NASA Technical Memorandum 100021

The Minimal Residual QR-Factorization Algorithm for Reliably Solving Subset Regression Problems

M. H. Verhaegen

September 1987

(NASA-TM-100021) THE MINIMAL RESIDUAL QR-FACTORIZATION ALGORITHM FOR RELIABLY SOLVING SUBSET REGRESSION PROBLEMS (NASA) 15 p Avail: NTIS HC A02/MF A01 CSCL 12A

N87-30131

Unclas

G3/65 0103494

National Aeronautics and Space Administration

NASA Technical Memorandum 100021

The Minimal Residual QR-Factorization Algorithm for Reliably Solving Subset Regression Problems

M. H. Verhaegen, Ames Research Center, Moffett Field, California

September 1987

National Aeronautics and
Space Administration

Ames Research Center
Moffett Field, California 94035

The minimal residual QR-factorization algorithm for reliably solving subset regression problems

M.H. Verhaegen*

September 10, 1987

Abstract

In this paper, a new algorithm to solve subset regression problems is described, called the *minimal residual QR-factorization* algorithm (MRQR). This scheme performs a QR factorization with a new column-pivoting strategy. Basically, this strategy is based on the change in the residual of the least-squares problem.

Furthermore, it is demonstrated that this basic scheme might be extended in a numerically efficient way to *combine* the advantages of existing numerical procedures, such as the singular value decomposition, with those of more classical statistical procedures, such as stepwise regression. In this paper this extension is presented as an advisory-expert system that guides the user in solving the subset regression problem.

The advantages of the new procedure are highlighted by a numerical example.

1 Introduction

The least-squares problem considered in this paper is formulated as

$$\min_{x} \|Ax - b\|_2 \tag{1}$$

where $A \in R^{m \times n}$ $(m \geq n)$,$b \in R^m$ and $\|.\|_2$ denotes the Euclidean norm of a vector.

For a number of applications, equation (1) is not solvable because of either the *redundancy* in the specified columns of the A-matrix in (1) or the specification of columns bearing *"little" relationship* with the right hand side (rhs) b in (1). These two types of difficulties have been treated in the literature along two different lines: The first difficulty could reliably be treated and analyzed using the singular-value decomposition (SVD) [9]. An attractive way to treat the second difficulty is via the stepwise regression technique (SRT) [12].

In practice, these two difficulties might arise in combination. Therefore, this certainly provides motivation for a procedure that handles both difficulties in a reliable way. For example, in the aerodynamic model identification problem [2], [11], this (new) procedure should address the following problem formulation

*Associate of the US National Research Council, NASA Ames Research Center, M.S. 210-9, Moffett Field, CA 94035

Select the *minimal* number of columns of A which *maximally* "explain" the rhs b (or which yield the smallest residual).

This particular problem is analyzed in this paper and referred to as the subset regression problem (SRP).

For this type of SRP, existing solutions have a number of major drawbacks. The drawbacks of the SVD are:

1. It does not perform a subset selection from the originally defined parameter vector x in (1), but generally finds estimates for *all* its components. This violates the minimality condition in the SRP formulation. To address this problem, an attempt to extend the capabilities of the SVD is made in [3]. Here additional use is made of the QR factorization with column pivoting as described in [8]. This not only increases the computational complexity of the overall procedure, but a number of examples remain where this pivoting strategy fails (e.g., in [4] pp. 791-792).

2. In [5] it was demonstrated that the use of the SVD in solving least-squares problems generally requires different column-scaling procedures which often are *contradictory.*

3. The ϵ-numerical rank, defined in [3], which is a key element in the solution based on the SVD, is determined based only on the A-matrix. The rhs b is not included in this process. However, in solving so-called "ill-defined, rank-deficient least squares problems" [1] the rhs b might supply the crucial information.

The SRT may be considered as the most suitable technique presently available to address the SRP. It certainly overcomes a number of the drawbacks of the SVD; however, its drawbacks are:

1. *Numerical robustness* may be limited because of the use of the normal equations.

