University of Maryland College Park

Institute for Advanced Computer Studies TR–97–02
Department of Computer Science TR–3864

On the Adjoint Matrix[*]

G. W. Stewart[†]

ABSTRACT

The adjoint A^{A} of a matrix A is the transpose of the matrix of the cofactors of the elements of A. The computation of the adjoint from its definition involves the computation of n^2 determinants of order $(n-1)$—a prohibitively expensive $O(n^4)$ process. On the other had the computation from the formula $A^{\mathrm{A}} = \det(A)A^{-1}$ breaks down when A is singular and is potentially unstable when A is ill-conditioned. In this paper we first show that the ajdoint can be perfectly conditioned, even when A is ill-conditioned. We then show that if due care is taken the adjoint can be accurately computed from the inverse, even when the latter has been inaccurately computed. In an appendix to this paper we establish a folk result on the accuracy of computed inverses.

[*] This report is available by anonymous ftp from `thales.cs.umd.edu` in the directory `pub/reports` or on the web at `http://www.cs.umd.edu/ stewart/`.

[†] Department of Computer Science and Institute for Advanced Computer Studies, University of Maryland, College Park, MD 20742. This work was supported by the National Science Foundation under grant CCR-95-03126.

On the Adjoint Matrix

G. W. Stewart

ABSTRACT

The adjoint A^{A} of a matrix A is the transpose of the matrix of the cofactors of the elements of A. The computation of the adjoint from its definition involves the computation of n^2 determinants of order $(n-1)$ — a prohibitively expensive $O(n^4)$ process. On the other had the computation from the formula $A^{\mathrm{A}} = \det(A)A^{-1}$ breaks down when A is singular and is potentially unstable when A is ill-conditioned. In this paper we first show that the ajdoint can be perfectly conditioned, even when A is ill-conditioned. We then show that if due care is taken the adjoint can be accurately computed from the inverse, even when the latter has been inaccurately computed. In an appendix to this paper we establish a folk result on the accuracy of computed inverses.

1. Introduction

Let A be a real matrix of order n and let A_{ij} denote the submatrix of A that is complementary to the element a_{ij}. Then the adjoint[1] of A is the matrix

$$A^{\mathrm{A}} = \begin{pmatrix} \det(A_{11}) & -\det(A_{12}) & \cdots & (-1)^{n+1}A_{1n} \\ -\det(A_{21}) & \det(A_{22}) & \cdots & (-1)^{n+2}A_{2n} \\ \vdots & \vdots & & \vdots \\ (-1)^{n+1}\det(A_{21}) & (-1)^{n+2}\det(A_{22}) & \cdots & (-1)^{2n}A_{1n} \end{pmatrix}^{\mathrm{T}}. \qquad (1.1)$$

In the language of determinant theory, the adjoint is the matrix whose (i,j)-element is the cofactor of the (j,i)-element of A. (For background see [6].)

The (i,j)-element of A^{A} is also the derivative $\partial\det(A)/\partial a_{ji}$, as can be seen by expanding $\det(A)$ in cofactors along the jth row of A and differentiating with respect to a_{ji}. Thus the adjoint is useful in optimization problems that involve determinants or functions of determinants.

But the adjoint is a remarkable creature, well worth studying for its own sake. It has an unusual perturbation theory, and it can be computed by a class of algorithms that at first glance appear unstable. These statements are a consequence of the well-known relation

$$A^{\mathrm{A}}A = AA^{\mathrm{A}} = \det(A)I,$$

[1] The adjoint treated here is not the adjoint of functional analysis, which corresponds to the transpose or conjugate transpose of A.

or when A is nonsingular

$$A^{\mathrm{A}} = \det(A)A^{-1}. \tag{1.2}$$

The perturbation theory is unusual because although A^{A} and A^{-1} differ only by a scalar factor the matrix A^{-1} has singularities while A^{A} is analytic — in fact, it is a multinomial in the elements of A. It turns out that multiplying by the determinant smooths out the singularities to give an elegant perturbation expansion.

The computational consequences of (1.2) are that we can, in principle, calculate the adjoint of a nonsingular matrix A by computing its inverse and determinant and multiplying. This approach has the advantage that it can be implemented with off-the-shelf software. However, if A is ill-conditioned — that is, if A is nearly singular — the inverse will be inaccurately computed. Nonetheless, in we will show that this method, properly implemented, can give an accurate adjoint, even when the inverse has been computed inaccurately.

This paper is organized as follows. In §2 we will treat the perturbation of the adjoint. In §3 we will describe a general algorithm for computing the adjoint and discuss the practicalities of its implementation. In §4 we will give some numerical examples. The paper concludes with an appendix on a folk theorem about the accuracy of computed inverses.

