A NOTE ON THE COMPUTATION OF
AN ORTHONORMAL BASIS FOR THE
NULL SPACE OF A MATRIX*

Thomas F. Coleman+
Danny C. Sorensen++

TR 82-510
August 1982

Department of Computer Science
Upson Hall
Cornell University
Ithaca, New York 14853

*Work supported in part by the Applied Mathematical Sciences Research Program (KC-04-02) of the Office of Energy Research of the U.S. Department of Energy under Contract W-31-109-Eng-38.

+Department of Computer Science, Cornell University, Ithaca NY 14853

++Mathematics and Computer Science Division, Argonne National Laboratory, Argonne IL 60

A Note on the Computation of
an Orthonormal Basis for the Null Space of a Matrix

Thomas F. Coleman

Computer Science Department

Cornell University

Ithaca, NY 14853

Danny C. Sorensen

Mathematics and Computer Science Division

Argonne National Laboratory

Argonne Illinois 60439

ABSTRACT

A highly regarded method to obtain an orthonormal basis, Z, for the null space of a matrix A^T is the QR decomposition of A, where Q is the product of Householder matrices. In several optimization contexts $A(x)$ varies continuously with x and it is desirable that $Z(x)$ vary continuously also. In this note we demonstrate that the **standard implementation** of the QR decomposition does **not yield** an orthonormal basis $Z(x)$ whose elements vary continuously with x. We suggest three possible remedies.

Work supported in part by the Applied Mathematical Sciences Research Program (KC-04-02) of the Office of Energy Research of the U.S. Department of Energy under Contract W-31-109-Eng-38.

1. Introduction

The question we are addressing in this short note is this: Let B be a ball around a point $x^* \epsilon R^n$. Suppose that $A(x)$ is an n by t matrix of rank t whose elements vary continuously with x on B. Is it possible to construct, stably and efficiently, a matrix $Z(x)$ with elements which **vary continuously** with x in B and with the additional properties

(1.1) $A(x)^T Z(x) = 0,$

(1.2) $Z(x)^T Z(x) = I_{(n-t)}$?

Several techniques for nonlinearly constrained optimization problems require the availability of a matrix $Z(x)$ with properties (1.1) and (1.2). (See, for example, Bartels and Conn[1982], Coleman and Conn[1982a,b], Kaufman[1975], Murray and Wright[1978], Murray and Overton[1980], Tanabe[1981], and Wright[1979]). Theoretical results given in Coleman and Conn[1982a,b] explicitly require that the elements of $Z(x)$ vary continuously in a ball around x^*, where x^* is a solution to the nonlinear programming problem. Kaufman **assumes** differentiability of $Z(x)$. The other references are not as explicit in their dependence on continuity however it would appear that possible future theoretical developments concerning projected quasi-Newton methods would also require that $Z(x)$ vary continuously. Surprisingly, the standard implementation of the QR factorization of $A(x)$, using Householder matrices (elementary reflectors), does not necessarily yield a matrix $Z(x)$ with continuously varying elements.

In section 2 we support this claim in detail. We suggest three possible remedies in section 3.

2. The Standard Implementation

A well-accepted procedure to obtain an orthonormal basis for the null space of A^T is given by Gill and Murray[1974]: Construct an orthogonal matrix $Q = (Q_1, Q_2)$ such that

(2.1) $Q_1^T A = R,$

where R is t by t and upper triangular, and

(2.2) $Q_2^T A = 0.$

We can then identify Z with Q_2. Unquestionably, the most popular method for obtaining such a

Q is the formation of a product of Householder matrices. Let us consider the simple case when $t = 1$ and $A = a = (a_1,a_2,...,a_n)^T$. The textbook rule for constucting Q is

$$Q \leftarrow I - \frac{2uu^T}{u^Tu}, \; where\; u = a + sgn(a_1)\cdot||\, a\, ||\cdot e^1,$$

and $sgn(a_1) = 1$ if $a_1 \geq 0$,

$= -1$ if $a_1 < 0$.