2. Even if its key decision parameters used in selecting individual columns of A, which often are taken as the partial F-ratios, are computed reliably, this selection process might become *indefinite* or lead to an *ill-conditioned* solution, as demonstrated in this paper.

In this paper, a new procedure is described that overcomes the drawbacks of the techniques outlined above. The basic algorithm in this procedure is called the *Minimal Residual QR factorization* algorithm (MRQR). This algorithm computes a QR factorization of A with a new column-pivoting strategy. The column pivoting is based on the change in residual of the least squares problem caused by the selection of the different "candidate" columns of A. It is shown that this technique can be used as a robust implementation of SRT. A description of the MRQR is given in section 2 and implementation aspects are briefly discussed in section 3. The extension of the MRQR is presented in this paper as an "advisory-expert" system; this is discussed in section 4. Some of the advantages of this new procedure are demonstrated by a numerical example in section 5. Finally, the concluding remarks are presented in section 6.

[1] Ill-defined, rank-deficient least squares problems arise when in the set of singular values of the A-matrix, ordered in decreasing magnitude, "small" values occur (relative to the errors on the entries of the A-matrix); however, there is no clear "gap" present between any set of two subsequent singular values.

2 A description of the MRQR

In this section an outline of the Minimal Residual QR algorithm is given in pseudo-programming-language form. The precise algorithmic details of computing the different parts of the core of the algorithm are discussed in the following section.

The MRQR

INITIALIZATION

$$\text{rank}^2 = n$$
$$A = A^1 = [a_1^1 \cdots a_{j-1}^1 a_j^1 a_{j+1}^1 \cdots a_n^1] \qquad b = b^1 \tag{2}$$

DO i = 1:n,

STEP 1. Select the column vector of A^i "most closely" related to b^i,
$\Rightarrow$ denote this column vector as a_j^i

STEP 2. Interchange the j^{th} and i^{th} column vector of A^i by the column permutation matrix π_i

STEP 3. Perform an orthogonal projection Q_i, such that

$$Q_i \left(\begin{array}{c|c} T^{(i-1)} & \star^{(i-1)\times(n-i+1)} \\ \hline 0 & a_i^i \cdots a_j^i \cdots a_n^i \end{array} \right) \pi_i = \left(\begin{array}{c|c} T^{(i)} & \star^{i\times(n-i)} \\ \hline 0 & a_{i+1}^{i+1} \cdots a_n^{i+1} \end{array} \right) = \left(\begin{array}{c|c} T^{(i)} & \star^{i\times(n-i)} \\ \hline 0 & A^{i+1} \end{array} \right) \tag{3}$$

and $Q_i b^i = \left(\begin{array}{c} \star^i \\ \hline b^{i+1} \end{array} \right)$

STEP 4. Rank determination test:
This can more generally be referred to as the "termination test" and here different methods may be desirable to terminate the MRQR depending on the goal of our SRP.
a) *Numerically* we could be interested in the ϵ-numerical rank of the matrix A [3].
b) *Practically* we could be interested in whether the residual of our least squares problem is sufficiently small or the remaining "candidate" columns of A are related to the rhs in a (statistically) insignificant way.

END:
STOP:

The main part of this algorithm are the four steps in the "DO-loop". A combined execution of these four steps is referred to as a "sweep" of the MRQR algorithm.

[2] The notion of rank in the MRQR may be considered for our purposes as the (minimal) number of columns of A that yield the minimal residual of the least squares problem (additional comments are given in STEP 4 of the DO-loop)

3 Implementation of the MRQR algorithm

A crucial task in the MRQR scheme is the selection of that column of A^i "most closely" related to the rhs vector b^i. This task is formulated in the first step of each sweep of the MRQR algorithm. A possible way to put this formulation into a mathematical framework is as follows:

Find the column vector of A^i having the minimal residual when solving the following sub-least squares problem:

$$\min_{\xi} \|a^i_\ell \xi - b^i\| \tag{4}$$

for each column vector of A^i.

