Singular value decomposition : A reliable tool in the algorithmic analysis of linear systems

Jan Staar, Joos Vandewalle *

SUMMARY

The Singular Value Decomposition is an eigenvalue-like concept with remarkable geometric and numerical qualities. It is becoming more and more important in control theory. We review some important facts and properties with respect to the SVD interpretation and manipulation of matrices, seen as linear operators, or seen as a collection of samples of a vector process.

1. INTRODUCTION

The basic purpose of this article is to familiarize the reader with Singular Value Decomposition (SVD), which is an eigenvalue-like concept, but with remarkable geometric and numerical qualities. Although the idea of Singular Value Decomposition dates back to the previous century, it was not until recently [2-4] that the numerical and algorithmic problems related to the computation of an SVD were described, and finally solved. This mathematical achievement [2], together with the availability of a standard FORTRAN subroutine [4] that performs the SVD of an arbitrary matrix in an extremely efficient and numerically stable way, has been the impetus of much interest in the domain of applied sciences. Today, SVD is more and more being considered as a basic tool in system and control engineering [1], and has proved to be useful in both conceptual and numerical aspects of problem analysis [5-10].

2. THE SINGULAR VALUE DECOMPOSITION THEOREM

Although there is no formal difficulty to extend the SVD statement to complex valued matrices, we will restrict ourselves to the case of MxN matrices over the field of real numbers.

Theorem 1

(The Autonne-Eckart-Young theorem, restricted to real matrices) :
For any real matrix A, there exists a factorization **

$$A_{MN} = U_{MM}\, \Sigma_{MN}\, V^t_{NN} \qquad (1)$$

in which the factors U and V^t are real, orthonormal matrices ($UU^t = I$, $VV^T = I$), and the factor Σ is a real, pseudodiagonal non-negative definite matrix. Proofs of this classical theorem can be found in [5] (numerical context), [6] (control application context). The diagonal elements of Σ are called the *singular values* σ_i of matrix A. Their set is called the *singular spectrum* $\sigma(A)$ of matrix A. Usually, the singular values are assumed to be ordered in decreasing magnitude.

The columns u_i of U and the rows v_i^t of V^t are called the left resp., right *singular vectors* of A and their set its *left resp., right singular basis*. Observe that Theorem 1 then guarantees the existence of the sum $A = \sum_{i=1}^{\text{rank } A} \sigma_i\, u_i\, v_i^t$, which is usually called the dyadic decomposition.

Example
The 2x2 matrix A has an SVD given by

$$\underset{A}{\begin{bmatrix} 0.577699 & -0.224144 \\ 1.190069 & 0.836516 \end{bmatrix}} = \underset{U}{\begin{bmatrix} \cos 75 & -\sin 75^\circ \\ \sin 75^\circ & \cos 75^\circ \end{bmatrix}} \cdot \underset{\Sigma}{\begin{bmatrix} 3/2 & 0 \\ 0 & 1/2 \end{bmatrix}} \cdot \underset{V^t}{\begin{bmatrix} \cos 30^\circ & \sin 30^\circ \\ -\sin 30^\circ & \cos 30^\circ \end{bmatrix}} \qquad (2)$$

The matrices U and V^t are clearly orthogonal and the singular values are 3/2 and 1/2. The entries of U, Σ, and V in practice are best computed using the subroutine SVD of [4]. More examples are given in [6]. The singular structure of A is very closely related to the eigenstructures of AA^t and A^tA. Indeed, from (1) we have :

$$AA^t = U\, \Sigma\, \Sigma^t\, U^t \qquad (3)$$

$$A^tA = V\, \Sigma^t\, \Sigma\, V^t \qquad (4)$$

Hence the left (resp., right) singular vectors are the eigen-

** The subscripts MN in A_{MN} indicate the size MxN of the matrix A and are often omitted. V^t denotes the transpose of V.

* Departement Elektrotechniek – ESAT Laboratory, Katholieke Universiteit Leuven, Kardinaal Mercierlaan 94, B-3030 – Heverlee, Belgium.

doctor's thesis,Université de Rennes, France.

