University of Maryland College Park

Institute for Advanced Computer Studies TR–95–91
Department of Computer Science TR–3533

The Triangular Matrices of Gaussian Elimination and Related Decompositions*

G. W. Stewart†

October, 1995

ABSTRACT

It has become a commonplace that triangular systems are solved to higher accuracy than their condition would warrant. This observation is not true in general, and counterexamples are easy to construct. However, it is often true of the triangular matrices from pivoted LU or QR decompositions. It is shown that this fact is closely connected with the rank-revealing character of these decompositions.

*This report is available by anonymous ftp from `thales.cs.umd.edu` in the directory `pub/reports`.

†Department of Computer Science and Institute for Advanced Computer Studies, University of Maryland, College Park, MD 20742. This work was supported in part by the National Science Foundation under grant CCR 95503126.

The Triangular Matrices Gaussian Elimination and Related Decompositions

G. W. Stewart

September 1995

ABSTRACT

It has become a commonplace that triangular systems are solved to higher accuracy than their condition would warrant. This observation is not true in general, and counterexamples are easy to construct. However, it is often true of the triangular matrices from pivoted LU or QR decompositions. It is shown that this fact is closely connected with the rank-revealing character of these decompositions.

1. Introduction

In 1961 J. H. Wilkinson [10] published a ground-breaking error analysis of Gaussian elimination. In the course of the paper he observed that triangular systems are frequently solved more accurately than their condition would warrant. In support of this observation he offered some examples and suggestive analyses, but no general theorems.

Wilkinson's observation has stood the test of time. But it has a touch of mystery about it. No general results can be proved, because it is easy to find innocuous looking matrices that are quite ill behaved — for example, an upper triangular matrix of standard normal deviates row scaled so that its diagonals are one. Thus any general bounds have to be weak. The weakness usually manifests itself by the appearance of a factor of 2^n in the bounds.

However, the matrices Wilkinson was chiefly concerned with were the unit lower triangular matrix L and the upper triangular matrix U that are produced by pivoted Gaussian elimination applied to a matrix A. It is the purpose of this paper to show that these matrices have special properties that derive from the rank revealing character of Gaussian elimination.

Specifically, it has been observed that Gaussian elimination with partial or complete pivoting tends to reveal the rank of a matrix in the sense that the diagonal elements of U are ball-park estimates of the smallest singular value of the corresponding leading principle submatrix of A. Under conditions that generally obtain for Gaussian elimination, this observation has two consequences.

1. Any ill-conditioning of U is artificial in the sense that if U is row scaled so that its diagonals are one, then its smallest singular value is near one.

2. The matrix L is well conditioned.

The first of these consequences implies that systems involving U will be solved accurately, since the best bounds on the accuracy of the solution do not depend on row scaling [3]. The second implies *ipso facto* that systems involving L will be solved accurately.

It might be objected that we have traded one mystery for another — the other being that Gaussian elimination tends to be rank revealing. The proper response is that it is not very mysterious. If A is exactly defective in rank, then Gaussian elimination must produce an exact zero on a diagonal of U, and by continuity the element will tend to remain small when A is perturbed slightly. In fact, most people are surprised to learn that Gaussian elimination can fail to detect a near degeneracy. Moreover, the examples on which Gaussian elimination does fail are invariably constructed by forming the product carefully chosen triangular factors. The mystery would be if Nature, who is ignorant of LU factorizations, should contrive to produce such a matrix.

In the next section we will establish the basic results on rank revealing triangular matrices. In the following section we will apply the results to Gaussian elimination, and in §4 to the QR and Cholesky factorizations. The paper concludes with a brief recapitulation.

In the following, $\|\cdot\|$ denotes the 2-norm. We will make free use of the properties of the singular value decomposition. For details see [2, 4, 6].

2. Lower bounds for singular values

The purpose of this section is to show that a triangular matrix whose principal minors reveal their rank becomes well conditioned when its rows are equilibrated. The heart of the development is a technical lemma that relates the smallest singular value of a triangular matrix to those of its largest leading principal submatrix.

