De Moor B., ``Generalizations of the OSVD: Structure, properties and applications", in In SVD and Signal Processing, II: Algorithms, Analysis and Applications, R. Vaccaro, ed., Elsevier Science Publishers B.V. (North-Holland), 1991, pp. 83-98., Lirias number: 181232.

Generalizations of the OSVD Structure, Properties and Applications

Bart De Moor *

Department of Electrical Engineering
Katholieke Universiteit Leuven
Kardinaal Mercierlaan 94
B-3030 Leuven
Belgium
tel:32/16/220931 fax: 32/16/221855
email: demoor@esat.kuleuven.be.ac
or na.demoor@na-net.stanford.edu

ESAT-SISTA report 1990-13

Abstract

In this paper, we provide a state-of-the-art survey of a recently discovered set of generalizations of the *Ordinary Singular Value Decomposition*, which contains all existing generalizations for 2 matrices (such as the *Product SVD* and the *Quotient SVD*) and for 3 matrices (such as the *Restricted SVD*), as special cases. We present the main Theorem and a discussion on some structural and geometrical properties of all *Generalized Singular Value Decompositions* (**GSVD**). A proposal for a standardized nomenclature is made as well. We conclude this paper with a survey of possible applications of these **GSVD**s, including a literature survey and a nice classification of numerical stochastic realization algorithms.

Keywords: Ordinary, Product, Quotient, Restricted Singular Value Decomposition, stochastic realization.

*Dr. De Moor is a *research associate* of the Belgian National Fund for Scientific Research (NFWO). Part of this research was supported by the Belgian Program on Interuniversity Attraction Poles initiated by the Belgian State Science Policy Programming (Prime Minister's Office) and the European Community Research Program ESPRIT, Basic Research Action nr.3280. The scientific responsibility is assumed by its author.

1. Introduction

The *ordinary singular value decomposition* (**OSVD**) has become an important tool in the analysis and numerical solution of numerous problems (see e.g. [13] [17] for properties and applications.) Not only does it allow for an elegant problem formulation, but at the same time it provides geometrical and algebraic insight together with an immediate numerically robust implementation [17]. In [20, p.78], credit for the first proofs of the **OSVD** is given to Beltrami [2], Jordan [19], Sylvester [28] and Autonne [1]. Recently, several generalizations to the **OSVD** have been proposed and their properties analysed. The best known example is the *Generalized SVD* for two matrices, as introduced in [32] and refined in [24], which we propose to rename as the *Quotient SVD* (**QSVD**) [7]. A specific reason for this name is the relation of this matrix factorization to the SVD of the 'quotient' of two matrices while the main motivation is of course the fact that there are several other similar generalizations. For instance, a *Product induced SVD*, also for two matrices, was proposed in [15], where it was called the ΠSVD. It was a formalization of ideas in [18]. We shall refer to it as the **PSVD** (see [7]). In [33], another generalization, this time for three matrices, was proposed. In [8] we have called it the *Restricted SVD* (**RSVD**) and analysed its properties in detail.
In [7] we have proposed a standardized nomenclature for all kinds of singular value decompositions. Besides the **OSVD**, **PSVD**, **QSVD** and **RSVD**, we have also included there the so-called *Structured Singular Value* which occurs in robust control theory and a (possibly) corresponding decomposition [14] (**SSVD**), the *Takagi Singular Value Decomposition* which applies for symmetric complex matrices (**TSVD**) [29] and the *Unordered Singular Value Decomposition* (**USVD**) [3] [10] which is useful when analysing analytic properties of the singular values and vectors of parametrized matrices.

This paper is organised as follows:

- In section 2, we present the main Theorem and explain in more detail its structure and some properties.
- Some additional structural and geometrical properties are summarized in section 3.
- In section 4, we discuss the potential numerical advantages of the **GSVDs** with some small examples. We also give a literature survey of possible applications.

2. A tree of generalizations of the OSVD

Theorem 1

Generalized Singular Value Decompositions for k matrices.

Consider a set of k matrices with compatible dimensions: A_1 $(n_0 \times n_1), A_2$ $(n_1 \times n_2), \ldots,$ A_{k-1} $(n_{k-2} \times n_{k-1}), A_k$ $(n_{k-1} \times n_k)$. Then there exist

- *Unitary matrices U_1 $(n_0 \times n_0)$ and V_k $(n_k \times n_k)$*
- *Matrices $D_j, j = 1, 2, \ldots, (k-1)$ of the form*

$$\underset{n_{j-1}\times n_j}{D_j} =$$

$$\begin{array}{l} \\ r_j^1 \\ r_{j-1}^1 - r_j^1 \\ r_j^2 \\ r_{j-1}^2 - r_j^2 \\ r_j^3 \\ \cdots \\ \cdots \\ r_j^j \\ n_{j-1} - r_{j-1} - r_j^j \end{array} \begin{array}{c} \begin{array}{ccccccc} r_j^1 & r_j^2 & r_j^3 & \cdots & \cdots & r_j^j & n_j - r_j \end{array} \\ \left(\begin{array}{ccccccc} I & 0 & 0 & \cdots & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & \cdots & 0 & 0 \\ 0 & I & 0 & \cdots & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & \cdots & 0 & 0 \\ 0 & 0 & I & \cdots & \cdots & 0 & 0 \\ 0 & \cdots & \cdots & \cdots & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & \cdots & \cdots & \cdots & I & 0 \\ 0 & 0 & \cdots & \cdots & \cdots & 0 & 0 \end{array} \right) \end{array} \tag{1}$$