When using this problem formulation, the first step is to first compute these residuals. The residuals resulting from solving the set of problems defined in (4) are denoted by $\|m^i_\ell\|$. The second step is to find the ***minimum*** of the obtained sequence $\{\|m^i_\ell\|$ for $\ell = i : n\}$. The calculation of the residuals when the entries of A are noise-free is addressed next.

For any set of two vectors $\{a^i_\ell, b^i\}$, that define problem (4), the residual is the length of the component of b^i orthogonal to a^i_ℓ. This component is denoted in Fig. 1 by m^i_ℓ.

There are a number of known procedures for computing $\|m^i_\ell\|_2$. A *first* procedure is simply based on the sine-rule of a rectangular triangle. In this case, $\|m^i_\ell\|_2$ (see Fig. 1) can be expressed as

$$\|m^i_\ell\|_2 = \|b^i\|_2 \sin\phi^i_\ell \tag{5}$$

or it can be written as

$$\|m^i_\ell\|_2^2 = \|b^i\|_2^2(1 - \cos^2\phi^i_\ell) \tag{6}$$

The quantities sin ϕ^i_ℓ and cos ϕ^i_ℓ in (5) and (6) may be calculated as follows:

$$\sin\phi^i_\ell = \frac{\|a^i_\ell \times b^i\|_2}{\|a^i_\ell\|_2 \|b^i\|_2} \tag{7}$$

$$\cos\phi^i_\ell = \frac{a^i_\ell \cdot b^i}{\|a^i_\ell\|_2 \|b^i\|_2} \tag{8}$$

Equation (8) can be computed rather straigtforwardly and substituted in (6). On the other hand, the cross-product in (7) requires some further explanation. A possible way to evaluate that cross-product is as follows:

$$\|a^i_\ell \times b^i\|_2^2 = \det(< a^i_\ell \quad b^i >' < a^i_\ell \quad b^i >) \tag{9}$$

A reliable procedure to find the determinant in (9) can use either the SVD or QR decomposition of $< a^i_\ell \quad b^i >$, since either gives

$$|\det < a^i_\ell \quad b^i > | = s_1 s_2 = |r_{11}||r_{22}| \tag{10}$$

where s_i denotes the i^{th} singular value and r_{ii} the i^{th} diagonal element of the upper triangular factor of the SVD and QR decomposition, respectively, of $< a_\ell^i \;\; b^i >$.

A *second* procedure is to compute the residual $\|m_\ell^i\|_2$ directly from solving the sub-least squares problems stated in (4) for all columns of the A^i-matrix. For example, this can be done by means of a QR decomposition of the matrix $< a_\ell^i \;\; b^i >$ [8].

From the set of residuals $\{\|m_i^i\|_2, \cdots, \|m_n^i\|_2\}$ of the different column vectors of A^i, we now have to select the *minimum*. The column corresponding to this minimum is represented by a_j^i in STEP 1 of the DO-loop described in section 2.

4 Extending the MRQR to an advisory-expert system

4.1 Difficulties with the column selection procedure in the MRQR

As outlined in section 3, the key parameters in the column selection procedure of the MRQR are the set of residuals $\{\|m_i^i\|_2, \cdots, \|m_n^i\|_2\}$. In [1] it has been demonstrated that these residuals are related in a very simple manner to the statistical partial F-ratios used in the SRT. Therefore, the MRQR might be used as a *robust* implementation of the SRT.

However, even if these F-ratios or residuals can be reliably computed, the following difficulties might arise:

1. The set of residuals $\{\|m_i^i\|_2, \cdots, \|m_n^i\|_2\}$ might have *several minima*, causing the column selection process of the MRQR to become indefinite (an example of this phenomenon is given in section 5).
2. The same set of residuals might show a clear minimum, say $\|m_j^i\|_2$; however, the inclusion of the corresponding column vector of A leads to a very badly conditioned submatrix $T^{(i)}$ of (3).

A possible way to overcome this set of difficulties is to *reformulate* the selection procedure into the following "min-max" formulation:

> Find the column vector a_j^i from the candidate columns $\{a_i^i, \cdots, a_n^i\}$ which has a *minimal* residual and which leads to the *maximal*, smallest singular value of the matrix $T^{(i)}$ in (3).