Lemma 2.1. *Let*

$$\hat{R} = \begin{pmatrix} R & r \\ 0 & 1 \end{pmatrix}$$

be upper triangular. Let ρ be the smallest singular value of R and $\hat{\rho}$ be the smallest singular value of $\hat{R}$. If for some $\beta, \delta \in (0,1]$ the smallest singular value σ

of the matrix

$$\begin{pmatrix} R & r \\ 0 & \delta \end{pmatrix}$$

satisfies

$$\sigma \geq \beta\delta, \tag{2.1}$$

then

$$\hat{\rho} \geq \frac{\beta\rho}{\sqrt{\beta^2 + \rho^2}}. \tag{2.2}$$

Proof. Let

$$\begin{pmatrix} R & r \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ \xi \end{pmatrix} = \hat{\rho} \begin{pmatrix} y \\ \eta \end{pmatrix},$$

where the vectors $(x^{\mathrm{T}} \ \ \xi)$ and $(y^{\mathrm{T}} \ \ \eta)$ have norm one. Set

$$e = -\hat{\rho} R^{-1} y \quad \text{and} \quad \tilde{x} = x + e,$$

so that

$$\begin{pmatrix} R & r \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \tilde{x} \\ \xi \end{pmatrix} = \hat{\rho} \begin{pmatrix} 0 \\ \eta \end{pmatrix}.$$

Let

$$\nu = \left\| \begin{pmatrix} \tilde{x} \\ \xi \end{pmatrix} \right\|.$$

Then it follows from the relation

$$\frac{1}{\nu} \begin{pmatrix} R & r \\ 0 & \delta \end{pmatrix} \begin{pmatrix} \tilde{x} \\ \xi \end{pmatrix} = \begin{pmatrix} 0 \\ \nu^{-1}\hat{\rho}\delta\eta \end{pmatrix}$$

that $\nu^{-1}\hat{\rho}\delta\eta \geq \sigma$. Hence by (2.1),

$$\frac{\hat{\rho}\eta\delta}{\nu} \geq \beta\delta,$$

or

$$\hat{\rho} \geq \frac{\beta\nu}{\eta}. \tag{2.3}$$

Now

$$\|e\| = \|\hat{\rho} R^{-1} y\| \leq \rho^{-1}\hat{\rho}\|y\| = \rho^{-1}\hat{\rho}\sqrt{1 - \eta^2}.$$

Hence

$$\nu \geq 1 - \rho^{-1}\hat{\rho}\sqrt{1 - \eta^2},$$

and it follows from (2.3) that

$$\hat{\rho} \geq \frac{\beta\rho}{\rho\eta + \beta\sqrt{1-\eta^2}}. \tag{2.4}$$

The maximum of the denominator of this expression occurs when

$$\eta^2 = \frac{\rho^2}{\beta^2+\rho^2},$$

and its value is $\sqrt{\beta^2+\rho^2}$. The inequality (2.2) follows on substituting this value in (2.4). ■

We wish to apply this lemma to get lower bounds on the singular values of a row-equilibrated upper triangular matrix. We will denote the original matrix by U and assume without loss of generality that its diagonal elements δ_i are positive. Let

$$D = \mathrm{diag}(\delta_1, \delta_2, \ldots, \delta_n),$$

so that the equilibrated matrix is $D^{-1}U$.

To pin down what we mean for a triangular matrix to be rank revealing, we make the following definition.

Definition 2.2. *Let U be upper triangular of order n and let σ be the smallest singular value of U. We say that U is rank revealing of quality β if*

$$\sigma = \beta|u_{nn}|.$$

The quality factor β in this definition is always less than or equal to one, since $\sigma \leq |u_{nn}|$. The nearer β is to one, the better $|u_{nn}|$ estimates σ. It should be stressed, that this definition has been tailored to the requirements of this paper and is not meant to preempt other definitions of what it means to reveal rank.

We are going to use our lemma to get a recursion for a lower bound on the singular values of the leading principal submatrices of $D^{-1}U$. Let U_k and D_k denote the leading principal submatrices of order k of U and D. Let σ_k be the smallest singular value of U_k and, suppose U_k reveals σ_k with quality β_k, so that

$$\sigma_k = \beta_k\delta_k.$$

Let σ be the smallest singular value of

$$U_k = \begin{pmatrix} U_{k-1} & u_k \\ 0 & \delta_k \end{pmatrix},$$

Since the diagonal elements of D_{k-1}^{-1} are all greater than one, the smallest singular value σ of

$$\begin{pmatrix} D_{k-1}^{-1}U_{k-1} & D_{k-1}^{-1}u_k \\ 0 & \delta_k \end{pmatrix}$$

