IMA Journal of Numerical Analysis (1982) **2**, 303–323

Numerical Solution of the Stable, Non-negative Definite Lyapunov Equation

S. J. HAMMARLING

Division of Information Technology and Computing, National Physical Laboratory, Teddington, Middlesex TW11 0LW

[Received 30 November 1981]

We discuss the numerical solution of the Lyapunov equation

$$A^H X + XA = -C, \qquad C = C^H$$

and propose a variant of the Bartels–Stewart algorithm that allows the Cholesky factor of X to be found, without first finding X, when A is stable and C is non-negative definite.

1. Introduction

LET A be a given n by n matrix with eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$, let A^T denote the transpose of A and A^H the complex conjugate of A^T and let C be a given Hermitian matrix. Then the equation

$$A^H X + XA = -C, \qquad C = C^H \tag{1.1}$$

is called the continuous-time Lyapunov equation and is of interest in a number of areas of control theory such as optimal control and stability (Barnett, 1975; Barnett & Storey, 1968).

The equation has a unique Hermitian solution, X, if and only if $\lambda_i + \bar{\lambda}_j \neq 0$ for all i and j (Barnett, 1975). In particular if every λ_i has a negative real part, so that A is stable, and if C is non-negative definite then X is also non-negative definite (Snyders & Zakai, 1970; Givens, 1961; Gantmacher, 1959). In this case, since X is non-negative definite, it can be factorized as

$$X = U^H U, \tag{1.2}$$

where U is an upper triangular matrix with real non-negative diagonal elements, this being the Cholesky factorization of X (Wilkinson, 1965).

The Cholesky factors of X can readily be used in place of X and in many situations they will be far more useful than X itself. Furthermore, when X is non-singular, if $\|X\|_2$ denotes the spectral norm of X and $c_2(X)$ denotes the condition number of X with respect to inversion given by

$$c_2(X) = \|X\|_2 \|X^{-1}\|_2, \tag{1.3}$$

then we have the well-known result that

$$c_2(X) = c_2^2(U) \tag{1.4}$$

and hence $\mathbf{X}$ may be considerably more ill-conditioned with respect to inversion than $\mathbf{U}$. If we could solve Equation (1.1) directly for $\mathbf{U}$, then by using $\mathbf{U}$ in place of $\mathbf{X}$ we might hope to avoid the loss of accuracy associated with the squaring of the condition number in Equation (1.4). Whether or not this hope can be realized will depend upon the application.

A number of methods for solving the Lyapunov equation have appeared in the literature (Rothschild & Jameson, 1970; Hagander, 1972; Pace & Barnett, 1972; Belanger & McGillivray, 1976; Hoskins *et al.*, 1977; Galeone & Peluso, 1979; Sima, 1980). One of the most effective methods from a numerical point of view is an algorithm due to Bartels & Stewart (1972) (see also Belanger & McGillivray, 1976, and Sima, 1980). The Bartels–Stewart algorithm can be used to solve the more general Sylvester equation

$$\mathbf{BX}+\mathbf{XA} = -\mathbf{C}, \tag{1.5}$$

where $\mathbf{C}$ is not necessarily Hermitian, and their method for this equation has been further refined by Golub *et al.* (1979). Here we shall only consider the special case of Equation (1.1), this case having also been discussed in the Bartels & Stewart paper.

We first describe the Bartels–Stewart algorithm, then we discuss the non-negative definite case and propose a variant of the algorithm that allows the Cholesky factor $\mathbf{U}$ to be obtained directly without first finding $\mathbf{X}$. The case where $\mathbf{A}$ is normal, the Kronecker product form of the Lyapunov equation and the sensitivity to perturbations of the Lyapunov equation are discussed. Finally, mention is made of the discrete-time Lyapunov equation and of the implicit Lyapunov equation.