(The vector $(1,0,...,0)^T$ is denoted by e^1.) Now suppose that each component $a_i(x)$ is a continuous function of x in B. We wish to examine the continuity of Q with respect to $a(x)$. To do this it is useful to partition Q in the following way:

$$Q = \begin{pmatrix} q_{11} & q_{*1}^T \\ q_{*1} & \overline{Q} \end{pmatrix}$$

(Note that $Q_1 = \begin{pmatrix} q_{11} \\ q_{*1} \end{pmatrix}$ and the columns of $Q_2 = \begin{pmatrix} q_{*1}^T \\ \overline{Q} \end{pmatrix}$ are orthonormal bases for the range space of $a(x)$ and null space of $a(x)^T$, respectively.) It is straightforward to show that

$u^Tu = 2\cdot||\, a\, ||\cdot\{||\, a\, ||+ sgn(a_1)\cdot a_1\}$, and hence

$$q_{11} = \frac{-sgn(a_1)a_1}{||\, a\, ||},$$

$$q_{j1} = \frac{-sgn(a_1)a_j}{||\, a\, ||}, \text{ for } j>1,$$

$$\overline{Q}_{ij} = \frac{-a_i a_j}{||\, a\, ||\cdot\{||\, a\, ||+ sgn(a_1)a_1\}}, \quad \text{for } i\neq j, \text{ and}$$

$$\overline{Q}_{ii} = 1 - \frac{a_i^2}{||\, a\, ||\cdot\{||\, a\, ||+ sgn(a_1)a_1\}} .$$

It is clear that q_{11} and $\overline{Q}$ are continuous with respect to $a(x)$, however q_{*1} is discontinuous at the plane $a_1 = 0$. It follows that Q_2 is discontinuous at the plane $a_1 = 0$.

Therefore we cannot, in general, assume continuity of Q_2 when Q is computed in the standard way - this is unfortunately true even for B of arbitrarily small radius. Note that when $t = 1$ the only situation that is troublesome (for B of arbitrarily small radius) is when $a_1(x^*) = 0$. This observation leads us to the first of three possible strategies described in section 3.

We note that the elements of Q_1 do not change continuously with x. However, this is of no great concern since a continuously varying orthonormal basis for the range space of $A(x)$ is trivially available given Q_1. It is only necessary to monitor the signs of the diagonal elements of R and the corresponding columns of Q_1. Such a simple solution is not available for $Z(x)$.

3. Variations of the Standard Method

a. Row Orderings

For simplicity of presentation, we initially restrict ourselves to the case $t = 1$. Suppose that x^* is the point of convergence and $||\, a(x^*) \,|| \neq 0$. Hence there is an ordering of the rows of $a(x^*)$ such that $a_1(x^*) \neq 0$. Therefore, if this ordering is used for all x in B then $sgn(a_1(x))$ is equal to $sgn(a_1(x^*))$ for $||\, x-x^* \,||$ sufficiently small. Considering the formula for Q given given above, it is clear that in this case Q varies continuously.

We now turn to the general case where $A(x^*)$ is an n by t matrix of rank t. Consider the QR decomposition of A where Q is the product of a sequence of elementary reflectors: Let $a_i^{(j)}(x^*)$ be the i^{th} component of column j at the **start** of the j^{th} step of the QR decomposition of $A(x^*)$. Provided the rows of $A(x^*)$ are suitably ordered, and using linear independence, $a_j^{(j)}(x^*) \neq 0$. Hence if this ordering is used for all x in B, then for $||\, x-x^* \,||$ suitably small, $sgn(a_j^{(j)}(x))$ is equal to $sgn(a_j^{(j)}(x^*))$ and it follows that Q will vary in a continuous way.

Therefore, maintaining a continuous matrix $Z(x)$ in a neighbourhood B of x^* is possible (for B of sufficiently small radius) by suitably ordering the rows of $A(x)$ and applying the standard QR decomposition rules. Unfortunately, a suitable ordering is not known a priori: However, it is clear that any of a number of row-interchange tests could be employed such that interchange would not occur for $||\, x-x^* \,||$ sufficiently small.

b. Maintaining The Sign Bit

The source of our problems is the sign bit used in the standard rule for computing Q. Is it necessary? That is, can we always define Q as

(3.0) $Q \leftarrow I - \frac{2uu^T}{u^T u}$, *where* $u = a + ||a|| e^1$?

