Mathematics in Signal Processing.
T. S. Durrani, J. B. Abbiss, J. E. Hudson, R. N. Madan,
J. G. McWhirter, and T. A. Moore (editors).
Oxford University Press, 1987.

THE NUMERICAL SOLUTION OF THE GENERAL GAUSS-MARKOV LINEAR MODEL

S. Hammarling
(Numerical Algorithms Group Ltd., Oxford)

1. INTRODUCTION

In this contribution we give an overview of the numerical solution of the general Gauss-Markov linear model given by

$$y = X\beta + e, \quad e \subset N(0,\sigma^2 W), \tag{1.1}$$

where the notation $e \subset N(0,\sigma^2 W)$ indicates that the noise vector e is assumed to come from a normal distribution with mean zero and variance-covariance matrix $\sigma^2 W$, W being a symmetric non-negative definite matrix.

We start by considering the standard case where W = I, then consider the more general case where W is any non-singular positive definite matrix and finally we consider the case where W is allowed to be singular. The material in this contribution is not new, but deserves an airing since it provides us with the basic tools for the reliable implementation of least-squares problems and the Kalman filter, and on a variety of architectures.

2. THE STANDARD GAUSS-MARKOV LINEAR MODEL

In this section we look at the standard model

$$y = X\beta + e, \quad e \subset N(0,\sigma^2 I). \tag{2.1}$$

The associated linear least-squares problem for this model is given by

$$\min_b \hat{e}^T\hat{e}, \quad \hat{e} = \hat{y} - \hat{X}b, \tag{2.2}$$

where $\hat{X}$ is the n by p matrix of observations, $\hat{y}$ the n element vector of dependent observations, b is the p element vector of regression coefficients and $\hat{e}$ is the n element error vector. Throughout this article we shall assume that there are at least as many observations as variables so that $n \geq p$, although much of the discussion is readily extended to the underdetermined case.

Tools for the reliable solution of (2.2) are the QR factorization and the singular value decomposition (SVD). In the next section we describe these factorizations. We shall not give the algorithmic detail of computing these factorizations, but instead give appropriate references.

3. THE QR FACTORIZATION AND THE SVD

The _QR_ _factorization_ of an n by p matrix X is given by

$$X = Q \begin{bmatrix} R \\ O \end{bmatrix} , \tag{3.1}$$

where Q is an n by n orthogonal matrix and R is a p by p upper triangular matrix. The SVD of X is given by

$$X = U \begin{bmatrix} D \\ O \end{bmatrix} V^T, \tag{3.2}$$

where U is an n by n orthogonal matrix, V is a p by p orthogonal matrix and D is a diagonal matrix with non-negative diagonal elements called the _singular values_ of X. The factorization can be chosen so that the singular values are in descending order down D. The first p columns of U are the left singular vectors of X and the columns of V are the right singular vectors of X.

If we perform the QR factorization (3.1) and then perform an SVD of R as

$$R = \tilde{U}\tilde{D}\tilde{V}^T \tag{3.3}$$

then

$$X = Q \begin{bmatrix} \tilde{U}\tilde{D}\tilde{V}^T \\ O \end{bmatrix} = Q \begin{bmatrix} \tilde{U} & O \\ O & I \end{bmatrix} \begin{bmatrix} \tilde{D} \\ O \end{bmatrix} \tilde{V}^T$$

$$= U \begin{bmatrix} D \\ C \end{bmatrix} V^T, \quad U = Q \begin{bmatrix} \tilde{U} & O \\ O & I \end{bmatrix} , \quad D = \tilde{D}, \quad V = \tilde{V} \tag{3.4}$$

and hence X and R have the same singular values and right singular vectors.

From (3.1) and (3.2) we find that

$$X^TX = R^TR \tag{3.5}$$

$$= VD^2V^T,$$

which are respectively the Cholesky and spectral factorizations of X^TX. Thus R and (V,D) give us alternative compact representations of X^TX without having to take the numerically damaging step of explicitly forming X^TX. (Golub, 1965; Hammarling, 1985.)