8. A.Shamir,"The strongest knapsack-based cryptosystem," presented at CRYPTO'82, Santa Barbara, California, U.S.A., August, 1982.

9. M.Davio,"Knapsack trapdoor functions : an introduction", presented at CISM Summer School on : Secure Digital Communications, Udine, Italy, June 7-11, 1982

10. E.F.Brickell,J.A.Davis and G.J.Simmons, "A Preliminary Report on the Cryptanalysis of the Merkle-Hellman Knapsack Cryptosystems", Crypto'82, Santa Barbara, CA, August 1982, appeared in : Advances in Cryptography, 1982.

11. A.J.Mc.Auley and R.M.Goodman,"Modifications to the Trapdoor-Knapsack Public Key Cryptosystem," IEEE Intern. Symp. Inform. Theory, St. Jovite, Québec, Canada, September 26-30, 1983, Abstract of papers, pp. 130.

12. B.Arazi,"A Trapdoor Multiple Mapping," IEEE Trans. Inform. Theory, vol. 26, no. 1, pp. 100-102, January 1980.

13. E.F.Brickell,"A New Knapsack Based Cryptosystem," presented at CRYPTO'83, Santa Barbara, California, U.S.A. August 21-24,1983.

14. L.M.Adleman,"On breaking generalized knapsack public key cryptosystems," CRYPTO'82, Santa Barbara, California, U.S.A. August, 1982.

15. G.Forsythe and C.Moler,"Computer solution of linear algebraic systems," Prentice Hall Inc., New Jersey, 1967.

16. A.Shamir,"On the Cryptocomplexity of Knapsack Systems," Proc. Stoc 11 ACM pp. 118-129, 1979.

17. E.F.Brickell,"Solving low density knapsacks," proceedings of CRYPTO'83, Santa Barbara, California, U.S.A. August 21-24, 1983.

18. M.R.Garey and D.S.Johnson,"Computers and Intactability : A Guide to the Theory of NP-Completeness," W.H.Freeman and Company, San Francisco, 1979.

19. G.Brassard,"A Note on the Complexity of Cryptography," IEEE Trans. Inform. Theory, vol. 25, no. 2, pp. 232-233, March, 1979.

20. R.L.Rivest, A.Shamir and L.Adleman, "A method for obtaining digital signatures and public key cryptosystems," Commun. ACM, vol. 21, pp. 294-299, April, 1978

21. H.W.Lenstra,"Integer Programming with a Fixed Number of Variables", University of Amsterdam, Dept. of Mathematics, Technical Report, 81-03, April, 1981.

22. L.M.Adleman,"On breaking iterated knapsack systems," CRYPTO'82, Santa Barbara, California, U.S.A. August 21-24, 1983.

23. A.K.Lenstra, H.W.Lenstra,Jr., and L.Lovasz,"Factoring Polynomials with Rational Coefficients," Mathematische Annalen 261, pp. 515-534, 1982.

24. J.W.S.Cassels,"An Introduction to the Geometry of Numbers," Springer-Verlag, Berlin, New-York, 1971.

25. R.Kannan and A.Bachem,"Polynomial Algorithms for Computing the Smith and Hermite Normal Forms of an Integer Matrix," SIAM J.Comput., vol. 8, no. 4, November, 1979, pp. 499-507.

26. Y.Desmedt, J.Vandewalle and R.Govaerts, "Linear Algebra and Extended Mappings Generalise Public Key Cruptographic Knapsack Algorithms," Electronics Letters, 12 May, 1983, vol. 19, no. 10, pp. 379-381.

27. Y.Desmedt, J.Vandewalle and R.Govaerts, "Fast Authentication Using Public Key Schemes," 1984 International Zurich Seminar on Digital Communications, March 6-8, 1984, to appear in the proceedings.

28. Y.Desmedt, J.Vandewalle and R.Govaerts, "The General Public Key Cryptographic Knapsack Algorithm," submitted for publication to the IEEE Trans. Inform. Theory.