The singular value decomposition will play a central role in this paper. We will write it in the form

$$A = U\Sigma V^{\mathrm{T}}, \tag{1.3}$$

where U and V are orthogonal and

$$\Sigma = \operatorname{diag}(\sigma_1, \sigma_2, \ldots, \sigma_n), \qquad \sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq 0.$$

The function $\|\cdot\|$ will denote the Euclidean norm of a vector or the spectral norm of a matrix; i.e.,

$$\|A\| = \max_{\|x\|=1} \|Ax\| = \sigma_1.$$

For more on norms and the singular value decomposition see [3]

For later use we note that (1.2) implies that for nonsingular A and B

$$(AB)^{\mathrm{A}} = B^{\mathrm{A}}A^{\mathrm{A}}.$$

By continuity, this relation continues to hold when A and B are singular.

Finally, we will also use the following characterization of the singular value decomposition of A^{A}. Assuming that A is nonsingular, we have from (1.3) that $A^{-1} = V\Sigma^{-1}U^{\mathrm{T}}$.

Moreover $\det(A) = \det(U)\det(V)\det(\Sigma) = \det(U)\det(V)\sigma_1 \cdots \sigma_n$. It follows that if we set

$$\gamma_i = \prod_{k \neq i} \sigma_i \quad \text{and} \quad \Gamma = \operatorname{diag}(\gamma_1, \ldots, \gamma_n), \tag{1.4}$$

then

$$A^{\mathrm{A}} = \det(U)\det(V) V \Gamma U^{\mathrm{T}} \tag{1.5}$$

is the singular value decomposition of A^{A}. By continuity, this result also holds for singular A.

2. Perturbation theory

We have already noted that on the space of $n \times n$ matrices, the inverse has singularities while the adjoint is analytic. Since the singularity of A is equivalent to σ_n being equal to zero, we should expect perturbation bounds for the adjoint not to depend on the inverse of σ_n — as do those for the inverse. In this section we will show that the sensitivity of A^{A} depends on the inverse of σ_{n-1}. (For the perturbation of matrix inverses, see [9, Ch. III].)

We will begin with a first-order perturbation expansion for the adjoint.

Theorem 2.1. *Let A have the singular value decomposition* (1.3). *Let γ_j and Γ be defined by* (1.4) *and let*

$$\gamma_{ij} = \frac{\gamma_i}{\sigma_j} = \prod_{k \neq i,j} \sigma_k.$$

Let E be a perturbation of A and let

$$(A + E)^{\mathrm{A}} = A^{\mathrm{A}} + F.$$

Then with $\hat{E} = U^{\mathrm{T}} E V$

$$V^{\mathrm{T}} F U = \det(U)\det(V) \begin{pmatrix} \sum_{k \neq 1} \hat{\epsilon}_{kk}\gamma_{k1} & -\hat{\epsilon}_{12}\gamma_{12} & \cdots & -\hat{\epsilon}\gamma_{1n} \\ -\hat{\epsilon}_{21}\gamma_{21} & \sum_{k \neq 2} \hat{\epsilon}_{kk}\gamma_{k2} & \cdots & -\hat{\epsilon}_{2n}\gamma_{2n} \\ \vdots & \vdots & & \vdots \\ -\hat{\epsilon}_{n1}\gamma_{n1} & -\hat{\epsilon}_{n2}\gamma_{n2} & \cdots & \sum_{k \neq n} \hat{\epsilon}_{kk}\gamma_{kn} \end{pmatrix} + O(\|E\|^2). \tag{2.1}$$

Proof. We have $A + E = U(\Sigma + \hat{E})V^{\mathrm{T}}$. Since the adjoint of an orthogonal matrix is its determinant times its transpose,

$$\begin{aligned}
(A+E)^{\mathrm{A}} &= [U(\Sigma + \hat{E})V^{\mathrm{T}}]^{\mathrm{A}} \\
&= (V^{\mathrm{T}})^{\mathrm{A}}(\Sigma + \hat{E})^{\mathrm{A}}U^{\mathrm{A}} \\
&= \det(U)\det(V)V(\Sigma + \hat{E})^{\mathrm{A}}U^{\mathrm{T}} \\
&\equiv \det(U)\det(V)V(\Gamma + \hat{F})U^{\mathrm{T}} \\
&= \det(U)\det(V)(A^{\mathrm{A}} + V\hat{F}U^{\mathrm{T}}) \\
&= \det(U)\det(V)(A^{\mathrm{A}} + F).
\end{aligned}$$

For definiteness we will assume that $\det(U)\det(V) = 1$. It then follows that the first-order approximation to $V^{\mathrm{T}}FU$ is just the first-order approximation to $\hat{F}$ in the equation $(\Sigma + \hat{E})^{\mathrm{A}} = \Gamma + \hat{F}$. Because Σ is diagonal this approximation may be obtained by computing the determinants in (1.1) and throwing out higher order terms.