For many applications we can use the QR factorization when X is not close to being rank deficient, but carry on to the SVD otherwise. The SVD is a powerful tool for the analysis of rank deficient and near rank deficient problems. For example, from (3.2)

$$Xv_j = d_ju_j,$$

where d_j, u_j and v_j are respectively the jth singular value, left singular vector and right singular vector of X. Thus

$$||Xv_j||_2 = d_j$$

so that right singular vectors that correspond to small singular values give information on the near linear dependencies in the columns of X.

Both the QR factorization and the SVD can be obtained by numerically stable methods, based upon orthogonal transformations, and there are a number of sources of quality software that implement these methods. For background on both factorizations, the algorithms to compute them and many further references see Golub and Van Loan (1983).

For signal processing applications, a useful transformation is the plane rotation. This transformation is flexible and, unlike the hyperbolic rotation, is orthogonal and therefore has the desired numerical stability. Plane rotations are described in the above reference and in Figure 3.1 we illustrate the progress from X to R using the standard sequence of plane rotations for the case n=4, p=3. (Givens, 1958.)

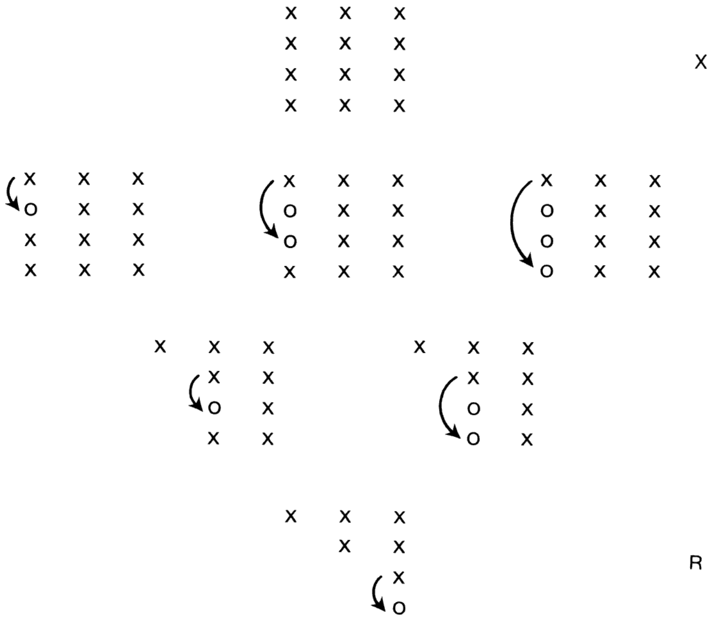

Fig. 3.1

A number of other sequences are possible, based upon introducing zeros into a column in different orders and two such alternative orderings are illustrated in Figure 3.2.

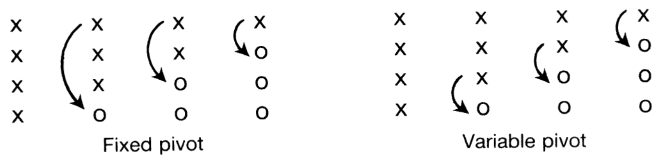

Fig. 3.2

Similarly we can perform plane rotations in disjoint planes in parallel, an idea that originates from a paper on error analysis for the QR factorization! (Gentleman, 1975; see also Modi and Clarke, 1984.) One such sequence for the case n=8, p=6 is illustrated in Figure 3.3, where rotations with the same index can be performed in parallel.

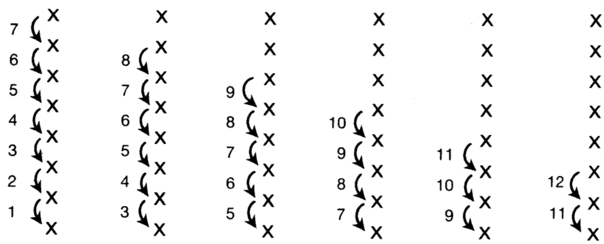

Fig. 3.3

The number of parallel operations for the illustrated sequence is (n+p-2) as opposed to ½p(2n-p-1) serial operations.

We can also readily update an existing QR factorization with additional observations as they become available, a particularly important feature for real time computing. This is illustrated in Figure 3.4 for the case p=3.