The selection of the column a_j^i as defined in the min-max formulation corresponds to finding the minimum of the following new sequence:

$$\left\{\frac{\|m_i^i\|_2}{\sigma_i^i}, \cdots, \frac{\|m_n^i\|_2}{\sigma_n^i}\right\} \tag{11}$$

where σ_ℓ^i corresponds to the i-th singular value of the matrix $\left(\begin{array}{c|c} T^{(i-1)} & \star \\ \hline 0 & a_\ell^i \end{array}\right)$.

Remark 1: The above min-max formulation is the "exact" mathematical or algorithmic translation of the objective of the SRP stated in the introduction.

Remark 2: The quantities σ_ℓ^i of (11) can be computed individually by the inverse iteration method [7]. This calculation becomes efficient when operating only on the upper triangular matrix $\left(\begin{array}{c|c} T^{(i-1)} & \star \\ \hline 0 & \|a_\ell^i\|_2 \end{array}\right)$.

From this discussion it is clear that the difficulties just described allow (or force) us to incorporate additional decision parameters in the column selection process. Furthermore, it was demonstrated that one such additional parameter, namely the smallest singular value σ_ℓ^i in (11), can efficiently and reliably be calculated during the operation of the MRQR.

Other decision parameters can be derived in a similar way. (1) Additional *statistical* parameters can be computed, such as the partial-correlation coefficients as demonstrated in [6] or the set of collinearity coefficients defined in [5]. (2) Additional *numerical* parameters that might be derived are an estimate of the condition number of the upper triangular matrix defined in remark 3, or simply the distance of the remaining candidate columns to the already-selected columns of A (as is given by the quantity $\|a_\ell^i\|_2$ using the notation in (3)).

These different parameters not only overcome the difficulties of the original column-selection process in STEP 1 of the DO-loop of the MRQR, but also allow us to realize both termination schemes or a "trade-off" between them as was desired in STEP 4 of the same DO-loop.

The challenge we are now facing is to combine and use all or some of these decision parameters. This can be done by the proposition of an advisory-expert system that is described in the next section.

4.2 The advisory-expert system for solving SRP

In the proposed advisory-expert system, more than two decision parameters could be incorporated in the column-selection of the MRQR by making this algorithm *interactive*.

A decision screen is used as depicted in Fig. 2. This decision screen consists of *two levels*, which could be represented either simultanously or subsequently upon request of the user. The first level presents the residual sequence $\{\|m_i^i\|_2, \cdots, \|m_n^i\|_2\}$. In Fig. 2 this information is first presented graphically, making it easy to scan these quantities;then the additional numerical and statistical parameters are presented. This information is presented in Fig. 2 in a tabular form.

The main advantage of presenting the decision parameters in this format is that this information "only" advises the user on which column to select and that it is actually the user who makes that selection. Therefore, this present system makes it possible to use the "expertise" of the user and furthermore allows for the incorporation of an important class of information often available in practical applications [11], [2] referred to as "a priori" information.

5 A numerical example

The implementation of the MRQR as well as the derived advisory-expert system outlined respectively in sections 2, 3, and 4 are now evaluated via a numerical example. This example is performed on a VAX 780 computer in double precision and is taken from [10], p. 219. Here, a least-squares problem is stated

$$\min_x \left\| \begin{pmatrix} 1 & 21 & 41 & 61 \\ 2 & 22 & 42 & 62 \\ 3 & 23 & 43 & 63 \\ \vdots & \vdots & & \vdots \\ 20 & 40 & 60 & 80 \end{pmatrix} x - \begin{pmatrix} 4.95 \\ 5 \\ 5 \\ \vdots \\ 5 \end{pmatrix} \right\|_2 \tag{12}$$