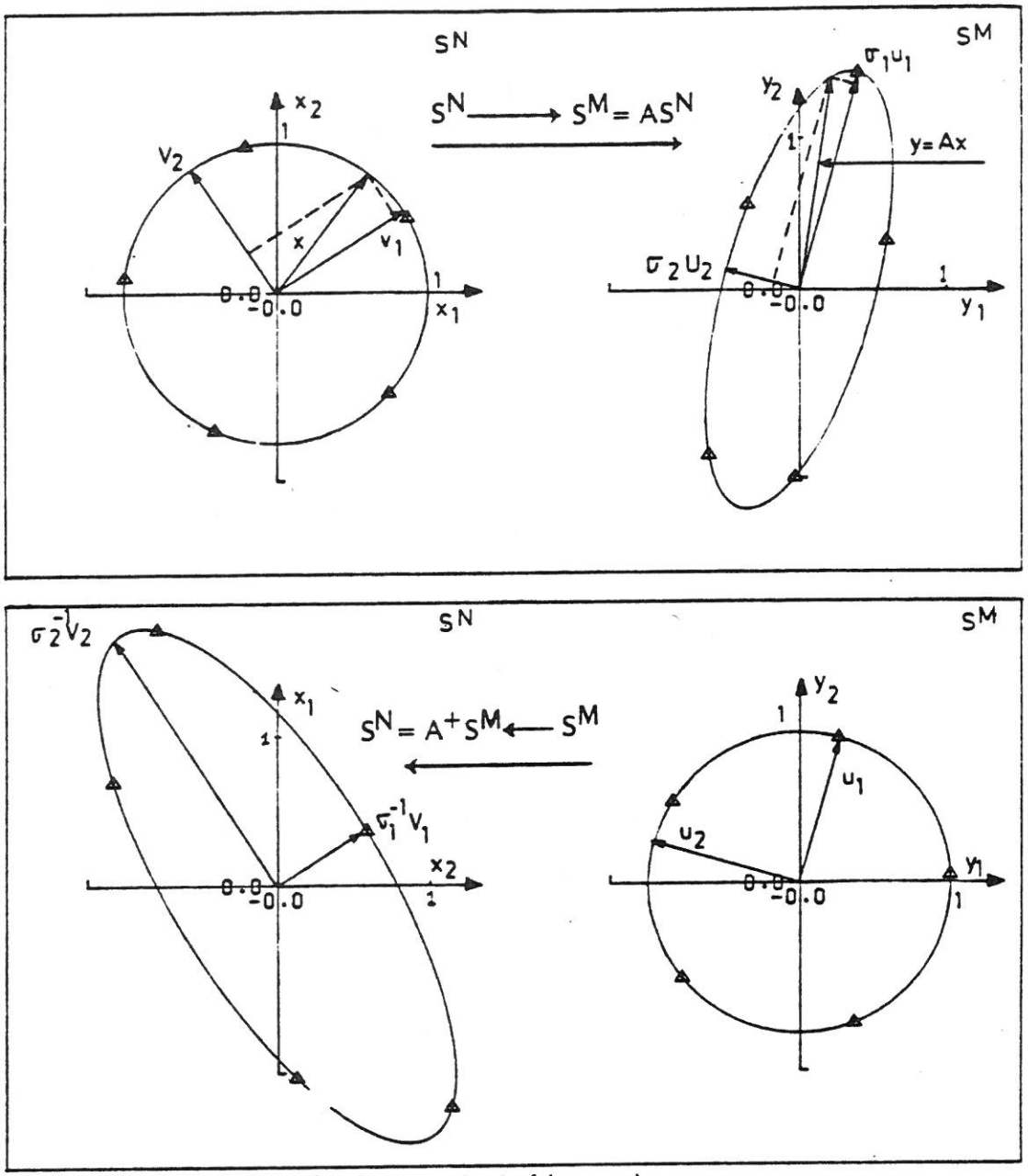


Fig. 1. Geometric interpretation of the linear mapping A of the example

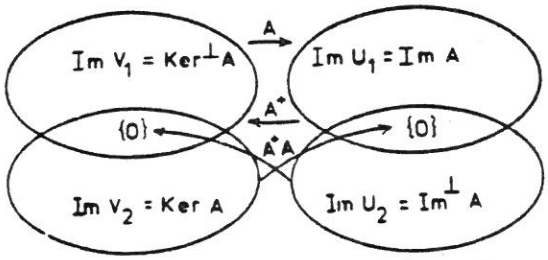


Fig. 2. The four fundamental subspaces of a matrix

the orthogonal complement of the kernel of A into the range of A. This operator is one-to-one and invertible :