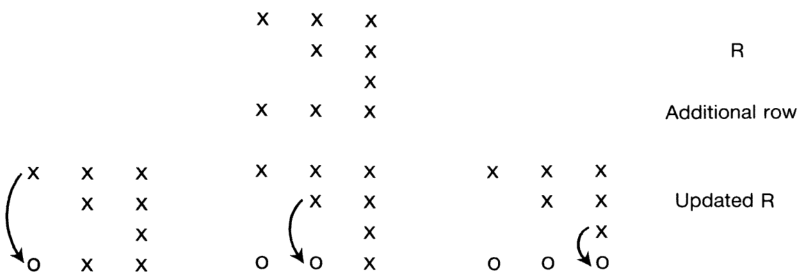

Fig. 3.4

These ideas form the basis of the systolic array implementation of the QR factorization. (Gentleman and Kung, 1981.)

For the SVD, as already mentioned, we can first apply the QR factorization and then find the SVD of R. The classical method of computing the SVD is to use the method originating with Golub and Kahan (1965) and realized as an Algol 60 algorithm by Golub and Reinsch (1970), in which R is reduced to bi-diagonal form and then the implicit QR algorithm is used to further reduce this to diagonal form. We note that R may be reduced to bi-diagonal form by applying plane rotations from both sides. This computation of the SVD is illustrated in Figure 3.5. The method corresponds to the reduction of R^TR to tri-diagonal form, followed by the QR algorithm applied to the tri-diagonal form.

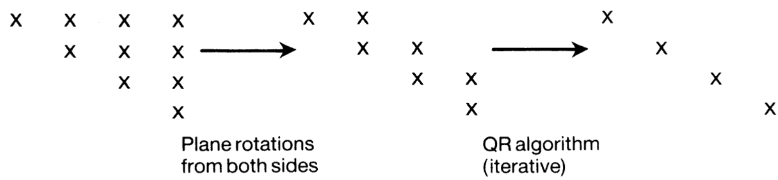

Fig. 3.5

More recently a method due to Kogbetliantz (1955) has been revived and in this method R is iteratively reduced to diagonal form, corresponding to Jacobi's method applied to R^TR. In Jacobi's method each plane rotation is chosen to solve the two by two symmetric eigenvalue problem so that $c=\cos\theta$ and $s=\sin\theta$ are such that

$$\begin{bmatrix} c & s \\ -s & c \end{bmatrix} \begin{bmatrix} x & y \\ y & z \end{bmatrix} \begin{bmatrix} c & -s \\ s & c \end{bmatrix} = \begin{bmatrix} x' & 0 \\ 0 & z' \end{bmatrix}$$

and a sweep of Jacobi's method applies $\frac{1}{2}n(n-1)$ such transformations with

$$\begin{bmatrix} x & y \\ y & z \end{bmatrix} = \begin{bmatrix} a_{ii} & a_{ij} \\ a_{ij} & a_{jj} \end{bmatrix}$$

for each i and j (j>1). The usual method chooses i and j cyclically, but other orderings are possible and, as with the QR algorithm we can apply transformations in parallel. (Modi, 1982; Modi and Parkinson, 1982.)

In the Kogbetliantz method (Figure 3.6) each transformation solves the two by two SVD problem, so that the left and right rotations are such that

$$\begin{bmatrix} c_1 & s_1 \\ -s_1 & c_1 \end{bmatrix} \begin{bmatrix} x & y \\ y' & z \end{bmatrix} \begin{bmatrix} c_2 & -s_2 \\ s_2 & c_2 \end{bmatrix} = \begin{bmatrix} x' & 0 \\ 0 & z' \end{bmatrix}$$

and one sweep of the method corresponds to one sweep of Jacobi. In the cyclic ordering one sweep makes R lower triangular and a second sweep restores the upper triangle. The Kogbetliantz approach is the basis of a systolic array implementation of the SVD. (Brent and Luk, 1985.)

$$\begin{matrix} x & x & x & x \\ & x & x & x \\ & & x & x \\ & & & x \end{matrix} \xrightarrow[\text{Iterative}]{} \begin{matrix} x & & & \\ & x & & \\ & & x & \\ & & & x \end{matrix}$$