By using the column notation of the A-matrix, as defined in (2), we clearly see that its *ϵ-numerical rank* (with ϵ taken equal to the machine precision) *is* 2, since

$$a_2 = \frac{2a_1 + a_4}{3} \qquad a_3 = \frac{a_1 + 2a_4}{3} \tag{13}$$

The MRQR is now applied to this example. Instead of automatically selecting the column of A having the minimal residual, the first level of the advisory-expert system as defined in section 4.2 was used to perform the column selection. This (basic) selection screen is shown in Fig. 3. From this figure it is obvious that the fourth column of A must be selected. The first level of the selection screen for the second sweep is shown in Fig. 4. At this time we requested the second level of the advisory-expert system. In the present example, the information supplied at this level consisted of the smallest singular value, defined in (11), and the distance between the candidate columns of A and the already-selected columns. This information is summarized in table 1. From this table, it again becomes clear that the original first column of the A-matrix is the correct choice.

	Original column number ℓ		
Decision parameter	1	2	3
Distance $\|\|a_\ell^2\|\|_2$	**21.87**	14.58	7.29
Smallest singular value σ_ℓ^2	**21.62**	13.36	5.92

Table 1: The second level of the second sweep of the advisory-expert system in the numerical example.

Finally, in the third sweep, the statistical parameters as well as the numerical parameters considered in the previous sweeps reveal either that the rhs has been explained entirely or that the ϵ-numerical rank of the matrix A is 2. Therefore, the algorithm is terminated at this point.

6 Concluding remarks

In this paper a new algorithmic solution is presented to solve subset regression problems (SRP). This solution is called the minimal-residual QR-factorization algorithm

(MRQR). It has been demonstrated that the MRQR could be extended to an advisory-expert system that also allows analysis of the SRP where the system matrix is ill-conditioned. This advisory-expert system combines the following advantages over existing techniques such as the singular-value decomposition and stepwise regression techniques.

1. It is a robust scheme that can also be implemented in a computationally-efficient manner.
2. It allows for the selection of a minimal subset of the *originally*-defined columns of the system matrix that lead to a minimum residual of the least squares problem. This has been achieved by the following unique features of the MRQR, that are exploited in the advisory-expert system:
   (a) The ability to combine numerical parameters, such as smallest singular values, as well as statistical parameters, such as partial F-ratios, in the column-selection process.
   (b) The versatility to terminate the algorithm, thereby allowing the user to focus on different problems that might arise with practical applications, such as a numerically ill-conditioned system matrix or candidate columns having no relationship (in a statistical sense) with the right-hand side.

In conclusion, the technique presented in this paper allows solution of the SRP or rank-deficient least squares problems where existing solutions fail. An example might be ill-defined rank-deficient least squares problems, which have been defined in the introduction. This topic is still under full investigation.

References

[1] M.H. Verhaegen, "Comparison of two numerical techniques for aerodynamic model identification," AIAA paper 87-2620-CP, Atmospheric Flight Mechanics Conference, Monterey, CA, Aug. 1987

[2] M.H. Verhaegen, "A new class of algorithms in linear system theory: with application to real-time aircraft model identification," Ph.D. dissertation, Catholic University, Leuven, Belgium, Nov. 1985

[3] G.H. Golub, V. Klema, and G.W. Stewart, "Rank degeneracy and least squares problems," Technical Report TR-456, Department of Computer Science, University of Maryland, College Park, MD, 1976

[4] W. Kahan, "Numerical linear algebra", Canadian Math. Bull. 9, No. 6, pp. 757-801, 1966

[5] G.W. Stewart, "Collinearity and least squares regression," Technical Report, Department of Computer Science, University of Maryland, College Park, MD (in preparation)

[6] G.H. Golub, "Matrix decomposition and statistical computation," in Statistical computation, R.C. Milton and J.A. Nelder, eds., Academic Press, N.Y., pp.365-397, 1969

[7] G.H. Golub and C.F. Van Loan, "Matrix computations," The Johns Hopkins University Press, Maryland, 1983

[8] J.J. Dongarra, C. Moler, J. Buch, and G. Stewart, "LINPACK user's guide," SIAM, Philadelphia, 1979