where

$$r_0 = 0$$

$$r_j = \sum_{i=1}^{j} r_j^i = rank(A_j) \tag{2}$$

- A matrix S_k of the form

$$\underset{n_{k-1}\times n_k}{S_k} =$$

$$\begin{array}{l} \\ r_k^1 \\ r_{k-1}^1 - r_k^1 \\ r_k^2 \\ r_{k-1}^2 - r_k^2 \\ r_k^3 \\ \cdots \\ \cdots \\ r_k^k \\ n_{k-1} - r_{k-1} - r_k^k \end{array} \begin{array}{c} \begin{array}{ccccccc} r_k^1 & r_k^2 & r_k^3 & \cdots & \cdots & r_k^k & n_k - r_k \end{array} \\ \left(\begin{array}{ccccccc} S_k^1 & 0 & 0 & \cdots & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & \cdots & 0 & 0 \\ 0 & S_k^2 & 0 & \cdots & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & \cdots & 0 & 0 \\ 0 & 0 & S_k^3 & \cdots & \cdots & 0 & 0 \\ 0 & \cdots & \cdots & \cdots & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & \cdots & \cdots & \cdots & S_k^k & 0 \\ 0 & 0 & \cdots & \cdots & \cdots & 0 & 0 \end{array} \right) \end{array} \tag{3}$$

where

$$r_k = \sum_{i=1}^{k} r_k^i = rank(A_k) \tag{4}$$

and the $r_k^i \times r_k^i$ matrices S_k^i are diagonal with positive diagonal elements. Expressions for the integers r_j^i are given in section 3. They are ranks of certain matrices in the constructive proof of this Theorem [11].

- Nonsingular matrices X_j $(n_j \times n_j)$ and Z_j, $j = 1, 2, \ldots, (k-1)$ where Z_j is either $Z_j = X_j^{-*}$ or either $Z_j = X_j$ (i.e. both choices are always possible)

such that the given matrices can be factorized as

$$\begin{aligned} A_1 &= U_1 D_1 X_1^{-1} \\ A_2 &= Z_1 D_2 X_2^{-1} \\ A_3 &= Z_2 D_3 X_3^{-1} \\ \cdots &= \cdots \end{aligned}$$

$$
\begin{aligned}
A_i &= Z_{i-1} D_i X_i^{-1} \\
\ldots &= \ldots \\
A_k &= Z_{k-1} S_k V_k^*
\end{aligned}
$$

Observe that the matrices D_j in (1) and S_k in (3) are in general not diagonal. Their only non-zero blocks however are diagonal block matrices. We propose to label them as *quasi-diagonal* matrices. The matrices $D_j, j = 1, \ldots, k-1$ are quasi-diagonal, their only nonzero blocks being identity matrices. The matrix S_k is quasi-diagonal and its nonzero blocks are diagonal matrices with positive diagonal elements. Observe that we always take the last factor in every factorization as the inverse of a nonsingular matrix, which is only a matter of convention (Another convention would result in a modified definition of the matrices Z_i). As to the name of a certain **GSVD**, we propose to adopt the following convention:

Definition 1

The nomenclature for GSVDs

If $k = 1$ in Theorem 1, then the corresponding factorization of the matrix A_1 will be called the ordinary singular value decomposition.

If for a matrix pair $A_i, A_{i+1}, 1 \leq i \leq k-1$ in Theorem 1, we have that

$$Z_i = X_i$$

then, the factorization of the pair will be said to be of P-type.

If on the other hand, for a matrix pair $A_i, A_{i+1}, 1 \leq i \leq k-1$ in Theorem 1, we have that

$$Z_i = X_i^{-*}$$

the factorization of the pair will be said to be of Q-type.

The name of a **GSVD** *of the matrices A_i, $i = 1, 2, \ldots k > 1$ as in Theorem 1, is then obtained by simply enumerating the different factorization types.*

Let us give some examples.

Example 1:

Consider two matrices A_1 $(n_0 \times n_1)$ and A_2 $(n_1 \times n_2)$. Then, we have two possible **GSVDs**:

	P-type	Q-type
A_1	$U_1 D_1 X_1^{-1}$	$U_1 D_1 X_1^{-1}$
A_2	$X_1 S_2 V_2^*$	$X_1^{-*} S_2 V_2^*$

The P-type factorization corresponds to the **PSVD** as in [15] (called ΠSVD there) and [7], while the Q-type factorization is nothing else than the **QSVD** in [17] [24] [32] (called generalized SVD there). This justifies the choice of names for the factorization of pairs: A P-type factorization is precisely the kind of transformation that occurs in the **PSVD** while a Q-type factorization occurs in the **QSVD**.

Example 2

The **RSVD** for three matrices (A_1, A_2, A_3) as introduced and analysed in [8] [33] has the form:

$$\begin{aligned} A_1 &= U_1 S_1 X_1^{-1} \\ A_2 &= X_1^{-*} S_2 X_2^{-1} \\ A_3 &= X_2^{-*} S_3 V_3^* \end{aligned}$$

where S_1, S_2, S_3 are certain *quasi-diagonal* matrices. It can be verified that this **RSVD** can be rearranged into a **QQ-SVD** that is conform with the structure of Theorem 1.