Fig. 3.6

4. SOLVING THE STANDARD LINEAR LEAST-SQUARES PROBLEM

In this section we consider the solution of the least-squares problem

$$\min \hat{e}^T\hat{e}, \quad \hat{e} = \hat{y} - \hat{X}b \tag{4.1}$$

using the QR factorization of (3.1) and the SVD. If we put

$$Q^T\hat{e} = f \equiv \begin{bmatrix} f_1 \\ f_2 \end{bmatrix}, \quad Q^T\hat{y} = z \equiv \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}$$

then, since $f^Tf = \hat{e}^T\hat{e}$, (4.1) becomes

$$\min f^Tf, \quad f = z - \begin{bmatrix} R \\ 0 \end{bmatrix} b = \begin{bmatrix} z_1 - Rb \\ z_2 \end{bmatrix} \tag{4.2}$$

If $\hat{X}$ has full rank so that R is non-singular, we can choose b as the solution of the upper triangular equations

$$Rb = z_1 \tag{4.3}$$

and hence this solves (4.2). Thus

$$\hat{e}^T\hat{e} = z_2^Tz_2. \tag{4.4}$$

If $\hat{X}$ does not have full rank we can proceed with the SVD of R

$$R = \tilde{U}DV^T \tag{4.5}$$

and (4.2) is equivalent to

$$\min f_1^T f_1, \quad f_1 = z_1 - \tilde{U}DV^T b \tag{4.6}$$

Putting

$$\tilde{U}^T f_1 = g \equiv \begin{bmatrix} g_1 \\ g_2 \end{bmatrix}, \quad \tilde{U}^T z_1 = w \equiv \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}, \quad V^T b = c \equiv \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \tag{4.7}$$

and

$$D = \begin{bmatrix} S & O \\ O & O \end{bmatrix}, \tag{4.8}$$

where S is non-singular, then (4.6) becomes

$$\min g^T g, \quad g = w - \begin{bmatrix} S & O \\ O & O \end{bmatrix} c, \tag{4.9}$$

and this is solved for any c such that

$$Sc_1 = w_1 \tag{4.10}$$

for which

$$\hat{e}^T \hat{e} = z_2^T z_2 + w_2^T w_2. \tag{4.11}$$

c_2 is arbitrary and the solution for which $b^T b$ is minimized is called the minimal length solution and is a best linear unbiased estimate of β. Since $b^T b = c^T c$, this solution is given by

$$c_2 = O \tag{4.12}$$

so that

$$b = V \begin{bmatrix} S^{-1} w_1 \\ O \end{bmatrix}, \quad S^{-1} = \mathrm{diag}\ (1/d_i). \tag{4.13}$$

Other solutions, such as the reduced parameter solution, can readily be obtained via the SVD and this allows the decision on the rank of X to be made in the most favourable circumstance. (See for example, Hammarling, 1985, and the references given there.)

5. THE GENERAL GAUSS-MARKOV LINEAR MODEL

In this section we look at the general model

$$y = X\beta + e, \quad e \subset N(0, \sigma^2 W). \tag{5.1}$$

When W is non-singular this is usually associated with the weighted least-squares problem

$$\min_b \hat{e}^T W^{-1} \hat{e}, \quad \hat{e} = \hat{y} - \hat{X}b \tag{5.2}$$

If we let F be any matrix for which

$$W = FF^T \tag{5.3}$$

and put

$$r = F^{-1}\hat{e} \quad \text{so that} \quad r \subset N(0, \sigma^2 I) \tag{5.4}$$

then (5.2) becomes

$$\min r^T r, \quad r = F^{-1}\hat{y} - F^{-1}\hat{X}b \tag{5.5}$$

which is in the form of the standard problem (2.2), and so we can solve (5.5) as before using the QR factorization and the SVD. <u>But</u> this approach is numerically unstable if W is ill-conditioned and (5.2) and (5.5) are not even defined if W is singular. Adding in artificial noise is a completely unsatisfactory numerical approach for circumventing this.