[9] C.L. Lawson and R.J. Hanson, "Solving least squares problems," Prentice-Hall, Englewood Cliffs, N.J., 1974

[10] G.E. Forsythe, M.A. Malcolm, C.B. Moler, "Computer methods for mathematical computations," Prentice-Hall, N.J., 1977

[11] V. Klein, J.G. Batterson, and P.C. Murphy, "Determination of airplane structure from flight data by using modified stepwise regression," NASA TP-1916, 1981

[12] N.R. Draper and H. Smith, "Applied regression analysis," John Wiley, N.Y., 1981

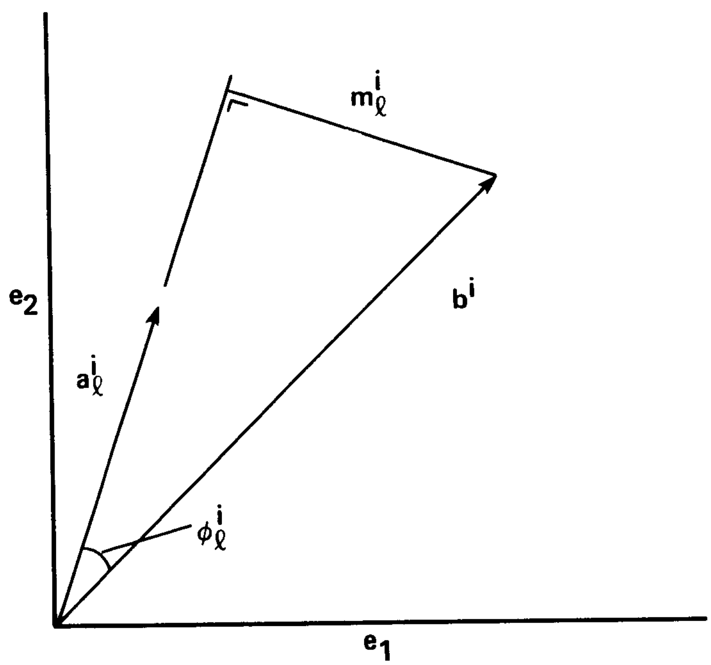

Figure 1: The residual m_ℓ^i of the set of vectors $\left\{a_\ell^i, b^i\right\}$.