Example 3
Let us write down the **PQQP-SVD** for 5 matrices:

$$\begin{aligned} A_1 &= U_1 D_1 X_1^{-1} \\ A_2 &= X_1 D_2 X_2^{-1} \\ A_3 &= X_2^{-*} D_3 X_3^{-1} \\ A_4 &= X_3^{-*} D_4 X_4^{-1} \\ A_5 &= X_4 S_5 V_5^* \end{aligned}$$

We also introduce the following notation, using powers, which symbolize a certain repetition of a letter or of a sequence of letters:

- $\mathbf{P}^3\mathbf{Q}^2$**-SVD** = **PPPQQ-SVD**
- $(\mathbf{PQ})^2\mathbf{Q}^3(\mathbf{PPQ})^2$**-SVD** = **PQPQQQQPPQPPQ-SVD**

Despite the fact that there are 2^{k-1} different sequences of letters **P** and **Q** at level $k > 1$, not all of these sequences correspond to different **GSVDs**. The reason for this is that for instance the **QP-SVD** of (A^1, A^2, A^3) can be obtained from the **PQ-SVD** of $((A^3)^*, (A^2)^*, (A^1)^*)$. Similarly, the $\mathbf{P}^2(\mathbf{QP})^3$**-SVD** of $(A^1, \ldots, A^9)$ is essentially the same as the $(\mathbf{PQ})^3\mathbf{P}^2$**-SVD** of $((A^9)^*, \ldots, (A^1)^*)$. The following table gives the number of *different* factorizations for k matrices.

	k even	k odd
number of different **GSVDs**	$\frac{1}{2}(2^{k-1} + 2^{k/2})$	$\frac{1}{2}(2^{k-1} + 2^{(k-1)/2})$

Finally, we'll spend some words on the proof of the main Theorem, a detailed exposition of which can be found in [11]. It is based on two basic ideas: First, there is an *inductive* argument which allows us to construct the **GSVD** of k matrices $A_1, \ldots, A_k$ from a corresponding one for $k-1$ matrices $A_1, \ldots, A_{k-1}$. A key result here is a certain block factorization lemma for partitioned matrices. Next, the already obtained **GSVD** of the $k-1$ matrices $A_1, \ldots, A_{k-1}$ has to be modified according to a certain algorithm, which we have called the *ripple-through-phenomenon* in [11]. For all details of the constructive proof, the interested reader is referred to [11].

3. Rank properties

It is possible to express the block dimensions of the quasi-diagonal matrices $D_j, j = 1, \ldots, k-1$ and S_k of Theorem 1, in terms of the ranks of the matrices $A_1, \ldots, A_k$ and concatenations and products thereof as was shown in [12].

Let's first consider the case of a **GSVD** that consists only of $P-type$ factorizations. Denote the rank of the product of the matrices A_i,A_{i+1}, ..., A_j with $i \leq j$ by

$$r_{i(i+1)\ldots(j-1)j} = rank(A_i A_{i+1} \ldots A_{j-1} A_j)$$

Theorem 2

On the structure of the $\mathbf{P^{k-1}}$-SVD

Consider a $\mathbf{P^{k-1}}$**-SVD** *of the matrices* $A_1, A_2, \ldots, A_k$. *Then, the block dimensions* r_i^j *that appear in Theorem 1 are given by:*

$$r_j^1 = r_{(1)(2)\ldots(j)} \tag{5}$$

$$r_j^i = r_{i(i+1)\ldots(j)} - r_{(i-1)(i)\ldots(j)} \tag{6}$$

with $r_j^i = r_i$ *if* $i = j$.

Next, consider the case of a **GSVD** that only consists of $Q-type$ factorizations. Denote the rank of the block bidiagonal matrix

$$\begin{pmatrix} A_i & 0 & 0 & \ldots & 0 & 0 & 0 \\ A_{i+1}^* & A_{i+2} & 0 & \ldots & 0 & 0 & 0 \\ 0 & A_{i+3}^* & A_{i+4} & \ldots & 0 & 0 & 0 \\ 0 & \ldots & \ldots & \ldots & \ldots & \ldots & 0 \\ 0 & \ldots & \ldots & \ldots & A_{j-3}^* & A_{j-2} & 0 \\ 0 & \ldots & \ldots & \ldots & 0 & A_{j-1}^* & A_j \end{pmatrix} \tag{7}$$

by $r_{i|i+1|\ldots|j-1|j}$.

Theorem 3

On the structure of the $\mathbf{Q^{k-1}}$-SVD

Consider a $\mathbf{Q^{k-1}-SVD}$ *of the matrices* $A_1, A_2, \ldots, A_k$. *Then*

- *If* $j-i$ *even*

$$r_{i|\ldots|j} = r_{i|\ldots|j-1} + (r_j^1 + r_j^2 + \ldots + r_j^i) + r_j^{i+2} + r_j^{i+4} + \ldots + r_j^{j-2} + r_j^j \tag{8}$$

- *If* $j-i$ *odd*

$$r_{i|\ldots|j} = r_{i|\ldots|j-1} + (r_j^{i+1} + r_j^{i+3} + \ldots + r_j^{j-2} + r_j^j) \tag{9}$$