$$Ax = y : U_1 \Sigma_1 V_1^t x = y \tag{9a}$$

$$x = A^+ y : x = V_1 \Sigma_1^{-1} U_1^t y \tag{9b}$$

which defines the *generalized inverse* in a very simple and natural way. Notice the surprising property of interchangeability of summation and generalized inversion for the dyadic decomposition :

$$A^+ = [\sum_{i=1}^{P} \sigma_i u_i v_i^t]^+ = \sum_{i=1}^{P} [\sigma_i u_i v_i^t]^+$$

Moreover with $\|x\|_2 = (\sum_{i=1}^{N} x_i^2)^{1/2}$ and Max and Min the maximum and minimum over appropriate sets we have :

$$\|A\|_2 \overset{\Delta}{=} \max_x \frac{\|Ax\|_2}{\|x\|_2} = \max_y \frac{\|y\|_2}{\|A^+ y\|_2} = \sigma_1 \tag{10.a}$$

$$\|A^+\|_2 \overset{\Delta}{=} \max_y \frac{\|A^+ y\|_2}{\|y\|_2} = \max_x \frac{\|x\|_2}{\|Ax\|_2} = \sigma_p^{-1} \tag{10.b}$$

$$K(A) \overset{\Delta}{=} \max_{\Delta y, y} \frac{(\|\Delta x\|_2 / \|x\|_2)}{(\|\Delta y\|_2 / \|y\|_2)} = \frac{\sigma_1}{\sigma_p} = \|A\|_2 \, \|A^+\|_2 \tag{10.c}$$

K(A) is called the *condition number of matrix A* in $\|\bullet\|_2$-norm. It is the most severe increase in relative inaccuracy when solving x from Ax = y for the worst position of y (i.e. $y // u_1$) and the unluckiest position of the error Δy (i.e. $\Delta y // u_p$). Clearly, if one solves a concrete linear problem Ax = y for a specific righthand side y, one can define a milder condition number. The *condition number $K_y(A)$ of the set of linear equations* Ax = y is defined as

$$K_y(A) \triangleq \underset{\Delta y}{\text{Max}} \frac{(\|\Delta x\|_2/\|x\|_2)}{(\|\Delta y\|_2/\|y\|_2)} \qquad (11.a)$$

for which it is easily verified that

$$K_y(A) = K(A) \,.\, \|y\|_2/(\|x\|_2\, \sigma_1) \leqslant K(A) \qquad (11.b)$$

3.3. Rank of a matrix

Algebraically, the rank of a matrix is P, i.e. the number of non-zero singular values (counting multiplicities). This rank is the dimension of both spaces $(\text{Ker } A)^\perp$ and Im A, related in a one-to-one manner by the operators A and A^+. The complements Ker A and $(\text{Im } A)^\perp$ of these spaces do not participate in a regular way in the maps A and A^+.

In *numerical* practice however the matrix and/or the righthandside are not exactly known due to inaccurate representation (machine precision), measurement inaccuracies, or rounding in the previous computations. ($\epsilon_y \triangleq \|\Delta y\|_2/\|y\|_2$ and $\epsilon_A \triangleq \|\Delta A\|_2/\|A\|_2$). Then generically all singular values will be non-zero, but a practical computation of the *numerical* rank of this matrix should not take into account those nonzero singular values, which are due to the inaccuracies.