Specifically, consider the $(1,1)$-element of $\Sigma + \hat{F}$, which is the determinant of

$$\mathrm{diag}(\sigma_2, \sigma_3, \ldots, \sigma_n) + \hat{E}_{11},$$

where $\hat{E}_{11}$ is the submatrix complementary to ϵ_{11}. Now it is easily seen that an off-diagonal perturbation of $\mathrm{diag}(\sigma_2, \sigma_3, \ldots, \sigma_n)$ leaves its determinant unchanged; i.e., the off-diagonal elements of $\hat{E}$ have no first-order effects. It follows that the $(1,1)$-element of the adjoint is approximated by

$$\prod_{k\neq 1}(\sigma_k + \hat{\epsilon}_{kk}) \cong \gamma_1 + \sum_{k\neq 1}\epsilon_{kk}\gamma_{k1},$$

so that the $(1,1)$ element of $\hat{F}$ is $\sum_{k\neq 1}\epsilon_{kk}\gamma_{k1}$. The other diagonal elements are treated similarly.

For the off-diagonal elements of $\hat{F}$, consider the $(3,1)$-element, which is the determinant of

$$\begin{pmatrix}
0 & \sigma_2 & 0 & \cdots & 0 \\
0 & 0 & 0 & \cdots & 0 \\
0 & 0 & \sigma_4 & \cdots & 0 \\
\vdots & \vdots & \vdots & & \vdots \\
0 & 0 & 0 & \cdots & \sigma_n
\end{pmatrix} + \hat{E}_{13}.$$

The determinant of first term in this sum is unaffected by the perturbations other than in its $(2,1)$-element. Thus the first-order approximation to the $(3,1)$-element of the adjoint is

$$\det\begin{pmatrix} 0 & \sigma_2 & 0 & \cdots & 0 \\ \hat{\epsilon}_{31} & 0 & 0 & \cdots & 0 \\ 0 & 0 & \sigma_4 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & \sigma_n \end{pmatrix} = -\epsilon_{31}\gamma_{31}.$$

The other off-diagonal elements are treated similarly.[2] ■

We can turn the perturbation expansion (2.1) into a first-order perturbation bound as follows. Let $D_{\hat{E}}$ be the diagonal of $\hat{E}$ and let

$$N = \begin{pmatrix} 0 & \gamma_{12} & \cdots & \gamma_{1n} \\ \gamma_{21} & 0 & \cdots & \gamma_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ \gamma_{n1} & \gamma_{n2} & \cdots & 0 \end{pmatrix}.$$

Then the matrix in (2.1) can be written in the form

$$M + (\hat{E} - D_{\hat{E}})\circ N,$$

where M is the diagonal of the matrix in (2.1) and "$\circ$" denotes the Hadamard (componentwise) product. Since $\gamma_{n,n-1}$ is an upper bound on the γ_{ij}, we have

$$\|M\| \le \|D_{\hat{E}}\|\gamma_{n,n-1} \le (n-1)\|E\|\gamma_{n,n-1}.$$

By a generalization of a theorem of Schur [5, p. 333], the the 2-norm of $(\hat{E} - D_{\hat{E}})\circ N$ is the product of $\|\hat{E} - D_{\hat{E}}\|$ and the 2-norm of the column of largest 2-norm. Hence

$$\|(\hat{E} - D_{\hat{E}})\circ N\| \le \sqrt{n-1}\|\hat{E} - D_{\hat{E}}\|\gamma_{n-1,n}.$$

It follows that

$$\|F\| \le (n-1+\sqrt{n-1})\gamma_{n-1,n}\|E\| + O(\|E\|^2).$$

[2] Another approach is to assume A is nonsingular and write $(\Gamma + \hat{F})(\Sigma + \hat{E}) = \det(\Sigma + \hat{E})I$. Setting $\delta = \det(\Sigma + E) - \det\Sigma$, we can show that

$$\hat{F} = \frac{\delta\Gamma}{\det(\Sigma)} - \frac{\Gamma^{\mathrm{A}}\hat{E}\Gamma^{\mathrm{A}}}{\det(\Sigma)} - \frac{FE\Gamma}{\det(\Sigma)}.$$

Replacing δ with a first-order expansion and discarding $FE\Gamma/\det(\Sigma)$ gives (after some manipulation) the first-order approximation to $\hat{F}$.