2. The Bartels–Stewart Algorithm

The Schur factorization of a square matrix $\mathbf{A}$ is given by

$$\mathbf{A} = \mathbf{QSQ}^{\mathrm{H}}, \tag{2.1}$$

where $\mathbf{Q}$ is unitary and $\mathbf{S}$ is upper triangular (Wilkinson, 1965). Since $\mathbf{S}$ is similar to $\mathbf{A}$ the diagonal elements of $\mathbf{S}$ are the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$. This factorization is important because it can be obtained by numerically stable methods; first $\mathbf{A}$ is reduced to upper Hessenberg form by means of Householder transformations and then the QR-algorithm is applied to reduce the Hessenberg form to $\mathbf{S}$ (Wilkinson, 1965), the transformation matrices being accumulated at each step to give $\mathbf{Q}$.

If we now put

$$\tilde{\mathbf{C}} = \mathbf{Q}^{\mathrm{H}}\mathbf{CQ} \quad \text{and} \quad \tilde{\mathbf{X}} = \mathbf{Q}^{\mathrm{H}}\mathbf{XQ} \tag{2.2}$$

then Equation (1.1) becomes

$$\mathbf{S}^{\mathrm{H}}\tilde{\mathbf{X}}+\tilde{\mathbf{X}}\mathbf{S} = -\tilde{\mathbf{C}} \tag{2.3}$$

and these equations can readily be solved by a process of forward substitution, as we now demonstrate. Partition $\mathbf{S}$, $\tilde{\mathbf{C}}$ and $\tilde{\mathbf{X}}$ as

$$\mathbf{S} = \begin{pmatrix} \lambda_1 & \mathbf{s}^{\mathrm{H}} \\ 0 & \mathbf{S}_1 \end{pmatrix}, \quad \tilde{\mathbf{C}} = \begin{pmatrix} \tilde{c}_{11} & \tilde{\mathbf{c}}^{\mathrm{H}} \\ \tilde{\mathbf{c}} & \mathbf{C}_1 \end{pmatrix}, \quad \tilde{\mathbf{X}} = \begin{pmatrix} \tilde{x}_{11} & \tilde{\mathbf{x}}^{\mathrm{H}} \\ \tilde{\mathbf{x}} & \mathbf{X}_1 \end{pmatrix}, \tag{2.4}$$

where $\tilde{c}_{11}$ and $\tilde{x}_{11}$ are scalars and $\mathbf{s}$, $\tilde{\mathbf{c}}$ and $\tilde{\mathbf{x}}$ are $(n-1)$ element vectors. Then Equation (2.3) gives the three equations

$$(\lambda_1+\bar{\lambda}_1)\tilde{x}_{11} = -\tilde{c}_{11}$$

$$\tilde{x}_{11}\mathbf{s}+(\mathbf{S}_1^{\mathrm{H}}+\lambda_1\mathbf{I})\tilde{\mathbf{x}} = -\tilde{\mathbf{c}}$$

$$\mathbf{s}\tilde{\mathbf{x}}^{\mathrm{H}}+\mathbf{S}_1^{\mathrm{H}}\mathbf{X}_1+\tilde{\mathbf{x}}\mathbf{s}^{\mathrm{H}}+\mathbf{X}_1\mathbf{S}_1 = -\mathbf{C}_1$$

and hence

$$\tilde{x}_{11} = -\tilde{c}_{11}/(\lambda_1+\bar{\lambda}_1) \tag{2.5}$$

$$(\mathbf{S}_1^{\mathrm{H}}+\lambda_1\mathbf{I})\tilde{\mathbf{x}} = -\tilde{\mathbf{c}}-\tilde{x}_{11}\mathbf{s} \tag{2.6}$$

$$\mathbf{S}_1^{\mathrm{H}}\mathbf{X}_1+\mathbf{X}_1\mathbf{S}_1 = -\mathbf{C}_1-(\mathbf{s}\tilde{\mathbf{x}}^{\mathrm{H}}+\tilde{\mathbf{x}}\mathbf{s}^{\mathrm{H}}). \tag{2.7}$$

Once $\tilde{x}_{11}$ has been found from Equation (2.5), Equation (2.6) can be solved, by forward substitution, for $\tilde{\mathbf{x}}$ and then Equation (2.7) is of the same form as (2.3), but of order $(n-1)$. The condition $\lambda_i+\bar{\lambda}_j \neq 0$ ensures that Equations (2.5) and (2.6) have unique solutions.