There are two apparent difficulties. Firstly, if $a = -||a|| e^1$, then u is the zero vector - let us ignore this problem temporarily. Secondly, if a is 'close' to $-||a|| e^1$, then it would appear that disastrous cancellation may occur in the computation of u and hence Q will be inaccurate. Parlett[1980, p.91] disputes the second claim and suggests that disastrous cancellation will not occur under these conditions if u is computed as follows:

(3.1) $s \leftarrow \sum_{i>1} a_i^2$,

(3.2) $u_1 \leftarrow \frac{-s}{(a_1 - ||a||)}$,

(3.3) $u_j \leftarrow a_j,\ j = 2,\ldots,n$.

Formula (3.1)-(3.3) does not involve the subtraction of nearly equal small quantities and thus we do not risk disastrous cancellation.

Therefore the following strategy seems appropriate: If $a_1 \geq 0$, then **compute** u_1 by

(3.4) $u_1 \leftarrow a_1 + ||a||$.

If $a_1 < 0$, then **compute** u_1 by (3.1) - (3.2). In either case we can obtain Q by (3.0).

Unfortunately, our problems are not over. Indeed the first difficulty, that Q is not defined at $\bar{a} = -||a|| e^1$, is rather troublesome. The kernal of the problem is this: Q (as defined by (3.0)) does not have a limit point at $\bar{a}$. Hence it is impossible to make an appropriate **definition** of Q($\bar{a}$). For example, consider that for $i \neq j$, $i \neq 1$, $j \neq 1$, $||a|| \neq |a_1|$,

$$Q_{ij} = \frac{-a_i a_j}{||a|| \cdot (||a|| + a_1)} = \frac{-a_i a_j (||a|| - a_1)}{||a|| (||a||^2 - a_1^2)}.$$

Hence

(3.5) $\lim_{a \to -e^1} Q_{ij}(a) = \lim_{a \to -e^1} \frac{-2 a_i a_j}{||\tilde{a}||^2}$,

where $\tilde{a} = (a_2, \ldots, a_n)^T$. But if a approaches $-e^1$ along the line $(-1, \epsilon, \epsilon, \ldots, \epsilon)$, then $Q_{ij} \rightarrow \frac{-2}{n-1}$.

However, if a approaches $-e^1$ along the line $(-1, \epsilon, \ldots, \epsilon, 0, \epsilon, \ldots, \epsilon, 0, \epsilon, \ldots, \epsilon)$, where the zeroes occur in positions i and j, then $Q_{ij} \to 0$.

Observe - these difficulties occur only when $a(x^*) = \pm \, || \, a(x^*) \, || \, e^1$. Also, if $a(x^*) = + \, || \, a(x^*) \, || \, e^1$ and $|| \, a(x^*) \, || > 0$, then there is a ball around x^* for which $a(x) \neq - || \, a(x^*) \, || \, e^1$, and vice versa. Therefore, if $a(x^*) \neq - || \, a(x^*) \, || \, e^1$, then formula (3.0) can be used for all x in a ball B around x^*. The elements of Q will vary continuously on B provided the radius of B is sufficiently small. Alternatively, if $a(x^*) \neq + \, || \, a(x^*) \, || \, e^1$, then (3.0) can be replaced with

(3.0') $\quad Q \leftarrow I - \dfrac{2uu^T}{u^T u}$, *where* $u = a - || \, a \, || \, e^1$.

If $a_1 < 0$,then we can **compute** u_1 by

(3.4') $\quad u_1 \leftarrow a_1 - || \, a \, || \, e^1$.

If $a_1 > 0$, then we can **compute** u_1 by

(3.2') $\quad u_1 \leftarrow \dfrac{-s}{a_1 + || \, a \, ||}$.

The elements of Q will vary continuously provided the radius of B is sufficiently small.

Unfortunately, one does not know, a priori, if $a(x^*) = \pm \, || \, a(x^*) \, || \, e^1$. However, it is clear that several switching rules could be employed in conjunction with (3.0) and (3.0') - if $x^k \to x^*$ the switching rule would become inactive for sufficiently large k.