In the practical context of the solution of a linear set of equations, one then defines the rank P_ϵ of the associated matrix to be the number of singular values (counting multiplicities), larger than $R_0\, \epsilon_y\, \sigma_1$, or $R_0\, \epsilon_A\, \sigma_1$ or both. The rank-P_ϵ approximants

$A_\epsilon = \sum_{i=1}^{P_\epsilon} \sigma_i\, u_i\, v_i^t$ and $A_\epsilon^+ = \sum_{i=1}^{P_\epsilon} \sigma_i^{-1}\, v_i\, u_i^t$ establish a one-to-one relationship between $(\text{Ker } A_\epsilon)^\perp$ and $\text{Im } A_\epsilon$ with a guaranteed relative accuracy of at least $\epsilon_y . \sigma_1/\sigma_{P_\epsilon} < R_0^{-1}$. The complements $\text{Ker } A_\epsilon$ and $(\text{Im } A_\epsilon)^\perp$ are considered not to participate in the linear maps A and A^+. More rank properties are described in the next section.

3.4. Other operator-related properties

Other reliable computations of geometric concepts such as intersections, linear sums and angles can be obtained from SVD but are beyond the scope of this review article [6].

4. SVD-BASED ENERGY PROPERTIES OF A SEQUENCE OF VECTORS

Any set of M-vectors $\{a_i\}$ e.g. a finite number of samples of a vector signal, can be put together as the columns in an MxN matrix

$$A = [a_1\ a_2 \ldots a_N] \qquad (12)$$

The sum of the squares of the projections $e^t a_i$ of these vectors on any unit direction e in S^M is defined as the *energy* of these vectors, *measured in direction* e [9] :

$$E_e \triangleq \sum_{i=1}^{N} (e^t a_i)^2 \qquad (13)$$

The SVD of A now provides very interesting information on the size of E_e as a function of the direction e.

The reader can easily generate dual properties for vectors put together as rows in a matrix.

Theorem 3

Consider any set of M-vectors which form the columns of the matrix A of (12) with singular value decomposition $A = U \Sigma V^t$, then we have

$$1)\ \sum_{i=1}^{N} \|a_i\|_2^2 = \sum_{k=1}^{P} \sigma_k^2 \qquad (14.a)$$

$$2)\ E_{u_k} = \sum_{i=1}^{N} (u_k^t\, a_i)^2 = \sigma_k^2,\ k = 1,2,\ldots M \qquad (14.b)$$

$$3)\ E_e = \sum_{i=1}^{N} (e^t\, a_i)^2 = \sum_{k=1}^{M} \mu_k^2\, \sigma_k^2, \text{ where } e = \sum_{k=1}^{M} \mu_k\, u_k, \qquad (14.c)$$

$$4)\ \underset{S^k \subseteq S^M}{\text{Max}} \sum_{i=1}^{N} \|\text{Proj}_{S^k} a_i\|_2^2 = \sum_{j=1}^{k} \sigma_j^2,\ k = 1,\ldots M \qquad (14.d)$$

$$5)\ \underset{S^k \subseteq S^M}{\text{Min}} \sum_{i=1}^{N} \|\text{Proj}_{S^k} a_i\|_2^2 = \sum_{j=M-k+1}^{M} \sigma_j^2,\ k = 1,\ldots M \qquad (14.e)$$

where P = Min (M,N), Max (resp., Min) denotes the maximum (resp., minimum) over all subspaces S^k of dimension k of the ambient space S^M, Proj_{S^k} denotes the orthogonal projection into S^k, and where $\sigma_k \triangleq 0$ when k > P.

Proof : See [7] [9].