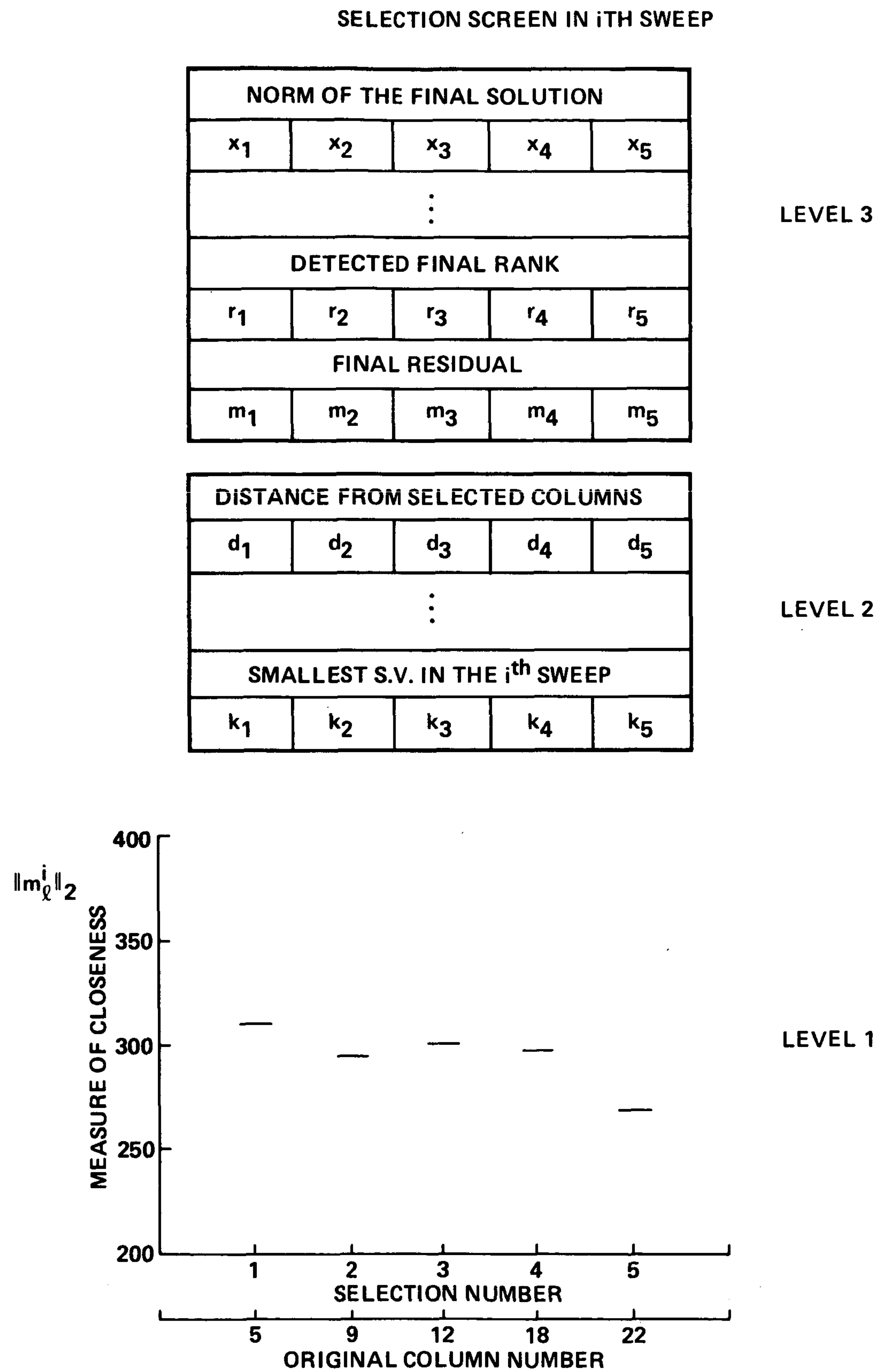

Figure 2: Possible outline of the advisory-expert system.

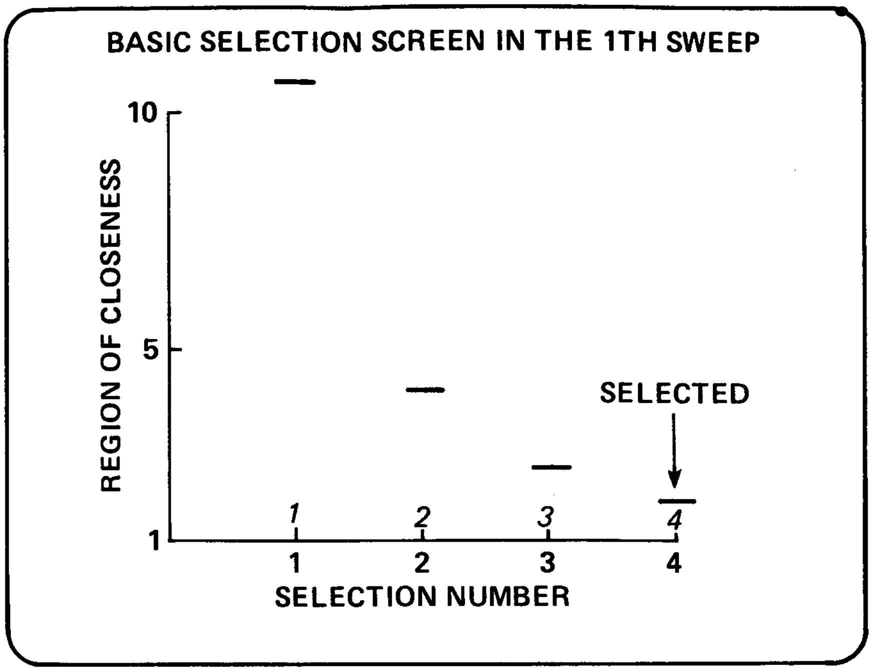

Figure 3: First level of the first sweep of the advisory-expert system in the numerical example.