It should be noted that although the matrix $\hat{\mathbf{C}}$ given by

$$\hat{\mathbf{C}} = \mathbf{C}_1+(\mathbf{s}\tilde{\mathbf{x}}^{\mathrm{H}}+\tilde{\mathbf{x}}\mathbf{s}^{\mathrm{H}})$$

is Hermitian, when $\tilde{\mathbf{C}}$ is positive definite the matrix $\hat{\mathbf{C}}$ is not necessarily positive definite so that $\mathbf{X}_1$ is positive definite by virtue of being a principal minor of $\tilde{\mathbf{X}}$ and not by virtue of being a solution of Equation (2.7). Such an example is given by

$$\mathbf{S} = \begin{pmatrix} -0\cdot5 & 1 & 1 \\ 0 & -0\cdot5 & -2 \\ 0 & 0 & -0\cdot5 \end{pmatrix}, \qquad \tilde{\mathbf{C}} = \mathbf{I}$$

for which we find that

$$\tilde{\mathbf{x}}_{11} = 1, \qquad \tilde{\mathbf{x}} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

and

$$\hat{\mathbf{C}} = \begin{pmatrix} 3 & 0 \\ 0 & -1 \end{pmatrix}$$

so that although $\tilde{\mathbf{C}}$ is positive definite $\hat{\mathbf{C}}$ is indefinite.

It is aesthetically displeasing that positive definiteness of $\hat{\mathbf{C}}$ is not guaranteed. We shall show in Section 5 that this deficiency is not present in the alternative approach of finding the Cholesky factor of $\mathbf{X}$.

When $\mathbf{A}$ and $\mathbf{C}$ are both real then $\mathbf{X}$ will also be real and it is possible to work entirely in real arithmetic by replacing (2.1) with the real Schur factorization

$$\mathbf{A} = \mathbf{Q}\mathbf{S}\mathbf{Q}^{\mathrm{T}}, \tag{2.8}$$

where now $\mathbf{Q}$ is orthogonal and $\mathbf{S}$ is block upper triangular with 1 by 1 and 2 by 2 blocks, the eigenvalues of a 2 by 2 block being a complex conjugate pair (Wilkinson, 1965). For this factorization Equations (2.5), (2.6) and (2.7) become

$$s_{11}^{\mathrm{T}}\tilde{x}_{11}+\tilde{x}_{11}s_{11} = -\tilde{c}_{11} \tag{2.9}$$

$$\mathbf{S}_1^{\mathrm{T}}\tilde{\mathbf{x}}+\tilde{\mathbf{x}}s_{11} = -\tilde{\mathbf{c}}-\mathbf{s}\tilde{x}_{11} \tag{2.10}$$

$$\mathbf{S}_1^{\mathrm{T}}\mathbf{X}_1+\mathbf{X}_1\mathbf{S}_1 = -\mathbf{C}_1-(\mathbf{s}\tilde{\mathbf{x}}^{\mathrm{T}}+\tilde{\mathbf{x}}\mathbf{s}^{\mathrm{T}}), \tag{2.11}$$

where s_{11}, $\tilde{x}_{11}$ and $\tilde{c}_{11}$ are either scalars or 2 by 2 matrices and $\mathbf{s}$, $\tilde{\mathbf{x}}$ and $\tilde{\mathbf{c}}$ are either vectors or matrices with two columns. In the 2 by 2 case Equation (2.9) defines three equations in the unknown elements of $\tilde{x}_{11}$ and Equation (2.10) can then be solved by forward substitution, a row of $\tilde{\mathbf{x}}$ being found at each step.