If we divide by $\|A^{\mathrm{A}}\| = \gamma_n$ and note that $\|A\| = \sigma_1$, we get

$$\frac{\|F\|}{\|A\|} \leq (n-1+\sqrt{n-1})\frac{\sigma_1\gamma_{n-1,n}}{\gamma_n}\frac{\|E\|}{\|A\|} + O(\|E\|^2).$$

Hence

$$\frac{\|F\|}{\|A\|} \leq (n-1+\sqrt{n-1})\frac{\sigma_1}{\sigma_{n-1}}\frac{\|E\|}{\|A\|} + O(\|E\|^2), \tag{2.2}$$

which is our first-order perturbation bound.

There are two comments to be made on this bound. First, it shows that the normwise relative perturbation of the adjoint depends on the ratio σ_1/σ_{n-1}. This should be contrasted with the ratio $\sigma_1/\sigma_n = \|A\|\|A^{-1}\|$ for perturbations of the inverse. Thus A can be arbitrarily ill conditioned or even singular while its adjoint is well conditioned.

Second, the factor $n-1$ in the bound is necessary. For let $A = I$ and $E = \epsilon I$. Then $F \cong (n-1)\epsilon I$, whose norm is $n-1$. It should be stressed that the factor was derived under the assumption that all the γ_{ij} are equal, and in practice it is likely to be an overestimate.

3. Computing the adjoint

As we pointed out in the introduction to this paper, the adjoint of a nonsingular matrix A can be computed by the following algorithm.

1. Compute A^{-1}.
2. Compute $\det(A)$.
3. $A^{\mathrm{A}} = \det(A)A^{-1}$.

Mathematically, this algorithm works for any nonsingular matrix A. Numerically, if A is ill-conditioned — that is, if σ_1/σ_n is large — the matrix A^{-1} will be computed inaccurately. Since the adjoint itself is insensitive to the size of σ_n, we are in danger of computing an inaccurate solution to a well conditioned problem. The remarkable fact is that if A^{-1} is computed with some care, the multiplication by $\det(A)$ wipes out the error.

Specifically, suppose we can factor A in the form

$$A = XDY,$$

where X and Y are well conditioned and D is diagonal. The singular value decomposition is such a factorization; however, as we shall see there are others. Given such a factorization, we have mathematically $A^{\mathrm{A}} = \det(X)\det(D)\det(Y)(X^{-1}D^{-1}Y^{-1})$. Algorithmically we proceed as follows.

```
1.  Factor $A = XDY$, where $X$ and $Y$ are well conditioned
        and $D$ is diagonal.
2.  Compute $U = X^{-1}$.
3.  Compute $V = D^{-1}U$.
4.  Compute $W = Y^{-1}X$.
5.  $A^{\mathrm{A}} = \det(X)\det(D)\det(Y)W$.
```

Two facts account for the success of this algorithm. First, the algorithms that compute the XDY factorizations to be treated later produce a computed factorization that satisfies

$$XDY = A + E \tag{3.1}$$

where $\|A\|/\|E\|$ is of the order of the rounding unit. Second, if the operations in steps 2–4 are properly implemented, then the computed W — call it $\tilde{W}$ — satisfies

$$\tilde{W} = W + F, \tag{3.2}$$

where $\|F\|/\|W\|$ is of the order of the rounding unit. In the parlance of rounding-error analysis, the XDY factorization is computed stably while the matrix W is computed accurately.

Combining these facts we see that the computed adjoint is near the adjoint of $A + E$. If A^{A} is well conditioned, then it is near the adjoint of $(A + E)$, which is near the computed adjoint. Thus well conditioned adjoints are computed accurately. The accuracy of the adjoint deteriorates as it becomes increasingly ill-conditioned, but the deterioration is of the same order of magnitude as that caused by the small perturbation E in A.

As we have said, the existence of E in (3.1) is a property of the algorithm used to compute an XDY factorization. The existence of F in (3.2) is a folk theorem in matrix computations, which is useful in explaining the behavior of computed inverses. Since this result has not, to my knowledge, been given a precise quantitative form, we state and prove it in the appendix.

Before turning to specific XDY factorizations, an observation on singular matrices is in order. Mathematically, a singular matrix must result in a matrix D in which at least one diagonal element is zero — in which case the algorithm cannot proceed. In the presence of rounding error this eventuality is unlikely. If, however, D does have a zero diagonal element, the cure is to perturb it slightly and proceed. Since X and Y are well conditioned, this perturbation will be equivalent to a small perturbation in E, and the algorithm will behave as described above.