For the general case, we shall need a mixture of the two preceding notations for block bidiagonal matrices, the blocks of which can be products of matrices, such as:

$$\begin{pmatrix} A_{i_0}A_{i_0+1}\ldots A_{i_1-1} & 0 & 0 & \ldots & 0 \\ (A_{i_1}\ldots A_{i_2-1})^* & A_{i_2}\ldots A_{i_3-1} & 0 & \ldots & 0 \\ 0 & (A_{i_3}\ldots A_{i_4-1})^* & A_{i_4}\ldots A_{i_5-1} & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & \ldots & \ldots & \ldots & A_{i_l}\ldots A_j \end{pmatrix}$$

where $1 \leq i_0 < i_1 < i_2 < i_3 < \ldots < i_l \leq j \leq k$. Their rank will be denoted by

$$r_{(i_0)\ldots(i_1-1)|i_1\ldots(i_2-1)|\ldots|i_l\ldots(j)}$$

For instance, the rank of the matrix

$$\begin{pmatrix} A_2A_3 & 0 & 0 \\ A_4^* & A_5A_6A_7 & 0 \\ 0 & (A_8A_9)^* & A_{10} \end{pmatrix}$$

will be represented by

$$r_{(2)(3)|4|(5)(6)(7)|(8)(9)|(10)}$$

Theorem 4

On the structure of a GSVD

The rank $r_{(i_0)(i_0+1)\ldots(i_1-1)|i_1\ldots(i_2-1)|\ldots|i_l\ldots j}$ can be expressed as follows:

1. *Calculate the following $l+1$ integers s_j^i, $i = 1, 2, \ldots, l+1$:*

$$\begin{aligned} s_j^1 &= r_j^1 + r_j^2 + \ldots + r_j^{i_0} \\ s_j^2 &= r_j^{i_0+1} + r_j^{i_0+2} + \ldots + r_j^{i_1} \\ \ldots &= \ldots \\ s_j^{l+1} &= r_j^{i_{l-1}+1} + r_j^{i_{l-1}+2} + \ldots + r_j^{i_l} \end{aligned}$$

2. *Depending on l even or odd there are two cases:*

l **even:**

$$\begin{aligned} & r_{i_0\ldots i_1-1|i_1\ldots i_2-1|\ldots|i_l\ldots j} \\ & = r_{i_0\ldots i_1-1|i_1\ldots i_2-1|\ldots|i_{l-1}\ldots i_l-1} + s_j^1 + s_j^3 + \ldots + s_j^{l+1} \end{aligned} \tag{10}$$

l **odd:**

$$\begin{aligned} & r_{i_0\ldots i_1-1|i_1\ldots i_2-1|\ldots|i_l\ldots j} \\ & = r_{i_0\ldots i_1-1|i_1\ldots i_2-1|\ldots|i_{l-1}\ldots i_l-1} + s_j^2 + s_j^4 + \ldots + s_j^{l+1} \end{aligned} \tag{11}$$

Observe that Theorems 2 and 3 are special cases of Theorem 4.

While Theorem 2 provides a direct expression of the dimensions r_j^i in terms of differences of ranks of products, Theorem 3 and 4 do so only implicitly. However, Theorem 3 and 4 can be used to set up a set of linear equations from which these block dimensions r_j^i

of the main Theorem can then be derived. Let us illustrate this with a couple of examples.

Example 4:

Let us determine the block dimensions of the quasi-diagonal matrix S_4 in a **QPP-SVD** of the matrices A_1, A_2, A_3, A_4 using Theorem 2, 3 and 4. From Theorem 2 we find:

$$\begin{aligned} r_4^4 &= r_4 - r_{34} \\ r_3^4 &= r_{34} - r_{234} \end{aligned}$$

From Theorem 4, we find:

$$\begin{aligned} s_4^1 &= r_4^1 \\ s_4^2 &= r_4^2 \end{aligned}$$

and

$$r_{(1)|(2)(3)(4)} = r_1 + s_4^2$$

so that

$$r_4^2 = r_{1|(2)(3)(4)} - r_1$$

Finally, since $r_4 = r_4^1 + r_4^2 + r_4^3 + r_4^4$, we find

$$r_4^1 = r_1 + r_{(2)(3)(4)} - r_{1|(2)(3)(4)}$$

Observe that this last relation can be interpreted geometrically as the dimension of the intersection between the row spaces of A_1 and $A_2A_3A_4$:

$$\begin{aligned} r_4^1 &= dim\ span_{row}(A_1) + dim\ span_{row}(A_2A_3A_4) \\ &\quad -dim\ span_{row}\begin{pmatrix} A_1 \\ A_2A_3A_3 \end{pmatrix} \end{aligned}$$

Example 5

Consider the determination of $r_5^1, r_5^2, r_5^3, r_5^4, r_5^5$ in a **PQQQ-SVD** of 5 matrices A_1, A_2, A_3, A_4, A_5 with Theorem 4:

r_5^i	s_5^i
$r_{4\|5}$	$s_5^1 = r_5^1 + r_5^2 + r_5^3 + r_5^4$ $s_5^2 = r_5^5$
$r_{3\|4\|5}$	$s_5^1 = r_5^1 + r_5^2 + r_5^3$ $s_5^2 = r_5^4$ $s_5^3 = r_5^5$
$r_{2\|3\|4\|5}$	$s_5^1 = r_5^1 + r_5^2$ $s_5^2 = r_5^3$ $s_5^3 = r_5^4$ $s_5^4 = r_5^5$
$r_{(1)(2)\|3\|4\|5}$	$s_5^1 = r_5^1$ $s_5^2 = r_5^2 + r_5^3$ $s_5^3 = r_5^4$ $s_5^4 = r_5^5$

These relations can be used to set up a set of equations for the unknowns $r_5^1, r_5^2, r_5^3, r_5^4, r_5^5$ as:

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} r_5^1 \\ r_5^2 \\ r_5^3 \\ r_5^4 \\ r_5^5 \end{pmatrix} = \begin{pmatrix} r_5 \\ r_{4|5} - r_4 \\ r_{3|4|5} - r_{3|4} \\ r_{2|3|4|5} - r_{2|3|4} \\ r_{(1)(2)|3|4|5} - r_{(1)(2)|3|4} \end{pmatrix}$$

the solution of which is

$$\begin{aligned} r_5^1 &= r_{3|4|5} - r_{3|4} + r_{(1)(2)|3|4} - r_{(1)(2)|3|4|5} \\ r_5^2 &= r_{(1)(2)|3|4|5} - r_{(1)(2)|3|4} - r_{2|3|4|5} + r_{2|3|4} \\ r_5^3 &= r_{2|3|4|5} - r_{2|3|4} - r_{4|5} + r_4 \\ r_5^4 &= r_5 - r_{3|4|5} + r_{3|4} \\ r_5^5 &= r_{4|5} - r_4 \end{aligned}$$

4. Applications

Most of the problems for which the **OSVD**, **PSVD**, **QSVD** etc ...provide an answer, can in principle be solved via a (generalized) eigenvalue problem. However, this always requires the *explicit* calculation of products or quotients of matrices, which can give raise to severe loss of numerical accuracy. Even if the eigenvalue algorithms would be numerically robust, it is in most cases the explicit formation of matrix products (which consists essentially of inner products) that causes loss of numerical accuracy.
As an example, consider the computation of the $\mathbf{P^3}$ – **SVD** of 4 matrices A_1, A_2, A_3, A_4 where

$$A_1 = \begin{pmatrix} 1 & \mu & 0 \\ 1 & 0 & \mu \end{pmatrix} \qquad A_2 = \begin{pmatrix} 1 & 1 \\ \mu & 0 \\ 0 & \mu \end{pmatrix}$$

$$A_3 = \begin{pmatrix} -1 & \mu & 0 \\ 1 & 0 & \mu \end{pmatrix} \qquad A_4 = \begin{pmatrix} -1 & 1 \\ \mu & 0 \\ 0 & \mu \end{pmatrix}$$

Assume that $\mu^2 < \epsilon_m < \mu$, where ϵ_m is the machine precision. Let $fl(.)$ represent the effect of performing a calculation on a finite precision machine so that $fl(1+\mu^2) = 1$. Then, it is easy to illustrate that matrix multiplication on a finite precision machine is not *associative*:

$$\begin{aligned} fl[\, fl(A_1A_2)\; fl(A_3A_4)\,] &= \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \\ fl[\, fl(\; fl(A_1A_2)\; A_3)A_4\,] &= \begin{pmatrix} \mu^2 & \mu^2 \\ \mu^2 & \mu^2 \end{pmatrix} \\ fl[\, fl(\; A_1\; fl(A_2A_3)\,)\; A_4\,] &= \begin{pmatrix} 2\mu^2 & 0 \\ 0 & 2\mu^2 \end{pmatrix} \end{aligned}$$

The first result has rank 0, the second result has rank 1 and the third result has rank 2! The correct result would be:

$$A_1A_2A_3A_4 = \begin{pmatrix} \mu^2(\mu^2+2) & 0 \\ 0 & \mu^2(\mu^2+2) \end{pmatrix}$$

and is of rank 2. Obviously, it is only a direct explicit factorization of every matrix separately that can preserve the fine numerical details that otherwise get irreversibly lost in matrix products.

For another example, suppose we want to compute the **QSVD** of a pair of matrices:

$$A_1 = \begin{pmatrix} 1 & 1 \\ -1 & 1 \\ \epsilon & \epsilon \end{pmatrix} \qquad A_2 = \begin{pmatrix} 1 & \mu & 0 \\ 1 & 0 & \mu \end{pmatrix}$$

where $\mu^2 < \epsilon_m < \mu$ and $\epsilon^2 < \epsilon_m < \epsilon$. The theoretically correct **QSVD** of this matrix pair is:

$$\begin{aligned} A_1 &= U_1 D_1 X_1^{-1} \\ &= \begin{pmatrix} \frac{1}{\sqrt{1+\epsilon^2}} & 0 & -\frac{\epsilon}{\sqrt{1+\epsilon^2}} \\ 0 & -1 & 0 \\ \frac{\epsilon}{\sqrt{1+\epsilon^2}} & 0 & \frac{1}{\sqrt{1+\epsilon^2}} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{2\sqrt{1+\epsilon^2}} & \frac{1}{2} \\ \frac{1}{2\sqrt{1+\epsilon^2}} & -\frac{1}{2} \end{pmatrix}^{-1} \\ A_2 &= X_1^{-t} S_2 V_2^t \\ &= \begin{pmatrix} \frac{1}{2\sqrt{1+\epsilon^2}} & \frac{1}{2} \\ \frac{1}{2\sqrt{1+\epsilon^2}} & -\frac{1}{2} \end{pmatrix}^{-t} \begin{pmatrix} \frac{\sqrt{1+2\mu^2}}{2\sqrt{1+\epsilon^2}} & 0 \\ 0 & \frac{\mu}{\sqrt{2}} \end{pmatrix} \end{aligned}$$

$$\times \begin{pmatrix} \frac{2}{\sqrt{4+2\mu^2}} & \frac{\mu}{\sqrt{4+2\mu^2}} & \frac{\mu}{\sqrt{4+2\mu^2}} \\ 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ -\frac{\mu}{\sqrt{\mu^2+2}} & \frac{1}{\sqrt{\mu^2+2}} & \frac{1}{\sqrt{\mu^2+2}} \end{pmatrix}$$