3. The Non-negative Definite Case

The case where $\mathbf{X}$ is non-negative definite generally arises when $\mathbf{A}$ is stable, that is $\mathbf{A}$ has eigenvalues with negative real parts, and when $\mathbf{C}$ is of the form

$$\mathbf{C} = \mathbf{B}^{\mathrm{H}}\mathbf{B} \tag{3.1}$$

where $\mathbf{B}$ is an m by n matrix. For example, when $\mathbf{A}$ is stable, the "reachability Grammian" given by

$$\mathbf{X} = \int_0^{\infty} (\mathbf{B}\, e^{t\mathbf{A}})^{\mathrm{H}}(\mathbf{B}\, e^{t\mathbf{A}})\, dt$$

satisfies the Lyapunov equation (see for instance Snyders & Zakai, 1970)

$$\mathbf{A}^{\mathrm{H}}\mathbf{H}+\mathbf{XA} = -\mathbf{B}^{\mathrm{H}}\mathbf{B}. \tag{3.2}$$

If we can find the Cholesky factor $\mathbf{U}$ via $\mathbf{B}$ rather than $\mathbf{B}^{\mathrm{H}}\mathbf{B}$ then we have added incentive to use this approach, since once again we should then hope to avoid the loss of accuracy associated with squaring the condition number. For example, consider the case where

$$\mathbf{A} = -\mathbf{I} \quad \text{and} \quad \mathbf{B} = \begin{pmatrix} 1 & 1 \\ 0 & \varepsilon \end{pmatrix}. \tag{3.3}$$

Since

$$\mathbf{B}^{\mathrm{H}}\mathbf{B} = \begin{pmatrix} 1 & 1 \\ 1 & 1+\varepsilon^2 \end{pmatrix}$$

we see that

$$\mathbf{X} = \tfrac{1}{2}\mathbf{B}^{\mathrm{H}}\mathbf{B} = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1+\varepsilon^2 \end{pmatrix}. \tag{3.4}$$

Because $\mathbf{B}$ is upper triangular we also have that

$$\mathbf{U} = (\tfrac{1}{2})^{\frac{1}{2}}\mathbf{B} = \left(\frac{1}{2}\right)^{\frac{1}{2}}\begin{pmatrix} 1 & 1 \\ 0 & \varepsilon \end{pmatrix}. \tag{3.5}$$

Whereas perturbations of order ε are required to destroy the ranks of $\mathbf{B}$ and $\mathbf{U}$, it only requires perturbations of order ε^2 to destroy the ranks of $\mathbf{B}^{\mathrm{H}}\mathbf{B}$ and $\mathbf{X}$. If ε^2 is smaller than the machine accuracy then we cannot even represent **X** on the machine as a full rank matrix.

We propose a variant of the Bartels–Stewart algorithm designed to allow the Cholesky factor **U** to be obtained directly. We shall assume that **A** is stable and that **C** is positive definite, and since we wish to be able to avoid the need to form $\mathbf{B}^H\mathbf{B}$ explicitly, we shall also assume that **C** is of the form of Equation (3.1). This is no real loss of generality since we can always take **B** to be the Cholesky factor of **C**. The assumptions imply that **B** is of full rank n and $m \geqslant n$. This simplifies the description of the method, but it should be noted that the method can easily be modified to allow these restrictions to be removed. We wish to solve the Lyapunov equation in the form

$$\mathbf{A}^H(\mathbf{U}^H\mathbf{U})+(\mathbf{U}^H\mathbf{U})\mathbf{A} = -\mathbf{B}^H\mathbf{B}, \tag{3.6}$$

for **U**, and first we show how we can transform this to a reduced equation, which is equivalent to Equation (2.3), of the form

$$\mathbf{S}^H(\tilde{\mathbf{U}}^H\tilde{\mathbf{U}})+(\tilde{\mathbf{U}}^H\tilde{\mathbf{U}})\mathbf{S} = -\tilde{\mathbf{R}}^H\tilde{\mathbf{R}}, \tag{3.7}$$

where $\tilde{\mathbf{U}}$ and $\tilde{\mathbf{R}}$ are upper triangular.

