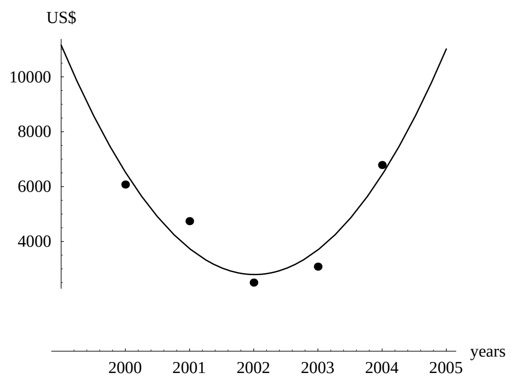

Fig. 4. The data points and the closest polynomial $f(x) = 3.69257\,10^9 - 3.68885\,10^6\,x + 921.286\,x^2$ fitted to them.

4 Conclusions

In this paper we have presented the "SVD-Fundamental Theorem" of Linear Algebra and used it in solving systems of linear equations. We hope that the geometrical interpretation of SVD makes our approach more appealing for educational purposes.

References

1. G. Strang, The Fundamental Theorem of Linear Algebra, *American Mathematical Monthly*, **100**(9), pp. 848–855, 1993.

2. B. Davis, J. Uhl, *Matrices, Geometry & Mathematica*, in: *Calculus & Mathematica*, Math Everywhere, Inc., 1999.

3. G. A. Akritas, G. I. Malaschonok, Applications of Singular Value Decomposition (SVD), *Mathematics and Computers in Simulation*, Special Issue on Applications of Computer Algebra in Science, Engineering, Simulation & Special Software, **67**, pp. 15–31, 2004.

4. E. Gidon, *The Fundamental Theorem of Linear Algebra*, `http://geosci.uchicago.edu/~gidon/geosci236/fundam`, Dept. of the Geophysical Sciences, Chicago University.

5. G. H. Golub, C. F. Van Loan, *Matrix Computations*, The Johns Hopkins University Press, Baltimore, 1983.

6. D. Kahaner, C. Moler, S. Nash, *Numerical Methods and Software*, Printice-Hall International, Inc., 1989.

7. S. Leach, *Singular value decomposition – A Primer*, Unpublished Manuscript, `http://citeseer.csail.mit.edu/651803.html`, Dept. of Computer Science, Brown University.

8. G. Strang, *Introduction to Linear Algebra*, Wellesley-Cambridge Press, 2003.

9. G. Strang, *Linear Algebra and Its Applications*, 4th edition, Brooks Cole, 2005.

10. R. Schmitt, *Notes on Matrix Subspaces*, `http://camelot.mssm.edu/~arc/Teaching/SupplementaryReading/`