Now, in a lot of applications [6] [22] [26], one is interested in the extrema of the so-called *oriented signal-to-signal ratio* of two vector sequences in the direction of a vector x, which is defined as:

$$E_x[A_1^t, A_2] = (x^t A_1^t A_1 x)/(x^t A_2 A_2^t x) \quad (12)$$

It is easy to verify that the extremal values of this quotient for our example, are given by the inverses of the diagonal elements of $S_2 S_2^t$:

$$max(E_x[A_1^t, A_2]) = \frac{4(1+\epsilon^2)}{(4+2\mu^2)}$$

$$min(E_x[A_1^t, A_2]) = \frac{2}{\mu^2}$$

If the vector sequence in the matrix A_1^t is considered to be signal + noise, and the one in A_2 contains the noise (disturbances) then it can be verified that the 'signal energy' [6] in the direction $x = [1 \; -1]^t$ is 1 while the noise energy is $\mu/\sqrt{2}$. On the other hand, if we would first calculate explicitly the matrix products $A_1^t A_1$ and $A_2 A_2^t$ and optimize (12) as a generalized eigenvalue problem of the matrix pair $(A_1^t A_1, A_2 A_2^t)$, then, the extremal values of

$$x^t(fl[A_1^t A_1])x/(x^t(fl[A_2 A_2^t])x)$$

are $(2+\epsilon^2)/2$ and 0! In this case, the signal energy in the direction $x = [1 \; -1]^t$ is 1 while the noise energy is 0. This would lead to the conclusion that this direction is noiseless while in fact, it is not!

The **OSVD** is so frequently used in signal processing and systems and control theory that we shall not attempt here to give a complete survey of all its applications. The interested reader may wish to consult [13] [17] in order to get a survey of applications and algorithms. A system identification application is treated in [21]. It is the *dynamic* counterpart of solving overdetermined sets of linear equations via *total linear least squares* using the **OSVD** [17].
The use of the **QSVD** is advocated in signal processing applications where strong 'desired' signals have to be separated from weak 'disturbing' signals. Typically, the frequency domain spectra are overlapping which complicates the use of frequency domain filtering techniques. The concept behind this separation technique is the *oriented signal-to-signal ratio* which coincides with the concept of *prewhitening* if noise covariance matrices are known [6]. Typical applications can be found in [4] [22] [26] [30]. In [22], a **QSVD** based system identification algorithm is explored, which gives unbiased results as compared to the **OSVD** version, when data are first treated prior to identification with some filter, as often happens in practice.

Applications of the **PSVD** are mentioned in [16] [18], including the computation of the Kalman decomposition of a linear system. Typically, the **PSVD** can be invoked whenever so-called *contragredient* transformations are involved as is the case in *open* (observability and controllability Lyapunov equations) and *closed loop balancing* (via the filter and control algebraic Riccati equation).
Applications of the **RSVD** (=**QQ-SVD**) are enumerated in [8]. A typical problem concerns the minimization of the rank of the matrix $A + BDC$ where A, B, C are given matrices, over all possible matrices D, such that a unitarily invariant norm of D is minimal. The answer is given in terms of the **QQ-SVD** of the matrix triplet (B, A, C). Relationships with the *shorted operator, generalized Schur complemenents, generalized Gauss-Markov estimation problems* and a generalization of *total linear least squares* are also pointed out in [8] (see also [31]). It's interesting to note that our **QQ-SVD** can be used to calculate the minimal rank matrix in a matrix ball, which is the solution set of a completion problem [5].
In [34], it is shown how the **PP-SVD** can increase the numerical robustness of the solution of matrix approximation problems of the form

$$\min_{rank(X) = r} \quad \|A(B - X)C\|_F^2$$

where A, B, C are given rectangular and possibly rank deficient matrices and X is to be found. The closeness of the approximation is measured by the semi-matrix norm with row weighting matrix A and column weighting matrix C. In [34] not only consistency conditions are derived for the problem but it is also shown how a subspace can be found using the **PP-SVD** so that the semi-norm becomes a matrix norm.
Finally, let us conclude this section by pointing out the connection between **GSVDs** and the *stochastic realization problem*, which is the following:
Given output signals $y_k, k = 1, \ldots, N$ or covariance sequences $\Lambda_i = \mathbf{E}(y_k y_{k+i}^t)$ generated by a stationary stochastic differential system of equations of the form:

$$\begin{aligned} x_{k+1} &= Ax_k + w_k \\ y_k &= Cx_k + v_k \end{aligned}$$

where $A \in \Re^{n \times n}, C \in \Re^{l \times n}$ and w_k, v_k are (possibly mutually correlated) white noise sequences. Find n, A, C and the covariance matrices $Q = \mathbf{E}(w_k w_k^t), S = \mathbf{E}(w_k v_k^t), R = \mathbf{E}(v_k v_k^t)$ from the data y_k (*data driven*) or the output covariances Λ_k (*covariance driven*).