In colloquial terms (14.a) says that the energy in the vector sequence is equal to the energy in the singular spectrum. According to (14.b) this energy is then decomposed over M orthogonal left singular directions u_i in S^M. The equation $E_e = \sum_{i=1}^{M} \mu_i^2\, \sigma_i^2$ shows that the energy along an arbitrary vector $e = \sum_{i=1}^{M} \mu_i\, u_i$ is a quadratic average of the energies σ_i^2 along the singular directions. In other words E_e is the quadratic sum of the projections of the vectors $\sigma_i\, u_i$ on the direction e. For the example (2) we have plotted the energy E_e

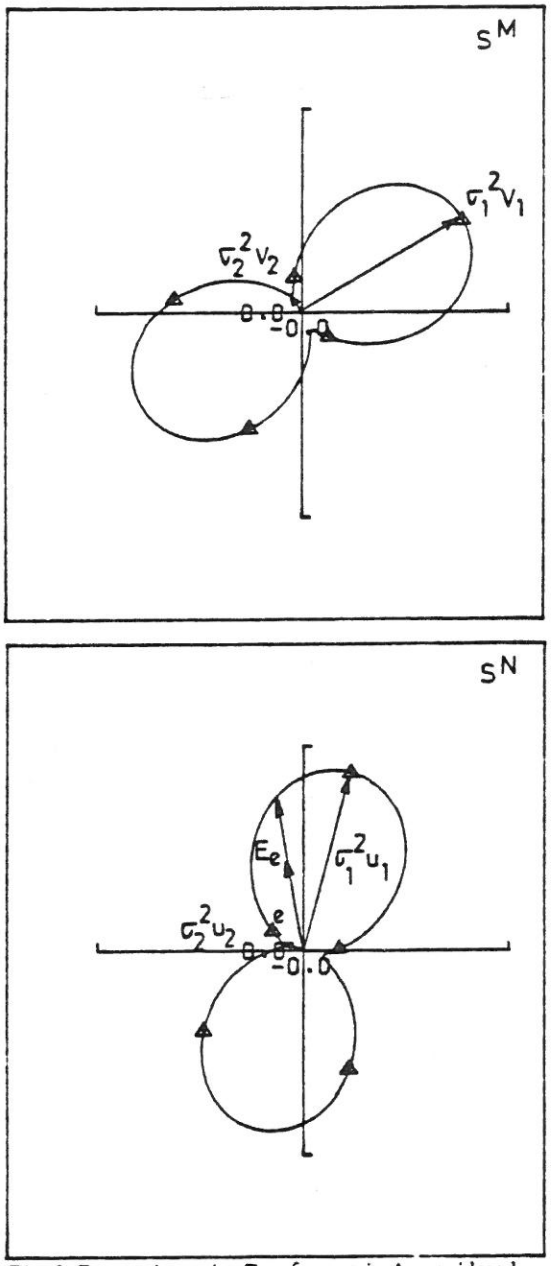


Fig. 3. Energy intensity E_e of a matrix A considered as set of row vectors (a) or column vectors (b)

along unit vectors e in Fig. 3(b). Extrema occur at the singular vectors u_i. This energy plot clearly has extrema at the singular vectors and shows a typical narrowing along the smallest singular vector even for mild variations $\sigma_1^2/\sigma_2^2 = 9$. From an algorithmic point of view this implies that minor inaccuracies on u_2 correspond to sizable variations in energy and thus that accurate computations of the energy require very precise computations of the singular vector. This is precisely what the SVD algorithm [4] does. The equations (14. d,e) finally show that the orthogonal decomposition of energies is canonical in the sense that it allows to find subspaces of dimension i where the projected sequence has maximal and minimal energy. A decomposition of the total space as a direct sum of a subspace of maximal, and a subspace of minimal energy for the given vector sequence, leads to very useful *rank considerations*.

Theorem 4

From the dyadic decomposition $A_{MN} = \sum_{i=1}^{P} \sigma_i u_i v_i^t$, it follows that

1) $\forall A^*_{MN}$: rank $A^*_{MN} = P_\epsilon < P$

$$\rightarrow \|A^*_{MN} - A_{MN}\|_2^2 \geq \|\sum_{i=1}^{P_\epsilon} \sigma_i u_i v_i^t - A_{MN}\|_2^2$$

$$= \sigma^2_{P_\epsilon+1}$$

2) $\forall A^*_{MN}$: rank $A^*_{MN} = P_\epsilon < P \rightarrow \|A^*_{MN} - A_{MN}\|_F^2$

$$\geq \|\sum_{i=1}^{P_\epsilon} \sigma_i u_i v_i^t - A_{MN}\|_F^2 = \sum_{j=P_\epsilon+1}^{P} \sigma_j^2$$

Theorem 4 gives a direct measure, in Frobenius norm $\|\bullet\|_F$, for the distance between the given matrix A_{MN}, and its closest approximant

$\sum_{j=1}^{P_\epsilon} \sigma_i u_i v_i^t$ of lower rank P_ϵ.