In place of (5.4) we can let r be any vector such that

$$Fr = \hat{e} \tag{5.6}$$

and rearrange (5.5) to give

$$\min r^T r, \text{ subject to: } \hat{y} = \hat{X}b + Fr. \tag{5.7}$$

While being superficially a trivial rearrangement this has the vital difference of not now requiring non-singularity of F and hence of W. Notice that if

$$r \subset N(0, \sigma^2 I) \quad \text{we still have} \quad \hat{e} \subset N(0, \sigma^2 W)$$

(5.7) is called the <u>generalized linear least-squares problem</u>. (Paige, 1978.) Tools for the reliable solution of (5.7) are the QR factorization, the SVD and the generalized singular value decomposition (GSVD). In the next section we indicate how the problem may be solved with the QR factorization and the SVD and in the following section we define the GSVD and give references to its computation and its use in solving (5.7).

6. SOLVING THE GENERALIZED LINEAR LEAST-SQUARES PROBLEM

Here we consider the solution of the problem

$$\min r^T r, \text{ subject to: } \hat{y} = \hat{X}b + Fr \tag{6.1}$$

using the QR factorization of (3.1). Denote the QR factorization of $\hat{X}$ by

$$\hat{X} = Q_x \begin{bmatrix} R_x \\ O \end{bmatrix} \tag{6.2}$$

and put

$$Q_x^T \hat{y} = z \equiv \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}, \quad Q_x = (Q_1 \; Q_2) \tag{6.3}$$

so that (6.1) becomes

$$\min r^T r, \text{ subject to: } z = \begin{bmatrix} R_x \\ O \end{bmatrix} b + \begin{bmatrix} Q_1^T F \\ Q_2^T F \end{bmatrix} \tag{6.4}$$

If $\hat{X}$ does not have full rank we can proceed to compute the SVD of R_x, otherwise for any r we can determine b as the solution of the upper triangular equations

$$R_x b = z_1 - (Q_1^T F) r \tag{6.5}$$

so that (6.4) becomes

$$\min r^T r, \text{ subject to: } z_2 = (Q_2^T F) r \tag{6.6}$$

If we now denote the QR factorization of $F^T Q_2$ as

$$F^T Q_2 = Q_F \begin{bmatrix} R_F \\ O \end{bmatrix} \tag{6.7}$$

and put

$$Q_F^T r = p \equiv \begin{bmatrix} p_1 \\ p_2 \end{bmatrix} \tag{6.8}$$

then (6.6) becomes

$$\min p^T p, \quad \text{subject to:} \quad z_2 = R_F^T p_1 . \tag{6.9}$$

If R_F does not have full rank we can proceed to compute its SVD, otherwise we can solve the lower triangular system $R_F^T p_1 = z_2$ for p_1 and hence p is given by

$$R_F^T p_1 = z_2, \quad p_2 = 0 \tag{6.10}$$

so that

$$r^T r = p_1^T p_1, \quad r = Q_F \begin{bmatrix} p_1 \\ 0 \end{bmatrix}$$

and b is given from (6.5). Details on this type of approach to the solution of (6.1) are given in Paige (1978, 1979a), error and perturbation analysis in Paige (1979b) and some statistical analysis in Kourouklis and Paige (1981). Further details on the SVD approach are given in Hammarling, Long and Martin (1983).

7. THE GENERALIZED SINGULAR VALUE DECOMPOSITION

To give a gentle introduction to the GSVD we suggest a generalization of the QR factorization of the single matrix to the matrix pair $(\hat{X},F)$. We define the generalized QR factorization as

$$\hat{X} = Q_x \begin{bmatrix} R_x \\ 0 \end{bmatrix}, \quad Q_x^T F = R_F Q_F^T, \tag{7.1}$$

where Q_x and Q_F are orthogonal and R_x and R_F are upper triangular. If F is non-singular we find that

$$F^{-1}\hat{X} = Q_F \begin{bmatrix} R_F^{-1} R_x \\ 0 \end{bmatrix} \quad , \tag{7.2}$$

which is the QR factorization of $F^{-1}\hat{X}$, the data matrix in (5.5). (7.1) gives all the information in (7.2) without relying on invertibility of F and without requiring the numerically damaging step of computing $F^{-1}\hat{X}$. Applying this to (6.1) and putting