To apply our algorithm we need a factorization XDY for which X and Y are well conditioned. In addition, the matrices X and Y must be such as to make the computations in steps 2, 4, and 5 efficient. The singular value decomposition is such a factorization. However, there are three other candidates: the LU decomposition with complete pivoting, the pivoted QR decomposition, and the pivoted QLP decomposition. We treat each of these four decompositions in turn.

- **The singular value decomposition.** If $A = U\Sigma V^{\mathrm{T}}$ is the singular value decomposition of A, then we may take $X = U$, $Y = V^{\mathrm{T}}$, and $D = \Sigma$ to get

$$A^{\mathrm{A}} = \det(U)\det(V)\det(\Sigma)V\Sigma^{-1}U^{\mathrm{T}} \tag{3.3}$$

The computation of $U\Sigma^{-1}V^{\mathrm{T}}$ is straightforward. Since U and V are orthogonal, their determinants are ±1; but it may be difficult to determine the sign. If the singular value decomposition is calculated by reduction to bidiagonal form followed by the QR algorithm, then U and V are products of Householder transformations, whose determinant is -1, and plane rotations, whose determinant is 1, followed by row or column scaling by a factors of -1 to make the σ_i positive. Thus in principle it is possible to calculate the determinants during the course of the algorithm by keeping track of the transformations. Unfortunately, off-the-shelf software (e.g., from LINPACK [2] or LAPACK [1]) does not do this.

- **The completely pivoted LU decomposition.** When Gaussian elimination with complete pivoting is used to compute an LU decomposition, we obtain a factorization of the form

$$A = \Pi_{\mathrm{R}}LDU\Pi_{\mathrm{C}},$$

where L and U are unit lower and upper triangular and Π_{R} and Π_{C} are permutation matrices (the subscripts stand for row and column permutation). Taking $X = \Pi_{\mathrm{R}}L$ and $Y = U\Pi_{\mathrm{C}}$, we have

$$A^{\mathrm{A}} = \det(\Pi_{\mathrm{R}})\det(\Pi_{\mathrm{C}})\det(D)\Pi_{\mathrm{C}}^{\mathrm{T}}U^{-1}D^{-1}L^{-1}\Pi_{\mathrm{R}}^{\mathrm{T}}.$$

Because L and U are triangular the computation of $\Pi_{\mathrm{C}}^{\mathrm{T}}U^{-1}\Gamma L^{-1}\Pi_{\mathrm{R}}$ is straightforward. Moreover,

$$\det(\Pi_{\mathrm{R}}) = (-1)^{\text{number of row interchanges}}$$

and

$$\det(\Pi_{\mathrm{C}}) = (-1)^{\text{number of column interchanges}}$$

and thus can be calculated from the pivot information that must be returned with the decomposition.

We have mentioned above, that the factors X and Y in our formula should be well conditioned. For the singular value decomposition, their orthogonality guaranteed their well conditioning. In the present case, the complete pivoting strategy tends to make L and U well conditioned. (The reasons, which have to do with how well the diagonal matrix D reflects the condition of A, are imperfectly understood. For more, see [8]). Additional security is provided by the fact that triangular systems are often solved more accurately than their condition warrants (See [4, Ch. 8] and the comments at the end of this paper.) If complete security is desired, a condition estimator [4, Ch. 14] can be used to check the status of L and U.

- **The pivoted QR decomposition.** The pivoted QR decomposition factors A in the form

$$A = QDR\Pi_{\mathrm{C}},$$

where Q is orthogonal, R is unit upper triangular, and Π_{C} is is a permutation. Setting $X = Q$ and $Y = R\Pi_{\mathrm{C}}$, we have

$$A^{\mathrm{A}} = \det(Q)\det(\Pi_{\mathrm{C}})\det(D)\Pi_{\mathrm{C}}^{\mathrm{T}}R^{-1}D^{-1}Q^{\mathrm{T}}.$$

Once again it is easy to calculate $\Pi_{\mathrm{C}}^{\mathrm{T}}R^{-1}D^{-1}Q^{\mathrm{T}}$. The usual algorithm uses $n-1$ Householder transformations to triangularize A, so that $\det(Q) = (-1)^{n-1}$. On the other hand,

$$\det(\Pi_{\mathrm{C}}) = (-1)^{\text{number of column interchanges}},$$

which can be computed from the output of the algorithm.

The algorithm generally produces a well-conditioned R, although there is a well-know counterexample. As with the the completely pivoted LU decomposition, we can use a condition estimator to check the status of R.