Algorithms to solve this realization problem are typically based on the observation that $rank[\mathbf{E}(Y_+ Y_-^t)] = n$, where Y_+ and Y_- are $li \times j$ $(j >> li)$ block Hankel matrices with future and past output signals. In [25], a classification was made of all existing stochastic realization schemes via the so-called RV-coefficient optimization problem. This problem consists essentially of finding linear combinations of the future as LY_+ and of the past MY_- and then maximizing the correlation between these two vector sequences subject to some constraints on L and M. In [27], we will describe in detail how all these optimization problems can be solved in terms of certain **GSVDs** while here we will only reproduce a summary. Mathematically, the problem can be formulated as a constrained optimization problem:

$max_{L,M} \|L^t Y_+ Y_-^t M\|_F^2$

Data driven	Constraints	Algorithm
- Symmetric - Multivariate Association and Similarity	$L^t Y_+ Y_+^t L = \Delta_z$ $M^t Y_- Y_-^t M = \Delta_x$	QPQ-SVD of Y_+^t, Y_+, Y_-^t, Y_-
	$L^t L = I$ $M^t M = I$	P-SVD of Y_+, Y_-^t
- Asymetric - Predictability	Backward $L^t Y_+ Y_+^t L = \Delta_z$ $M^t M = I$	QP-SVD of Y_+^t, Y_+, Y_-^t
	Forward $L^t L = I$ $M^t Y_- Y_-^t M = \Delta_x$	PQ-SVD Y_+, Y_-^t, Y_-

In some applications like astronomy, often only covariance information is available. Define here the covariance matrices $R_- = \mathbf{E}(Y_- Y_-^t)$, $R_+ = \mathbf{E}(Y_+ Y_+^t)$, $H = \mathbf{E}(Y_+ Y_-^t)$. Then the optimization problem becomes

$max_{L,M} \|L^t H M\|_F$

subject to

Covariance driven	Constraints	Algorithm
Symmetric	$L^t R_+ L = \Delta_z$ $M^t R_- M = \Delta_x$	QQ-SVD of R_+, H, R^-
	$L^t L = I$ $M^t M = I$	O-SVD of H
Asymmetric	Backward $L^t R_+ L = \Delta_z$ $M^t M = I$	Q-SVD of R_+, H
	Forward $L^t L = I$ $M^t R_- M = \Delta_x$	Q-SVD of R_-, H

5. Conclusions

In this paper, we have given a state-of-the-art survey of recently discovered *generalizations* of the *ordinary singular value decomposition.* While we have already revealed much of the structure and interesting properties of this infinite tree of GSVDs, much work remains to be done. We are convinced that this tree of generalized singular value decompositions will open new and exciting fields of research with respect to numerical algorithms, geometrical interpretations and applications. For instance, the uniqueness issues, the relation to generalized eigenvalue problems, genericity and sensitivity analysis and numerical and implementational aspects definitely deserve more attention as well as a continuous concern to demonstrate the usefulness in applications of practical interest. We have implement a straightforward algorithmic procedure in *MATLAB*, which allows to calculate any **GSVD** of any number of matrices of compatible dimensions. It basically follows the constructive proof of [11]. On the other hand, recently developed *updating techniques* for the **OSVD**, **PSVD** and **QSVD** [23] might prove useful when implementing algorithms for **GSVDs**.

References

[1] Autonne L. *Sur les groupes linéaires, réels et orthogonaux.* Bull. Soc. Math., France, vol.30, pp.121-133, 1902.

[2] Beltrami E. *Sulle Funzione Bilineari.* Giornale di Mathematiche, Battagline G., Fergola E. (Editors), Vol.11, pp.98-106, 1873.

[3] Boyd S., Balakrishnan V. *A Regularity Result for the Singular Values of a Transfer Matrix and a Quadratically Convergent Algorithm for Computing its L_∞-norm.* Technical Report No. L-106-89-1, Information Systems Lab, Stanford University, March 1989.

[4] Callaerts D., De Moor B., Vandewalle J., Sansen W. *Comparison of SVD-methods to extract the fetal electrocardiogram from cutaneous electrode signals.* Medical and Biological Engineering and Computing, 28, 1990, pp.217-224.

[5] Davis C., Kahan W.M., Weinberger H.F. *Norm-preserving dilations and their application to optimal error bounds.* SIAM J. Numer. Anal. 19, pp.445-469, 1982.

[6] De Moor B., Vandewalle J. *Oriented energy and oriented signal to signal ratios : The (Generalized) Singular Value Decomposition* in 'SVD and Signal Processing: Algorithms, Applications and Architectures.' Edited : E.F.Deprettere, North Holland, Elseviers Science Publishers B.V., 1988, pp. 209-232.

[7] De Moor B., Golub G. *Generalized singular value decompositions: A proposal for a standardized nomenclature.* ESAT-SISTA report 1989-10, 15 pp., April 1989, Department of Electrical Engineering, Katholieke Universiteit Leuven, Belgium. (Also Numerical Analysis Project Manuscript NA-89-04, Department of Computer Science, Stanford University.)

[8] De Moor B., Golub G.H. *The restricted singular value decomposition: properties and applications.* ESAT-SISTA report 1989-09, 55 pp., April 1989, Department of Electrical Engineering, Katholieke Universiteit Leuven, Belgium. (Also Numerical Analysis Project Manuscript NA-89-03, Department of Computer Science, Stanford University.) (Accepted for publication in the Siam Journal on Matrix Analysis and Applications.)

[9] De Moor B. *On the structure and geometry of the product singular value decomposition.* ESAT-SISTA report 1989-12, 60 pp., May 1989, Department of Electrical Engineering, Katholieke Universiteit Leuven, Belgium. (Also Numerical Analysis Project Manuscript NA-89-05, Department of Computer Science, Stanford University.) (submitted for publication)

[10] De Moor B., Boyd S. *Analytic properties of singular values and vectors.* ESAT-SISTA Report 1989-28, 17 pp., December 1989, Department of Electrical Engineering, Katholieke Universiteit Leuven, Belgium.