4. Transforming to Reduced Cholesky Form

The tool that enables us to avoid forming matrices of the form $\mathbf{B}^H\mathbf{B}$ is the QU-factorization (frequently called the QR-factorization, but not to be confused with the QR-algorithm). For an m by n matrix **B** of full rank with $m \geqslant n$ the QU-factorization is given by

$$\mathbf{B} = \mathbf{P}\begin{pmatrix}\mathbf{R}\\0\end{pmatrix}, \tag{4.1}$$

where **P** is an m by m unitary matrix and **R** is an n by n non-singular upper triangular matrix. The factorization can be chosen so that **R** has positive diagonal elements, although this is not usually important. The QU-factorization may be obtained in a numerically stable manner by means, for example, of Householder transformations and is in common use for solving linear least squares problems (Golub, 1965), because it avoids the need to form the normal, or Gram, matrix $\mathbf{B}^H\mathbf{B}$. The matrix **C** of Equation (3.1) can now be expressed as

$$\mathbf{C} = \mathbf{R}^H\mathbf{R} \tag{4.2}$$

so that **R** is the Cholesky factor of **C** and hence, as with Equation (1.4),

$$c_2(\mathbf{C}) = c_2^2(\mathbf{R}). \tag{4.3}$$

With the factorization of Equation (2.1) Equations (2.2) now become

$$\tilde{\mathbf{C}} = (\mathbf{RQ})^H(\mathbf{RQ}) \quad \text{and} \quad \tilde{\mathbf{X}} = (\mathbf{UQ})^H(\mathbf{UQ}). \tag{4.4}$$

If we let the QU-factorization of **RQ** be given by

$$\mathbf{RQ} = \tilde{\mathbf{P}}\tilde{\mathbf{R}}, \tag{4.5}$$

where $\tilde{\mathbf{P}}$ is unitary and $\tilde{\mathbf{R}}$ is upper triangular, then

$$\tilde{\mathbf{C}} = \tilde{\mathbf{R}}^{\mathrm{H}}\tilde{\mathbf{R}} \tag{4.6}$$

so that $\tilde{\mathbf{R}}$ is the Cholesky factor of $\tilde{\mathbf{C}}$ and is the matrix required in Equation (3.7). If we can solve Equation (3.7) for $\tilde{\mathbf{U}}$ then, by comparison with Equation (2.3), $\tilde{\mathbf{U}}$ will be the Cholesky factor of $\tilde{\mathbf{X}}$ so that

$$\tilde{\mathbf{X}} = \tilde{\mathbf{U}}^{\mathrm{H}}\tilde{\mathbf{U}}. \tag{4.7}$$

Once $\tilde{\mathbf{U}}$ has been found we can obtain the required Cholesky factor $\mathbf{U}$ by performing a QU-factorization of the matrix $\tilde{\mathbf{U}}\mathbf{Q}^{\mathrm{H}}$.

We note that we can also obtain the Cholesky factor $\tilde{\mathbf{R}}$ by forming $\mathbf{BQ}$ and then finding the QU-factorization of $\mathbf{BQ}$, but this involves more multiplications when $m > 7n/6$ and even when this is not true the difference is hardly significant. We note also that when $\mathbf{B}$ does not have full rank then one or more diagonal elements of $\mathbf{R}$ will be zero, and when $m < n$ if we partition $\mathbf{B}$ as

$$\mathbf{B} = (\hat{\mathbf{B}}\,\tilde{\mathbf{B}}),$$

where $\hat{\mathbf{B}}$ is an m by m matrix with the QU-factorization

$$\hat{\mathbf{B}} = \mathbf{P}\hat{\mathbf{R}},$$

then the Cholesky factor, $\mathbf{R}$, of $\mathbf{B}^{\mathrm{H}}\mathbf{B}$ is given by

$$\mathbf{R} = \begin{pmatrix} \hat{\mathbf{R}} & \mathbf{P}^{\mathrm{H}}\tilde{\mathbf{B}} \\ 0 & 0 \end{pmatrix}.$$