- **The pivoted QLP decomposition.** This decomposition, which can be computed from by two applications of orthogonal triangularization with column pivoting, can be written in the form

$$A = \Pi_{\mathrm{R}}QLDP\Pi_{\mathrm{C}},$$

where P and Q are orthogonal, L is unit lower triangular, and Π_{R} and Π_{C} are permutation matrices [7]. If we set $X = \Pi_{\mathrm{R}}QL$ and $Y = P\Pi_{\mathrm{C}}$, then from our formula

$$A^{\mathrm{A}} = \det(Q)\det(P)\det(\Pi_{\mathrm{R}})\det(\Pi_{\mathrm{C}})\det(D)\Pi_{\mathrm{C}}^{\mathrm{T}}P^{\mathrm{T}}L^{-1}D^{-1}Q^{\mathrm{T}}\Pi_{\mathrm{R}}^{\mathrm{T}}.$$

The calculation of $\Pi_{\mathrm{C}}^{\mathrm{T}}P^{\mathrm{T}}L^{-1}D^{-1}Q^{\mathrm{T}}\Pi_{\mathrm{R}}^{\mathrm{T}}$ is routine. The determinants of Q and P are $(-1)^{n-1}$, and the determinants of P_{R} and P_{C} can be determined from the interchanges made in the course of the algorithm.

There remains the question of which decomposition to use in practice. The singular value decomposition is obviously the safest since X and Y are perfectly conditioned. Unfortunately, off-the-shelf software does not provide the wherewithal to implement the formula. Of the alternatives, Gaussian elimination with complete pivoting is the cheapest. However, the standard packages do not have a complete pivoting option. The pivoted QR decomposition can be implemented with off-the-shelf software, and in all but contrived examples the triangular factor will be well conditioned. The pivoted QLP decomposition is relatively new, but experience suggests that it is close to the singular value decomposition in safety, and, like the pivoted QR decomposition, the formula can be implemented with off-the-shelf software.

4. Numerical examples

In this section we will give some numerical examples to show that our algorithms can actually compute an accurate adjoint from an inaccurate inverse. They were performed in MATLAB with a rounding unit of about 10^{-16}.

The first example was constructed as follows. Let A_0 be a matrix of standard normal deviates of order 50, normalized so that $\|A_0\| = 1$. In the singular value decomposition $A_0 = U\Sigma_0 V^{\mathrm{T}}$, set $\sigma_n = 10^{-15}$ and $\sigma_{n-1} = 10^{-1}$ to get Σ and set $A = U\Sigma V^{\mathrm{T}}$. Thus A^{-1} has a condition number $\sigma_1/\sigma_n = 10^{15}$, and we can expect a computed inverse of A to be almost completely inaccurate. The condition number of A^{A}, on the other hand, is $\sigma_1/\sigma_{n-1} = 10$, so we should be able to compute it accurately. The following table gives the results of some computations with this kind of matrix, repeated five times over.

$a_{11}^{(\mathrm{A})}$	SVD	QRD	LUD	QRSVD
7.5e−25	8.0e−15	5.0e−15	3.4e−15	4.3e−02
−1.5e−24	4.8e−16	2.7e−15	2.7e−15	6.2e−02
1.2e−25	2.8e−14	5.3e−15	2.5e−15	5.1e−02
−1.2e−24	2.1e−15	3.7e−15	6.0e−15	5.7e−02
−2.5e−25	4.4e−15	7.5e−15	4.0e−15	1.7e−02

The first column shows the $(1,1)$-element of the adjoint. The second shows the relative error in the approximation to that element computed from the singular value decomposition. The third column contains the relative normwise error in the adjoint computed from the pivoted QR decomposition (the QRD adjoint) compared to the SVD adjoint. The fourth column contains the same for the adjoint computed from a partially-pivoted LU decomposition. We will discuss the fifth column a little later.

The small size of the first column serves to remind us that in computing determinants we must be careful to avoid overflows and underflows. In particular, for any scalar μ, $(\mu A)^{\mathrm{A}} = \mu^{n-1}A^{\mathrm{A}}$. Hence, minor rescalings of A are magnified greatly in the adjoint.

It is an $O(n^4)$ process to compute the adjoint directly from its determinantal definition (1.1) — something too time consuming for a matrix of order 50. We have therefore let the adjoint computed from the singular value decomposition stand for the actual adjoint in assessing the QRD and LUD adjoints. The fact that the the second column shows that $(1,1)$-element is almost fully accurate gives us some confidence in this procedure. The fact that the QRD adjoint tracks the SVD adjoint so well (column 3) gives us additional confidence. Thus we conclude that the QRD adjoint, which is easy to compute, gives accurate results.

Although we have treated the completely pivoted LU decomposition above, the fourth column of the table shows that for this class of problems Gaussian elimination with partial pivoting works just as well.