For example, let $\{x^k\}$ be a sequence which converges to x^*. Denote $a(x^k)$ by a^k. A corresponding sequence of elementary reflectors could be defined by

(3.6) $\quad \theta_k \leftarrow \dfrac{\sigma_{k-1} \cdot a_1^k}{|| \, a^k \, ||}$,

(3.7) $\quad$ if $\theta_k \geq -\delta$ then $\sigma_k \leftarrow \sigma_{k-1}$ else $\sigma_k \leftarrow sgn(a_1^k)$,

(3.8) $\quad u^k \leftarrow a^k + \sigma_k || \, a^k \, || \, e^1$, (computed as above),

(3.9) $\quad Q_k \leftarrow I - \dfrac{2u^k (u^k)^T}{(u^k)^T u^k}$.

To begin, choose $\sigma_0 = sgn(a_1^0)$. The parameters θ and δ are introduced in an attempt to maintain the previous sign bit σ_{k-1}. This, in turn, results in the elements of $Z(x)$ (or $Q_2(x)$) behaving in a continuous manner. The parameter δ must satisfy $\delta < 1$, and should be positive in order to express a reluctance to change signs: say $\delta = .9$.

The analysis and procedures described in this section are given under the assumption that $t = 1$. The extension to the general case is straightforward and we will not go into detail.

c. Elementary Rotation Matrix

The third strategy that we investigate shares some features with the approach described above but is based upon elementary rotation matrices rather than reflectors. If q_1 , q_2 are two unit vectors with $q_1 \neq -q_2$ then the elementary rotation matrix sending q_1 into q_2 is

$$(3.10) \qquad P = I - \left(q_1, q_2\right) D \left(q_1, q_2\right)^T$$

where $D = \begin{pmatrix} 1 & 1 \\ -(1+2\gamma) & 1 \end{pmatrix}$, and $\gamma = q_1^T q_2$. Some properties of P are

(i) $P^T P = I$, (ii) $P q_1 = q_2$, (iii) $\lim_{q_1 \to q_2} P = I$.

Also, it can be readily verified that P rotates vectors in the plane, spanned by the vectors q_1 and q_2, through an angle of $\cos^{-1}(\gamma)$ with vectors orthogonal to this plane left untouched. Property (iii) is not shared by general elementary reflectors; it is this property which avoids the need for two definitions of the same transformation which are typically used to implement an elementary reflector stably. In fact if Q is of the form (3.0), with any nonzero vector u, then $\| Q - I \|_2 = 2$ - hence Q is never close to the identity transformation.

In the special case $q_1 = \dfrac{a}{\| a \|}$ and $q_2 = e_1$ the formula for P simplifies to

$$(3.11) \qquad P = \begin{pmatrix} \dfrac{a_1}{\| a \|} & \dfrac{\tilde{a}^T}{\| a \|} \\ \dfrac{-\tilde{a}}{\| a \|} & \bar{P} \end{pmatrix}$$

where $\tilde{a} = (a_2, \ldots, a_n)^T$,

$$\bar{P} = I - (\frac{1}{1+\gamma})\frac{\tilde{a}\tilde{a}^T}{||a||^2} \text{ and } \gamma = \frac{a_1}{||a||}.$$

This formula is briefly discussed by Parlett[1980, p.92, ex. 6-3-6], Note that P as defined in (3.11) can be stored and applied to a vector with the same efficiency as an elementary reflector. In fact only trivial modifications to existing QR codes are required to change from reflectors to rotators.

The elementary rotator P is not defined by (3.10) at points satisfying $a = \alpha e^1$, $\alpha < 0$: A strategy similar to that employed in part b must be used here also. That is, formula (3.10) can be used in a ball around x^*, provided $a(x^*) \neq -||a(x^*)||e^1$. If $a(x^*) \neq +||a(x^*)||e^1$, then P can be defined by (3.10) with the signs of the first row and column reversed and the definition of $\bar{P}$ changed to

$$\bar{P} = I - (\frac{1}{1-\gamma})\frac{\tilde{a}\tilde{a}^T}{||a||^2}.$$

It is clear that the elements of P will vary continuously in a ball B around x^* provided the radius of B is sufficiently small.

4. Concluding Remarks

We have suggested three different strategies for maintaining a continuous orthonormal basis for the null space of a matrix A^T which varies continuously with x. The first method has the attraction that the standard QR decomposition implementation can be employed. However, it has the disadvantage that row interchanges may be necessary in order to maintain continuity of $Z(x)$. Nevertheless, if the element of maximum modulus is initially pivoted into the first row (in the case $t = 1$) it seems highly unlikely that many subsequent interchanges would be necessary.