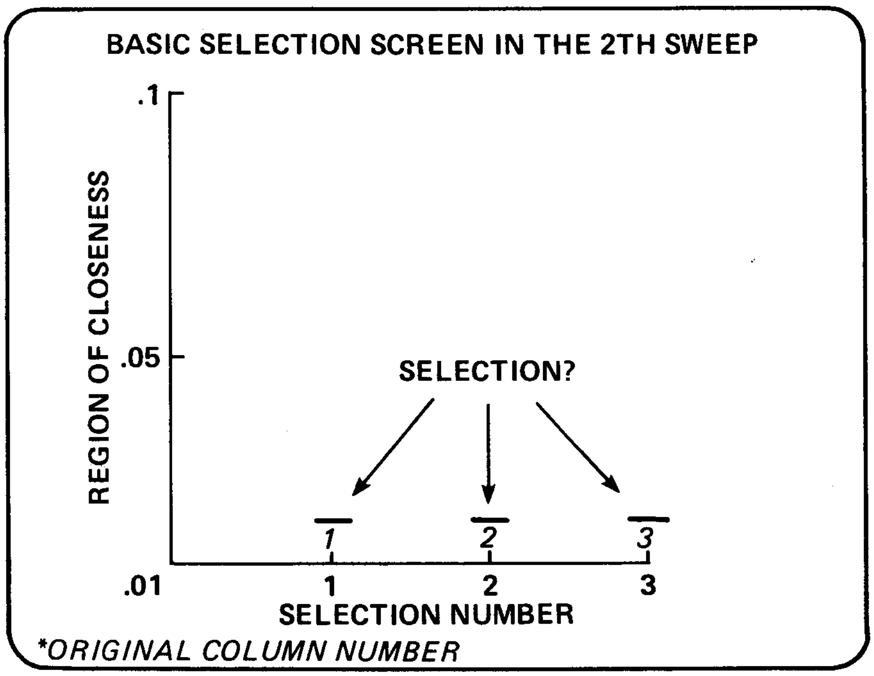

Figure 4: First level of the second sweep of the advisory-expert system in the numerical example.

NASA
National Aeronautics and Space Administration

Report Documentation Page

1. Report No. NASA TM 100021	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle The Minimal Residual QR-Factorization Algorithm for Reliably Solving Subset Regression Problems		5. Report Date September 1987
		6. Performing Organization Code
7. Author(s) M. H. Verhaegen		8. Performing Organization Report No. A-87334
		10. Work Unit No. 505-66-41
9. Performing Organization Name and Address Ames Research Center Moffett Field, CA 94035-5000		11. Contract or Grant No.
		13. Type of Report and Period Covered Technical Memorandum
12. Sponsoring Agency Name and Address National Aeronautics and Space Administration Washington, D.C. 20546-0001		14. Sponsoring Agency Code

15. Supplementary Notes

Point of Contact: M. H. Verhaegen, Ames Research Center, MS 210-9, Moffett Field, CA 94035-5000
(415)694-5983 or FTS 464-5983

16. Abstract

In this paper, a new algorithm to solve subset regression problems is described, called the minimal residual QR-factorization algorithm (MRQR). This scheme performs a QR factorization with a new column-pivoting strategy. Basically, this strategy is based on the change in the residual of the least-squares problem.

Furthermore, it is demonstrated that this basic scheme might be extended in a numerically efficient way to combine the advantages of existing numerical procedures, such as the singular value decomposition, with those of more classical statistical procedures, such as stepwise regression. In this paper this extension is presented as an advisory-expert system that guides the user in solving the subset regression problem.

The advantages of the new procedure are highlighted by a numerical example.

17. Key Words (Suggested by Author(s)) Numerical analysis Subset regression	18. Distribution Statement Unclassified–Unlimited Subject Category: 65

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of pages	22. Price
Unclassified	Unclassified	14	A02