$$Q_X^T\hat{y} = z \equiv \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}, \quad Q_F^T r = p \equiv \begin{bmatrix} p_1 \\ p_2 \end{bmatrix}, \quad R_F = \begin{bmatrix} R_{11} & R_{12} \\ 0 & R_{22} \end{bmatrix} \tag{7.3}$$

we get

$$\min p^T p, \text{ subject to: } z = \begin{bmatrix} R_x \\ 0 \end{bmatrix} b + R_F p \tag{7.4}$$

and if R_x and R_{22} are of full rank we solve (7.4) as

$$R_{22}p_2 = z_2, \; p_1 = 0, \; R_x b = z_1 - R_{12}p_{22}, \; r^T r = p_2^T p_2. \tag{7.5}$$

There are variations in the way one can define the GSVD, but the appropriate form for the generalized linear least-squares problem (Paige, 1985) is to define the generalized SVD for the matrix pair $(\hat{X}, F)$ as

$$\hat{X} = Q \begin{bmatrix} R \\ 0 \end{bmatrix} \begin{bmatrix} I & 0 & 0 \\ 0 & C & 0 \\ 0 & 0 & 0 \end{bmatrix} V^T, \quad F = Q \begin{bmatrix} R \\ 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & S & 0 \\ 0 & 0 & I \end{bmatrix} U^T, \tag{7.6}$$

where Q, U and V are orthogonal, R is a non-singular upper triangular matrix and C and S are diagonal with

$$C^2 + S^2 = I, \quad c_i > 0, \quad s_i > 0.$$

We can choose the diagonal elements of C, the c_i, to be in descending order, in which case the diagonal elements of S are of course in ascending order. The pairs (c_i, s_i) are called the generalized singular values of $(\hat{X}, F)$.

If F is non-singular then

$$F^{-1}\hat{X} = U \begin{bmatrix} S^{-1}C & O \\ O & O \end{bmatrix} V^T, \tag{7.7}$$

which is the SVD of $F^{-1}\hat{X}$, and

$$\hat{X}\hat{X}^T = (QR) \begin{bmatrix} C^2 & O \\ O & O \end{bmatrix} (QR)^T, \quad FF^T = (QR) \begin{bmatrix} S^2 & O \\ O & I \end{bmatrix} (QR)^T, \tag{7.8}$$

so that (QR) is the congruence matrix that simultaneously diagonalizes $\hat{X}\hat{X}^T$ and FF^T. The values (c_i^2/s_i^2) are the non-zero eigenvalues of the generalized symmetric eigenvalue problem

$$\hat{X}\hat{X}^T z = \lambda FF^T z.$$

For background on the GSVD the reader is strongly recommended to consult Van Loan (1976) and Paige and Saunders (1981).

THE GSVD can readily be applied to (6.1) to yield the solution of the generalized linear least-squares problem and just as the SVD is a valuable tool for the solution and analysis of the standard linear least-squares problem, so the GSVD plays the same role for the generalized problem as is deomstrated in Paige (1985). Stable numerical methods are emerging for computing the GSVD, (Stewart, 1983; Van Loan, 1984; Paige, 1984), the last of these, being based on implicit Kogbetliantz approach, has potential for systolic implementation (Brent, Luk, and Van Loan, 1983.)

8. CONCLUSION

We have discussed some of the numerical tools that are necessary for the reliable solution of the general Gauss-Markov linear model and have given references to the details necessary to implement this method. The crux of such methods is the use of orthogonal transformations, which are numerically stable (for example, Wilkinson, 1965), and the avoidance of the numerically damaging steps of forming normal matrices of the form $\hat{X}^T\hat{X}$ and of matrix inversion.

The Kalman filter can be viewed as a generalized least-squares problem and the ideas here form the basis of reliable and efficient methods for the numerical solution of the Kalman filter problem. (Hammarling, 1985.)

9. REFERENCES

Brent, R.P. and Luk, F.T., (1985) The solution of singular value and symmetric eigenvalue problems on multiprocessor arrays. *SIAM J. Sci. Stat. Comput.*, **6**, 69-84.