The second procedure (b.) does not require interchanges. It is based on the observation (Parlett[1980]) that disastrous cancellation need not occur when u is computed without the 'sign bit' provided the computation is done correctly. Discrete changes are necessary only in the extreme case when $a(x)$ oscillates between $+||a||e^1$ and $-||a||e^1$ - a highly unlikely scenario. On the negative side, this procedure cannot use a standard black box QR decomposition routine.

Finally, the third procedure (c.) has all of the advantages and some of the disadvantages attributed to method b. The elementary rotator (as described in c.) has some additional geometric appeal however. If the vector a is 'close' to $+ \| a \| e^1$ $(- \| a \| e^1)$, then a is rotated into $+ \| a \| e^1$ $(- \| a \| e^1)$. The opposite is true for elementary reflectors.

One may also require that $Z(x)$ have additional smoothness properties such as Lipschitz continuity or perhaps differentiability. It is clear that the strategies discussed in this note will allow $Z(x)$ to inherit all of the smoothness of A, in a ball around x^*, provided the rank of $A(x^*)$ is t.

Another popular way to obtain the QR decomposition of a matrix A is by using a sequence of Givens transformations. In the dense case the Givens procedure is more expensive than the elementary reflector approach. However, if A is sparse and the transformations are computed properly this may be the preferred method. An efficient way to compute and use Givens transformations in the sparse case is reported by Gentleman [1973] with further motivation and error analysis given in Gentleman [1975]. More discussion on the use of Givens transformations in the sparse situation may be found in George and Heath [1980]. Unfortunately, continuity difficulties also occur when Givens transformations are used. To see this suppose that $t = 1$ and both $a_1(x) \rightarrow 0$ and $a_n(x) \rightarrow 0$, where we assume that elements 1 and n define the Givens transformation that introduces a zero into position n. Depending on the manner in which a_1 and a_n converge, the corresponding Givens matrices may jump around wildly - this spells trouble for the continuity of Z(x). Continuity of Z(x) can be achieved in conjunction with the use of Givens transformations however, if a row interchange strategy (a.) was followed.

REFERENCES

R.H. Bartels and A.R. Conn[1982], An approach to nonlinear l_1 data fitting, Numerical Analayis Proceedings, 3rd IIMAS Workshop, Cocoyoc, Mexico.

T.F. Coleman and A.R. Conn[1982a], Nonlinear programming via an exact penalty function: Asymptotic analysis, Mathematical Programming, 24.

T.F. Coleman and A.R. Conn[1982b], On the local convergence of a quasi-Newton method for the nonlinear programming problem, TR 82-509, Department of Computer Science, Cornell University.

J.A. George and M.T. Heath[1980], Solution of sparse linear least squares problems using Givens rotations, Linear Algebra and Its Applications, 34, 69-83.

W.M. Gentleman[1973], Least squares computations by Givens transformations without square roots, J. Inst. Maths Applics 12, 329-336.

W.M. Gentleman[1975], Error analysis of QR decompositions by Givens transformations, Linear Algebra and its Applications 10, 189-197.

P. Gill and W. Murray[1974], *Numerical Methods for Constrained Optimization*, Academic Press, London.

L. Kaufman[1975], A variable projection method for solving separable nonlinear least squares problems, BIT 15, 49-57.

W. Murray and M. Overton[1980], A projected lagrangian algorithm for nonlinear minimax optimization, SIAM J. Sci. Stat. Comput., 1, 345-370.

W. Murray and M. Wright[1978], Projected lagrangian methods based on the trajectories of penalty and barrier functions, SOL 78-23, Stanford University.

B.N. Parlett[1980], *The Symmetric Eigenvalue Problem*, Prentice-Hall, Englewood Cliffs, N.J.

K. Tanabe[1981], Feasibility-improving gradient-acute-projection methods: a unified approach to nonlinear programming, Lecture Notes in Num. Appl. Anal., 3, 57-76.

M. Wright[1979], Algorithms for nonlinearly constrained optimization, SOL 79-24, Stanford University.