5. Solving the Reduced Equation

We now turn to the solution of Equation (3.7) and partition $\tilde{\mathbf{U}}$ and $\tilde{\mathbf{R}}$ as

$$\tilde{\mathbf{U}} = \begin{pmatrix} \tilde{u}_{11} & \tilde{\mathbf{u}}^{\mathrm{H}} \\ 0 & \mathbf{U}_1 \end{pmatrix}, \qquad \tilde{\mathbf{R}} = \begin{pmatrix} \tilde{r}_{11} & \tilde{\mathbf{r}}^{\mathrm{H}} \\ 0 & \mathbf{R}_1 \end{pmatrix}. \tag{5.1}$$

From Equations (2.4), (4.6) and (4.7) we see that

$$\tilde{c}_{11} = |\tilde{r}_{11}|^2, \qquad \tilde{\mathbf{c}} = \tilde{r}_{11}\,\tilde{\mathbf{r}}, \qquad \mathbf{C}_1 = \tilde{\mathbf{r}}\tilde{\mathbf{r}}^{\mathrm{H}} + \mathbf{R}_1^{\mathrm{H}}\mathbf{R}_1 \tag{5.2}$$

and

$$\tilde{x}_{11} = |\tilde{u}_{11}|^2, \qquad \tilde{\mathbf{x}} = \tilde{u}_{11}\,\tilde{\mathbf{u}}, \qquad \mathbf{X}_1 = \tilde{\mathbf{u}}\tilde{\mathbf{u}}^{\mathrm{H}} + \mathbf{U}_1^{\mathrm{H}}\mathbf{U}_1. \tag{5.3}$$

Substituting these into Equations (2.5), (2.6) and (2.7) gives

$$|\tilde{u}_{11}|^2 = -|\tilde{r}_{11}|^2/(\lambda_1 + \bar{\lambda}_1),$$
$$\tilde{u}_{11}(\mathbf{S}_1^{\mathrm{H}} + \lambda_1\mathbf{I})\tilde{\mathbf{u}} = -\tilde{r}_{11}\tilde{\mathbf{r}} - |\tilde{u}_{11}|^2\mathbf{s}$$

and

$$\mathbf{S}_1^{\mathrm{H}}\tilde{\mathbf{u}}\tilde{\mathbf{u}}^{\mathrm{H}} + \mathbf{S}_1^{\mathrm{H}}(\mathbf{U}_1^{\mathrm{H}}\mathbf{U}_1) + \tilde{\mathbf{u}}\tilde{\mathbf{u}}^{\mathrm{H}}\mathbf{S}_1 + (\mathbf{U}_1^{\mathrm{H}}\mathbf{U}_1)\mathbf{S}_1 = -(\tilde{\mathbf{r}}\tilde{\mathbf{r}}^{\mathrm{H}} + \mathbf{R}_1^{\mathrm{H}}\mathbf{R}_1) - (\tilde{u}_{11}^{\mathrm{H}}\mathbf{s}\tilde{\mathbf{u}}^{\mathrm{H}} + \tilde{u}_{11}\tilde{\mathbf{u}}\mathbf{s}^{\mathrm{H}}).$$