The fifth column column illustrates a subtle point in implementing these algorithms. In the SVD formula

$$A^{\mathrm{A}} = \det(U)\det(V)\det(\Sigma)V\Sigma^{-1}U^{\mathrm{T}},$$

the determinant $\det(\Sigma)$ must be the determinant of Σ computed to low relative error. It might be thought that we could substitute $\pm\det(A)$, where $\det(A)$ is computed from, say the QR decomposition. But this determinant will in general be different from $\det(\Sigma)$, the relative error approaching one as σ_n approaches the rounding unit. The fifth column gives the relative error in the SVD adjoint when this substitution is made. There is practically no accuracy.

To illustrate our perturbation theory, the same example was run with $\sigma_{n-1} = 10^{-5}$. Thus the condition number of the adjoint is 10^5, and we should expect a loss of four or five figures in our computed values. The following table shows that this is exactly what happens.

$a_{11}^{(\mathrm{A})}$	SVD	QRD	LUD	QRSVD
2.8e−29	2.9e−11	5.3e−12	4.9e−12	4.9e−02
5.9e−30	2.2e−11	1.8e−11	9.3e−12	7.1e−02
−3.3e−27	1.6e−12	2.4e−12	8.1e−12	1.7e−02
−2.5e−28	1.1e−12	5.9e−12	3.2e−12	2.2e−02
2.4e−30	2.1e−11	1.1e−11	1.0e−11	1.1e−01

Although we cannot expect our algorithms to give fully accurate results in this case, at least the deterioration is no worse that is warranted by condition of the problems.

It is worth noting that I have been unable to construct counterexamples to make the QRD adjoint and the LUD adjoint fail. We will return to this point at the end of the appendix to this paper

A. Appendix: The accuracy of computed inverses

Let $A = XDY$, where D is diagonal and X and Y are presumed to be well conditioned. In this appendix we will be concerned with the computation of $A^{-1} = Y^{-1}D^{-1}X^{-1}$ in floating-point arithmetic with rounding unit ϵ_{M}. Without loss of generality, we may assume that

$$\|X\| = \|S\| = \|Y\| = 1.$$

We will assume that the computations are arranged so that products like $W = Y^{-1}V$ have a forward error analysis of the form

$$\tilde{W} = \mathrm{fl}(Y^{-1}V) = U + H,$$

where

$$\frac{\|\tilde{W} - W\|}{\|W\|} \leq \alpha\|Y^{-1}\|\epsilon_{\mathrm{M}} = \|Y^{-1}\|\epsilon. \tag{A.1}$$

Here α is a constant that depends on the order of the matrices and the details of the algorithm, but does not depend on Y and V. Since we are concerned with the broad outlines of the analysis, not specific bounds, we will introduce an adjustable constant ϵ into which constants like α may be merged — as in (A.1) above. In particular, if $\|\tilde{W} - W\|/\|W\|$ is small, we can replace it with $\|\tilde{W} - W\|/\|\tilde{W}\|$ by adjusting ϵ slightly in (A.1).

The first step is to compute $U = X^{-1}$. If $\tilde{U} = \mathrm{fl}(X^{-1})$, then

$$\tilde{U} = X^{-1} + F, \qquad \frac{\|F\|}{\|U\|} \leq \|X^{-1}\|\epsilon. \tag{A.2}$$

Now consider $V = D^{-1}U$. What we compute is

$$\tilde{V} = \mathrm{fl}(D^{-1}\tilde{U}) = D^{-1}\tilde{U} + G.$$

Since D is diagonal the elements of $\tilde{V}$ are relative perturbations of those of $D^{-1}\tilde{U}$ of order ϵ_{M}. It follows that

$$\frac{\|G\|}{\|\tilde{V}\|} \leq \epsilon. \tag{A.3}$$

Since

$$\tilde{V} = D^{-1}X + D^{-1}F + G = V + D^{-1}F + G, \tag{A.4}$$

we have

$$\|\tilde{V} - V\| \le (\|S^{-1}\|\|F\| + \|G\|)\epsilon.$$

Since $VX = S^{-1}$,

$$\|V\| \ge \|S^{-1}\|.$$

Hence by (A.2)

$$\|\tilde{V} - V\| \le (\|F\|\|X^{-1}\| + \|\tilde{V}\|)\epsilon \le (1 + \|X^{-1}\|)\max\{\|V\|, \|\tilde{V}\|\}\epsilon.$$

By adjusting ϵ if necessary, we get

$$\frac{\|\tilde{V} - V\|}{\|\tilde{V}\|} \le (1 + \|X^{-1}\|)\epsilon.$$

Finally,

$$\tilde{W} = \mathrm{fl}(Y^{-1}\tilde{V}) = Y^{-1}\tilde{V} + H, \qquad \frac{\|H\|}{\|\tilde{W}\|} \le \|Y^{-1}\|\epsilon.$$