[11] De Moor B., Zha H. *A tree of generalizations of the singular value decomposition.* ESAT-SISTA Report 1990-11 (Revised version of ESAT-SISTA Report 1989-21), Department of Electrical Engineering, Katholieke Universiteit Leuven, Belgium, 24 pp., June 1990 (Accepted for publication in the special issue on " Matrix Canonical Forms" of "Linear Algebra and its Applications".)

[12] De Moor B. *On the structure of generalized singular value decompositions.* ESAT-SISTA report 1990-12, June 1990, Department of Electrical Engineering, Katholieke Universiteit Leuven, Belgium.

[13] Deprettere E. (editor) *SVD and Signal Processing: Algorithms, Applications and Architectures.* North Holland, 1988.

[14] Fan K., Tits A. *Toward a Structured Singular Value Decomposition.*, Report Electrical Engineering Department and System Research Center, University of Maryland, College Park, MD 20742, USA, pp.1-5., 1988.

[15] Fernando K.V., Hammarling S.J. *A product Induced Singular Value Decomposition for two matrices and balanced realisation.* NAG Technical Report, TR8/87.

[16] Fernando K.V. *The Kalman Reachability-Observability Canonical Form and the* Π*SVD.* NAG Technical Report TR9/87, July 1987.

[17] Golub G.H., Van Loan C. *Matrix Computations.* Johns Hopkins University Press, Baltimore 1983.

[18] Heath M.T., Laub A.J., Paige C.C., Ward R.C. *Computing the singular value decomposition of a product of two matrices.* SIAM J. Sci. Stat. Comput., Vol.7, pp.1147-1159, 1986.

[19] Jordan C.J. Math. pures appl.II, Vol.19, pp.35-54, 1874.

[20] MacDuffee C.C. *The theory of matrices.* Springer Berlin, 1933, also, Chelsea Publishing Company, New York (corrected reprint of First Edition).

[21] Moonen M., De Moor B., Vandenberghe L., Vandewalle J. *On- and off-line identification of linear state space models.* International Journal of Control, Vol.49, No.1, pp.219-239.

[22] Moonen M., Vandewalle J. QSVD *approach to on- and off-line state space identification.* Int.J.Control 1990, Vol. 51, No. 5, pp. 1133-1146

[23] Moonen M. *Jacobi-type updating algorithms for signal processing, system identification and control.* Doct. Dissertation, Department of Electrical Engineering, Katholieke Universiteit Leuven, Belgium, 1990.

[24] Paige C.C., Saunders M.A. *Towards a generalized singular value decomposition.* SIAM J. Numer. Anal., 18, pp.398-405, 1981.

[25] Ramos J.A., Verriest E.I., Rao S.G. *A unified RV-coefficient approach for solving the covaraicne based stochastic realization problem.* School of Electrical Engineering, Georgia Institute of Technology, Atlanta, Georgia, 1987.

[26] Roy R. **ESPRIT** - *Estimation of signal parameters via rotational invariance techniques,* Ph.D. Thesis, Department of Electrical Engineering, Stanford, Ca, USA, 1987.

[27] Suykens J., De Moor B., Vandewalle J. *Generalized singular value decompositions and the stochastic realization problem.* ESAT-SISTA Report, Department of Electrical Engineering, Katholieke Universiteit Leuven, 1990 (in preparation).

[28] Sylvester J.J. *Sur la réduction biorthogonale d'une forme linéo-lin'eaire à sa forme canonique.* Comptes Rendus, CVIII., pp.651-653, 1889 (see also Sylvester J.J. *A new proof that a general quadric may be reduced to its canonical form (that is, a linear function of squares) by means of a real orthogonal substitution.* Messenger of Mathematics, XIX, pp.1-5, 1890 and Sylvester J.J. *On the reduction of a bilinear quantic of the n^{TH} order to the form of a sum of n products by a double orthogonal substitution.* Messenger of Mathematics, XIX, pp.42-46, 1890.)

[29] Takagi T. *On an algebraic problem related to an analytic theorem of Caratheodory and Fejer and on an allied theorem of Landau.* Japan. J. Math., 1, pp.83-93, 1925.

[30] Van Huffel S., Vandewalle J. *Analysis and properties of the Generalized total least squares problem $AX \approx B$ when some or all columns of A are subject to errors.* SIAM J. Matrix Anal. Appl. 10 (1989), pp. 294-315.

[31] S. Van Huffel, Zha H. *Restricted Total Least Squares: A unified approach for solving (generalized) (total) least squares problems with(out) equality constraints.* ESAT-SISTA report 1989-05, 20 pp., March 1989, Department of Electrical Engineering, Katholieke Universiteit Leuven, Belgium. (also Numerical Analysis Project Manuscript NA-90-02, Computer Science Department, Stanford University, Stanford, CA 94305, U.S.A., February 1990)(accepted for publication in SIAM J. Matrix Anal. Appl.).

[32] Van Loan C.F. *Generalizing the singular value decomposition.* SIAM J. Numer.Anal., 13, pp.76-83, 1976.

[33] Zha H. *Restricted SVD for matrix triplets and rank determination of matrices.* Scientific Report 89-2, ZIB, Berlin (submitted).

[34] Zha H. *Analysis and Numerical Solution of Fixed Rank Matrix Approximation with General Weighting Matrices.* ESAT-SISTA Report 1989-22, 28 pp., October 1989, Department of Electrical Engineering, Katholieke Universiteit Leuven, Belgium.