5. APPLICATIONS

The singular value decomposition has already been used in the sixties in numerical linear algebra and statistics basically for solving equations (least squares and generalized inverse) and computing the condition, the rank and the fundamental spaces of a matrix. The availability in the software package LINPACK has recently promoted its use in linear system theory [1, 8, 9] especially for reliable software for systems. It has also recently turned out to be useful in comparing the numerical qualities of certain representations of electrical circuits [10] and in computing principal axes and moments of inertia of mechanical systems [7]. The singular values of the transfer matrix of a system with multiple inputs and outputs have recently been shown to be of critical importance in the design of robust controllers [1, 5 papers].

REFERENCES

1. Special issue on linear multivariable control systems, IEEE Trans., AC-*26*, (1), (1981).
2. GOLUB G. H. and KAHAN W. : "Calculating the Singular Values and pseudo-inverse of a Matrix", SIAM Numerical Analysis, (2), 202-224 (1965).
3. GOLUB G. H. and REINSCH C. : "Singular Value Decomposition and least squares solutions", Numer. Math., *14*, 403-420 (1970).
4. DONGARRA J. J. et al. : "LINPACK User's Guide", Philadelphia, P. A., SIAM, (1979).
5. FORSYTHE G. E. and MOLER C. B. : "*Computer Solution of Linear Algebraic Systems*", Prentice Hall, Englewood Cliffs, N. J., (1967).
6. KLEMA V. C. and LAUB A. J. : "The Singular Value Decomposition : Its computation and some Applications", IEEE Trans., AC-*25*, (2), (1980).

7. STAAR J. and VANDEWALLE J. : "On the relationship between singular value decomposition and principal axes and moments of inertia of a rigid body", to be submitted.

8. VAN DOOREN P. : "Numerical linear algebra : an increasing interest in linear system theory", Proc. Europ. Conf. on Circuit Theory and Design, ECCTD, The Hague, August 1981.

9. STAAR J., VANDEWALLE J. and WEMANS M. : "Realization of truncated impulse response sequences with prescribed uncertainty", Proc. IFAC World Congress VIII, Kyoto, Japan, Aug. 1981.

10. STAAR J. and VANDEWALLE J. : "Numerical implications of the choice of a reference node in the nodal analysis of large circuits", Int. Journal of Circuit Theory and Applications, (1981).



Jan STAAR was born in 1953. He received the Electrical Engineering degree from the Katholieke Universiteit Leuven, Belgium, in 1976, and is current ly working towards his doctoral degree in the E.S.A.T. Laboratory, Katholieke Universiteit Leuven.
From 1976 to 1977 he obtained a grant from I.W.O.N.L., and during 1977-1979 and 1980-1982 a scholarship from the N.F.W.O. (Belgian National Science foundation). Jan Staar received the governmental gold medal in 1971, the KVIV student award in 1976, and a third price in the IEEE student paper contest (region 8). His research interests are mainly in mathematical system theory and its applications in circuit theory and control.



Joos VANDEWALLE was born in 1948. He received the Electrical Engineering degree and the doctoral degree in applied sciences from the Katholieke Universiteit Leuven, Belgium, in 1971 and 1976, respectively.
From 1972 to 1976 he was assistant at the E.S.A.T. laboratory, Katholieke Universiteit Leuven. From 1976 to 1978 he was a Research Associate and from July 1978 to July 1979 he was Visiting Assistant Professor at the university of California, Berkeley. Since july 1979 he has a research and teaching position with the E.S.A.T. Laboratory of th Katholieke Universiteit Leuven, where he teaches Linear and Non-linear system theory courses. His research interests are mainly in mathematical system theory and its applications in circuit theory and control.