Hence

$$\begin{aligned}\|\tilde{W} - W\| &\le \|Y^{-1}\|\|\tilde{V} - V\| + \|H\| \\ &\le [\|Y^{-1}\|(1 + \|X^{-1}\|)\|\tilde{V}\| + \|Y^{-1}\|\|\tilde{W}\|]\epsilon \\ &\le [\|Y^{-1}\|(1 + \|X^{-1}\|)\|W\| + \|Y^{-1}\|\|\tilde{W}\|]\epsilon.\end{aligned}$$

It follows that

$$\frac{\|\tilde{W} - W\|}{\|W\|} \le \|Y^{-1}\|(2 + \|X^{-1}\|)\epsilon$$

Reintroducing the norms of X and Y, we have the following theorem.

Theorem A.1. *Let $A = XDY$, where D is diagonal. Let $A^{-1} = Y^{-1}D^{-1}X^{-1}$ be computed in floating-point arithmetic with rounding unit ϵ_{M} in such a way that each step has a forward error analysis analogous to* (A.1). *Then if $\tilde{A}^{-1}$ denotes the computed value of A^{-1},*

$$\frac{\|\tilde{A}^{-1} - A^{-1}\|}{\|A^{-1}\|} \le \gamma\kappa(Y)[2 + \kappa(X)]\epsilon_{\mathrm{M}},$$

where γ is a constant independent of X, S, and Y and

$$\kappa(X) = \|X\|\|X^{-1}\| \quad \text{and} \quad \kappa(Y) = \|Y\|\|Y^{-1}\|.$$

The key feature of this result is that the condition number $\kappa(S)$ does not appear in the final bound. This implies that as long as the ill-conditioning of A is confined to S its inverse will be accurately computed. Decompositions that have this property are called rank-revealing decompositions, and it is no coincidence that the decompositions treated §3 are regarded as generally rank revealing.

The diagonality of S is essential to the above analysis. Without it the bound (A.3) would depend on $\|S^{-1}\|$. However, there is a subtle point that is easy to miss: $\|V\|$ must be adequately large. For otherwise the error $D^{-1}F$ in (A.4) could overwhelm V. Fortunately, we have $\|V\| \geq \|S^{-1}\|$ from the relation $VX = S^{-1}$.

This point explains why we cannot use the above analysis to claim that solutions of $(XDY)b = d$ are computed accurately. For even if we compute $u = X^{-1}b$ accurately in a normwise sense, we cannot guarantee that $\tilde{v}^{-1}\tilde{u}$ is normwise accurate unless $v = S^{-1}X^{-1}b$ is sufficiently large. Unfortunately X^{-1} is now trapped between S^{-1} and b, and we cannot derive a bound like $\|v\| \geq \|S^{-1}\|$. We can only hypothesize it.

Finally, we point out that assuming that the condition numbers of X and Y are small essentially says that $\tilde{U}$ and $\tilde{W}$ are computed accurately [see (A.1) and (A.2)]; and, in fact, this assumption of accuracy is sufficient to establish the accuracy of the computed inverse. In particular, it is known that triangular systems — even ill-conditioned ones — are often solved with high accuracy [4, Ch. 8]. The fact the X- and Y-factors in the decompositions of §§3–4 are either well-conditioned (i.e., orthogonal) or triangular, may account for the difficulty the author had in constructing counterexamples.

References

[1] E. Anderson, Z. Bai, C. Bischof, J. Demmel, J. Dongarra, J. Du Croz, A. Greenbaum, S. Hammarling, A. McKenney, S. Ostrouchov, and D. Sorensen. *LAPACK Users' Guide.* SIAM, Philadelphia, second edition, 1995.

[2] J. J. Dongarra, J. R. Bunch, C. B. Moler, and G. W. Stewart. *LINPACK User's Guide.* SIAM, Philadelphia, 1979.

[3] G. H. Golub and C. F. Van Loan. *Matrix Computations.* Johns Hopkins University Press, Baltimore, Maryland, 3rd edition, 1996.

[4] N. J. Higham. *Accuracy and Stability of Numerical Algorithms.* SIAM, Philadelphia, 1996.

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

[6] P. Lancaster and M. Tismenetski. *The Theory of Matrices.* Academic Press, New York, 1985.

[7] G. W. Stewart. The qlp approximation to the singular value decomposition. Technical Report TR-3840, Department of Computer Science, University of Maryland, 1997.

[8] G. W. Stewart. The triangular matrices of Gaussian elimination and related decompositions. *IMA Journal on Numerical Analysis*, 17:7–16, 1997.

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