The true Cholesky factor has $\tilde{u}_{11}$ real and positive so that if we make this choice and we put

$$\mathbf{v} = \mathbf{S}_1^H\tilde{\mathbf{u}} + \tilde{u}_{11}\mathbf{s}, \tag{5.4}$$

then corresponding to Equations (2.5), (2.6) and (2.7) we have

$$\tilde{u}_{11} = |\tilde{r}_{11}|/[-(\lambda_1+\bar{\lambda}_1)]^{\frac{1}{2}} \tag{5.5}$$

$$(\mathbf{S}_1^H+\lambda_1\mathbf{I})\tilde{\mathbf{u}} = -(\tilde{r}_{11}/\tilde{u}_{11})\tilde{\mathbf{r}} - \tilde{u}_{11}\mathbf{s} \tag{5.6}$$

$$\mathbf{S}_1^H(\mathbf{U}_1^H\mathbf{U}_1)+(\mathbf{U}_1^H\mathbf{U}_1)\mathbf{S}_1 = -\mathbf{R}_1^H\mathbf{R}_1 - \tilde{\mathbf{r}}\tilde{\mathbf{r}}^H - (\mathbf{v}\tilde{\mathbf{u}}^H+\tilde{\mathbf{u}}\mathbf{v}^H). \tag{5.7}$$

Equations (5.5) and (5.6) enable us to obtain $\tilde{u}_{11}$ and $\tilde{\mathbf{u}}$, and, as we shall show, Equation (5.7) is of the same form as Equation (3.7), but of course of order $(n-1)$. For Equation (5.7) to be of the same form as Equation (3.7) we have to establish that the matrix $\mathbf{Z}$ given by

$$\mathbf{Z} = \mathbf{R}_1^H\mathbf{R}_1 + \tilde{\mathbf{r}}\tilde{\mathbf{r}}^H + (\mathbf{v}\tilde{\mathbf{u}}^H+\tilde{\mathbf{u}}\mathbf{v}^H) \tag{5.8}$$

is positive definite. This is an important distinction between Equation (2.7) and Equation (5.7). Let us put

$$\alpha = \tilde{r}_{11}/\tilde{u}_{11} \quad \text{and} \quad \mathbf{y} = \tilde{\mathbf{r}} - \bar{\alpha}\tilde{\mathbf{u}}. \tag{5.9}$$

Using Equation (5.5) we note that

$$|\alpha| = [-(\lambda_1+\bar{\lambda}_1)]^{\frac{1}{2}}. \tag{5.10}$$

Now, from Equations (5.4) and (5.6)

$$\mathbf{v} = -\alpha\tilde{\mathbf{r}} - \tilde{u}_{11}\mathbf{s} - \lambda_1\tilde{\mathbf{u}} + \tilde{u}_{11}\mathbf{s} = -\alpha\tilde{\mathbf{r}} - \lambda_1\tilde{\mathbf{u}}$$

so that

$$\begin{aligned}
\mathbf{v}\tilde{\mathbf{u}}^H + \tilde{\mathbf{u}}\mathbf{v}^H &= -\alpha\tilde{\mathbf{r}}\tilde{\mathbf{u}}^H - \bar{\alpha}\tilde{\mathbf{u}}\tilde{\mathbf{r}}^H - \lambda_1\tilde{\mathbf{u}}\tilde{\mathbf{u}}^H - \bar{\lambda}_1\tilde{\mathbf{u}}\tilde{\mathbf{u}}^H \\
&= -(\lambda_1+\bar{\lambda}_1)\tilde{\mathbf{u}}\tilde{\mathbf{u}}^H - [\tilde{\mathbf{r}}(\bar{\alpha}\tilde{\mathbf{u}})^H + (\bar{\alpha}\tilde{\mathbf{u}})\tilde{\mathbf{r}}^H] \\
&= -(\lambda_1+\bar{\lambda}_1)\tilde{\mathbf{u}}\tilde{\mathbf{u}}^H - [\tilde{\mathbf{r}}\tilde{\mathbf{r}}^H + |\alpha|^2\tilde{\mathbf{u}}\tilde{\mathbf{u}}^H - (\tilde{\mathbf{r}}-\bar{\alpha}\tilde{\mathbf{u}})(\tilde{\mathbf{r}}-\bar{\alpha}\tilde{\mathbf{u}})^H] \\
&= -\tilde{\mathbf{r}}\tilde{\mathbf{r}}^H + \mathbf{y}\mathbf{y}^H
\end{aligned}$$