is not less than σ_k. Thus

$$\sigma \geq \beta_k \delta_k,$$

which is just the hypothesis (2.1) of Lemma 2.1. Applying the lemma with

$$\hat{R} = \begin{pmatrix} D_{k-1}^{-1}U_{k-1} & D_{k-1}^{-1}u_k \\ 0 & 1 \end{pmatrix},$$

we find that if ρ_k is the smallest singular value of $D_k^{-1}U_k$ — i.e., the equilibrated principal minor — then

$$\rho_k \geq \frac{\beta_k \rho_{k-1}}{\sqrt{\beta_k^2 + \rho_{k-1}^2}}. \tag{2.5}$$

We sum up these results in the following theorem.

Theorem 2.3. *If for $k = 2, 3, \ldots, n$ the leading principal minor U_k of U reveals its rank with quality β_k, and ρ_k the smallest singular values of U_k row-scaled so that its diagonal elements are one, then the ρ_k satisfy the recursion* (2.5).

There is a corresponding theorem for lower triangular matrices. The only difference is that the scaling is by columns rather than rows.

The recursion, starting with $\rho_1 = 1$, allows us to compute lower bounds on the smallest singular value of the leading principal minors of the equilibrated submatrices. In general if the β_k are uniformly bounded away from 1, the recursion converges to zero. However, if the β_k are uniformly bounded away from zero, the convergence is sublinear in the sense that

$$\frac{\rho_{k+1}}{\rho_k} \to 1.$$

Typically, a sequence converging sublinearly to zero shows an initial sharp decrease followed by an increasingly slow approach to zero. Our recursion is no exception. The following table exhibits values of the lower bounds for ρ_k when β_k is held constant.

k	β 0.50	0.10	0.01
5	2.4e−1	5.0e−2	5.0e−3
10	1.6e−1	3.3e−2	3.3e−3
100	5.0e−2	1.0e−2	1.0e−3
1000	1.6e−2	3.2e−2	3.2e−3

In fact it can be shown that in the limit the iterates approach $\beta/\sqrt{k}$. (For a general theory of sublinear convergence see [7].)

The price we pay for these slowly decreasing bounds is the requirement that U satisfy a strong rank-revealing condition. The diagonals of U must not just approximate the corresponding singular values of U but instead must approximate the smallest singular value of the corresponding leading principal submatrix. By the interleaving theorem for singular values, the latter can be smaller than the former.

To illustrate the bounds, let us consider an upper-triangular matrix whose elements are standard normal deviates and also some triangular matrices that can be obtained from them by computing factorizations. Each column of Table 2.1 contains five replications of an experiment involving an upper triangular matrix of order 25. The source of these triangular matrices is explained in the legend. Each double entry consists of the smallest singular value of the equilibrated matrix and the lower bound computed from the recursion (2.5).

The bounds are remarkably sharp, which suggests that we gave little away in their derivation. If we had used the minimum of the ratios $\beta = \sigma_k/\delta_k$ for the β_k, the bounds would not have been much worse. Owing to the sublinear convergence of the recursion, the bound would quickly drop to a little below β and then stagnate.

Turning now to the kinds of matrix used in the experiments, we note from the first column that a random normal matrix U is not good at revealing its rank, as evidenced by the small lower bound. Since the bound is nearly attained, we observe a small singular value in $D^{-1}U$.

The numbers in the second column come from the R factor in a pivoted QR decomposition of the U in column one. Such a decomposition is generally an excellent rank revealer [5], and indeed we observe that the bounds and the smallest singular value are near one.

We will discuss the experiments of the third and fourth columns after we apply our results to Gaussian elimination.

1	2	3	4
2.6e−10	3.6e−01	3.1e−02	1.1e−10
1.0e−10	1.4e−01	2.1e−02	6.7e−11
5.6e−07	3.4e−01	1.9e−02	1.1e−06
3.0e−07	1.3e−01	1.1e−02	2.8e−07
1.8e−08	3.6e−01	3.6e−02	1.7e−08
1.6e−08	1.4e−01	2.6e−02	1.5e−08
2.1e−11	3.4e−01	5.3e−02	3.2e−11
1.3e−11	1.3e−01	3.2e−02	8.1e−12
8.2e−07	3.6e−01	1.8e−02	4.8e−07
5.4e−07	1.3e−01	1.5e−02	2.7e−07

1. U consisting of standard normal deviates.
2. R from the pivoted QR decomposition of U in 1.
3. The upper triangular part of the LU decomposition of UQ, U from 1 and Q a random orthogonal matrix.
4. The upper triangular part of the LU decomposition of QU, U from 1 and Q a random orthogonal matrix.