Brent, R.P., Luk, F.T. and Van Loan, C., (1983) Computation of the generalized singular value decomposition using mesh-connected processors. *Proc. SPIE,* Vol. 431, Real-Time Signal Processing, VI.

Gentleman, W.M., (1975) Error analysis of QR decompositions by Givens transformations. *Linear Algebra Applic.*, **10**, 189-197.

Gentleman, W.M. and Kung, H.T., (1981) Matrix triangularization by systolic arrays. *Proc. SPIE,* Vol. 298, Real-Time Signal Processing IV.

Givens, W., (1958) Computation of plane unitary rotations transforming a general matrix to triangular form. *SIAM J. Appl. Math.*, **6**, 26-50.

Golub, G.H., (1965) Numerical methods for solving linear least-squares problems. *Num. Math.*, **7**, 206-216.

Golub, G.H. and Kahan, W., (1965) Calculating the singular values and pseudo-inverse of a matrix. *SIAM J. Num. Anal.*, **2**, 202-224.

Golub, G.H. and Reinsch, C., (1970) Singular value decomposition and least-squares solutions. *Num. Math.*, **14**, 403-420.

Golub, G.H. and Van Loan, C.F., (1983) Matrix Computations. North Oxford Academic, Oxford.

Hammarling, S.J., (1985) The singular value decomposition in multivariate statistics, Signum Newsletter, No. 3.

Hammarling, S.J., (1985) The Numerical solution of the Kalman filtering problem. NAG Technical Report, TR1/85, NAG Central Office, 256 Banbury Road, Oxford, OX2 7DE, UK.

Hammarling, S.J., Long, E.M.R. and Martin, D.W., (1983) A generalized linear least-squares algorithm for correlated observations with special reference to degenerate data. NPL Report DITC 33/83, National Physical Laboratory, Teddington, Middlesex, TW11 OLW, UK.

Kogbetliantz, E.G., (1955) Solution of linear equations by diagonalization of coefficients matrix. *Quart. Appl. Math.*, **13**, 123-132.

Kourouklis, S. and Paige, C.C., (1981) A constrained least-squares approach to the general Gauss-Markov linear model. *J. Amer. Statist. Assoc.*, **76**, 620-625.

Modi, J.J., (1982) Jacobi methods for eigenvalue and related problems - in a parallel computing environment. Ph.D. Thesis, University of London.

Modi, J.J. and Clarke, M.R.B., (1984) An alternative Givens Ordering. *Numer. Math.*, **43**, 83-90.

Modi, J.J. and Parkinson, D., (1982) Study of Jacobi methods for eigenvalues and singular value decomposition on DAP. Computer Physics Communications, 26, 317-320.

Paige, C.C., (1978) Numerically stable computations for general univariate linear models. Commun. Statist.-Simula. Computa., B7(5), 437-453.

Paige, C.C., (1979a) Fast stable computations for generalized linear least-squares problems. *SIAM J. Num. Anal.*, **16**, 165-171.

Paige, C.C., (1979b) Computer solution and perturbation analysis of generalized linear least-squares problems. *Maths Comp.*, **33**, 171-183.

Paige, C.C., (1985) The general linear model and the generalized singular value decomposition. *Linear Algebra Applic.*, **70**, 269-284.

Paige, C.C., (1984) Computing the generalized singular value decomposition. School of Computer Science, McGill University, Montreal, Quebec, Canada, H3A 2K6. (Submitted to *SIAM J. Sci. Stat. Comput.*)

Paige, C.C. and Saunders, M.A., (1981) Towards a generalized singular value decomposition. *SIAM J. Num. Anal.*, **18**, 398-405.

Stewart, G.W., (1983) A method for computing the generalized singular value decomposition. In "Matrix Pencils". Eds. Kagstrom, B. and Ruhe, A., Springer-Verlag, Berlin.

Van Loan, C.F., (1976) Generalizing the singular value decomposition, *SIAM J. Num. Anal.*, **13**, 76-83.

Van Loan, C.F., (1984) Computing the CS and generalized singular value decomposition. Technical Report CS-604, Dept. Computer Science, Cornell University, Ithaca, New York.

Wilkinson, J.H., (1965) The Algebraic Eigenvalue Problem. Oxford University Press, London.