Table 2.1: Some Experiments with Triangular Matrices

3. Gaussian elimination

In applying our results to Gaussian elimination, we will have to make use of some empirical facts about the growth of elements in course of the algorithm. For experiments and analyses concerning this important topic, see the paper by Trefethen and Schreiber [9].

Let the matrix A of order n be decomposed by Gaussian elimination with pivoting, so that

$$P^{\mathrm{T}} AQ = LDU,$$

where P and Q are permutation matrices, L is a unit lower triangular matrix, and U is unit upper triangular. When partial pivoting is used, $Q = I$ and the elements of L are less than one in magnitude. When complete pivoting is used both L and U have elements less than one in magnitude.

We will assume that $\|L\|$ and $\|U\|$ are slowly growing functions of n. For complete pivoting, we have the bound $\|L\|, \|U\| \leq n$, which is often an overestimate.

For partial pivoting, the bound continues to hold for $\|L\|$. The fact that $\|U\|$ grows slowly is related to the slow growth of elements in Gaussian elimination.

Let A_k, L_k, D_k, and U_k be the leading principal submatrices of A, L, D, and U. Let σ_k denote the smallest singular value of D_kU_k, and τ_k the smallest singular value of A_k. Define

$$\beta_k = \frac{\sigma_k}{|\delta_k|} \quad \text{and} \quad \gamma_k = \frac{\tau_k}{|\delta_k|}$$

Thus γ_k is the quality of $|\delta_k|$ as a revealer of the rank of A.

Now since $A_k = L_k(D_kU_k)$, we have $\tau_k \leq \sigma_k\|L_k\|$, or

$$\beta_k \geq \frac{\gamma_k}{\|L_k\|}.$$

By our hypothesis on the size of $\|L\|$, if U_k reveals the rank of the A_k in the sense that γ_k is near one, the matrix D_kU_k also reveals its own rank. Since these statements are true for all k, it follows from the considerations of the last section that the smallest singular value of U is near one.

The well-conditioning of L can be deduced by the same argument applied to A^{T}.

The last two columns in Table 2.1 show two aspects of these results. In the third, the columns of the matrix U of standard normal deviates was scrambled by postmultiplication by a random orthogonal matrix Q, and Gaussian elimination with partial pivoting was applied to obtain a new upper triangular matrix. This new matrix is rank-revealing and consequently scaling its rows makes its smallest singular value near one. The bounds also show that Gaussian elimination with partial pivoting is not as good at revealing rank as pivoted QR.

In the fourth column, the matrix U is replaced by QU and subjected to Gaussian elimination with partial pivoting. Here the new upper triangular matrix completely fails to reveal the rank of the original, and when it is scaled it has small singular value. The reason is easy to see. If we let $Q = LR$ be a partially pivoted LU decomposition of Q, then RU is the upper triangular matrix computed from QU. But as it turns out Q, L, and R are well conditioned, hence RU remains rank concealing.

These results have implications for the perturbation theory of the LU decomposition. Let $A + E$ have the LU decomposition $(L + F_L)(U + F_U)$ and let S be an arbitrary nonsingular diagonal matrix. The author has shown [8] that for any absolute norm $\|\cdot\|$

$$\frac{\|F_U\|}{\|U\|} \leq \kappa(L)\kappa(SU)\frac{\|E\|}{\|A\|},$$

where as usual $\kappa(X) = \|X\|\|X^{-1}\|$. The results derived here say that if U reveals the rank of A the factor $\kappa(L)\kappa(SU)$ can be made near one — i.e., U is insensitive to perturbations in A. However, it is important to keep in mind that the insensitivity is normwise. Small elements of U are generally quite sensitive, as common sense would dictate.

4. Related Decompositions

Stronger results can be obtained for the triangular matrices produced by orthogonal decompositions. For example, let A be an $m \times n$ matrix with $m \geq n$. The pivoted QR decomposition factors A in the form

$$AP = QR,$$

where P is a permutation matrix, Q is an $m \times n$ matrix with orthonormal columns and R is upper triangular with positive diagonal elements. The pivoting insures that

$$r_{kk}^2 \geq \sum_{i=k}^{j} r_{ij}^2, \qquad j = k+1, \ldots n.$$

The pivoted QR decomposition is known empirically to reveal the rank of A in the following sense [5]. If A_k denotes the matrix consisting of the first k columns of A, then r_{kk} is an approximation to the smallest singular value of A_k. Since the singular values of the leading principal matrix R_k of R are the same as the singular values of A_k, the matrix R_k will also be rank revealing. Consequently, the smallest smallest singular value of the matrix obtained by scaling so its diagonal elements are one will be near one. Thus the R factor from a pivoted QR factorization is another source of triangular systems that tend to be solved accurately.

These results possibly explain an observation of Golub on the use of Householder transformations to solve least squares problems [1]. He noted that column pivoting slightly improved the accuracy of the computed solutions. From the point of view taken here, pivoting would make the R more rank revealing and hence the QR equations for the least squares solution would be solved more accurately.

Since the R factor of AP is the Cholesky factor of $P^{\mathrm{T}}A^{\mathrm{T}}AP$, pivoted Cholesky factorization with diagonal pivoting of positive definite matrices also give rise to systems that can be solved accurately. The same can be said of the systems resulting from two-sided orthogonal decompositions like the URV and ULV decompositions.

5. Conclusions

We do not claim to have solved the mysteries of Gaussian elimination in this paper. The basic result is that if Gaussian elimination produces a rank-revealing LDU factorization with L and U of modest norm, then L and U, suitably scaled, must be well conditioned, and systems involving them will be solved accurately. That pivoted Gaussian elimination should be rank-revealing is not in itself surprising, and the all counterexamples I am aware of are obtained by starting with rank-concealing triangular matrices (cf. column four in Table 2.1). The fact — critical to our analysis — that L and U are of modest size is guaranteed for complete pivoting, but for partial pivoting what inhibits growth of the elements of U is imperfectly understood [9].

For the pivoted QR and Cholesky factorizations we need no auxiliary hypothesis about the sizes of the triangular factor. All we need to believe is that the factorizations reveal rank.

Perhaps the most unusual feature of the analysis is the nature of the recursion (2.5). It is a great leveler, eager to reduce ρ's that are greater that β and reluctant to reduce them much further. For example, if $\rho = 2\beta$, then ρ is reduced by a factor of 0.45, whereas if $\rho = \frac{1}{2}\beta$ it is reduced by a factor of only 0.89. Thus, even if a triangular matrix consistently overestimates the rank of its principal submatrices, the overestimates have only a one-time effect and do not propagate exponentially in the bounds. May we have more bounds of this nature!

References

[1] G. H. Golub. Numerical methods for solving least squares problems. *Numerische Mathematik*, 7:206–216, 1965.

[2] G. H. Golub and C. F. Van Loan. *Matrix Computations*. Johns Hopkins University Press, Baltimore, Maryland, 2nd edition, 1989.

[3] N. J. Higham. *Accuracy and Stability of Numerical Algorithms*. SIAM, Philadelphia, 1995. To appear.

[4] R. A. Horn and C. R. Johnson. *Topics in Matrix Analysis*. Cambridge University Press, Cambridge, 1991.

[5] G. W. Stewart. The efficient generation of random orthogonal matrices with an application to condition estimators. *SIAM Journal on Numerical Analysis*, 17:403–404, 1980.

[6] G. W. Stewart and J.-G. Sun. *Matrix Perturbation Theory*. Academic Press, Boston, 1990.

[7] G. W. Stwart. On sublinear convergence. Technical Report CS-TR-3534 UMIACS-TR-95-92, University of Maryland, Department of Computer Science, 1995.

[8] G. W. Stwart. On the perturbation of LU and Cholesky factors. Technical Report CS-TR-3535 UMIACS-TR-95-93, University of Maryland, Department of Computer Science, 1995.

[9] L. N. Trefethen and R. S. Schreiber. Average-case stability of Gaussian elimination. *SIAM Journal on Matrix Analysis and Applications*, 11:335–360, 1990.

[10] J. H. Wilkinson. Error analysis of direct methods of matrix inversion. *Journal of the ACM*, 8:281–330, 1961.