and hence

$$\mathbf{Z} = \mathbf{R}_1^H\mathbf{R}_1 + \mathbf{y}\mathbf{y}^H. \tag{5.11}$$

Since $\mathbf{R}_1^H\mathbf{R}_1$ is positive definite $\mathbf{Z}$ must also be positive definite and hence $\mathbf{Z}$ has a Cholesky factorization, say

$$\mathbf{Z} = \hat{\mathbf{R}}^H\hat{\mathbf{R}}, \tag{5.12}$$

where $\hat{\mathbf{R}}$ is a non-singular upper triangular matrix. As we shall illustrate below $\hat{\mathbf{R}}$ can readily be obtained from $\mathbf{R}_1$ without the need to form $\mathbf{Z}$. Assuming that we can obtain $\hat{\mathbf{R}}$, Equations (5.5)–(5.7) which allow us to determine $\tilde{\mathbf{U}}$ can be summarized as

$$\tilde{u}_{11} = |\tilde{r}_{11}|/[-(\lambda_1+\bar{\lambda}_1)]^{\frac{1}{2}} \tag{5.13}$$

$$(\mathbf{S}_1^H+\lambda_1\mathbf{I})\tilde{\mathbf{u}} = -\alpha\tilde{\mathbf{r}} - \tilde{u}_{11}\mathbf{s} \tag{5.14}$$

$$\mathbf{S}_1^H(\mathbf{U}_1^H\mathbf{U}_1)+(\mathbf{U}_1^H\mathbf{U}_1)\mathbf{S}_1 = -\hat{\mathbf{R}}^H\hat{\mathbf{R}} \tag{5.15}$$

where

$$\hat{\mathbf{R}}^H\hat{\mathbf{R}} = \mathbf{R}_1^H\mathbf{R}_1 + \mathbf{y}\mathbf{y}^H, \qquad \mathbf{y} = \tilde{\mathbf{r}} - \bar{\alpha}\tilde{\mathbf{u}}, \qquad \alpha = \tilde{r}_{11}/\tilde{u}_{11}. \tag{5.16}$$

The problem of updating the Cholesky factorization of a matrix when the matrix is subject to a rank-one change and the equivalent problem of updating the QU-factorization of a matrix when a row is added to, or removed from, the matrix appears in a number of applications such as least squares, linear programming and non-linear optimization, and the techniques have received considerable attention in recent years (Golub, 1965; Lawson & Hanson, 1974; Saunders, 1972; Gill & Murray, 1974; Gill *et al.*, 1974; Gill & Murray, 1977; Paige, 1980; Stewart, 1979; Dongarra *et al.*, 1979).

Here we have a straightforward updating, as opposed to downdating, problem since we are making a strict rank-one addition to $\mathbf{R}_1^H\mathbf{R}_1$. If we let $\mathbf{F}$ be the matrix

$$\mathbf{F} = \begin{pmatrix} \mathbf{R}_1 \\ \mathbf{y}^H \end{pmatrix} \tag{5.17}$$

then

$$\mathbf{F}^H\mathbf{F} = \mathbf{R}_1^H\mathbf{R}_1 + \mathbf{y}\mathbf{y}^H = \hat{\mathbf{R}}^H\hat{\mathbf{R}} \tag{5.18}$$

and if we perform a QU-factorization of $\mathbf{F}$ then the upper triangular matrix will be the required Cholesky factor $\hat{\mathbf{R}}$ since, if

$$\mathbf{F} = \tilde{\mathbf{P}}\begin{pmatrix} \hat{\mathbf{R}} \\ 0 \end{pmatrix}, \tag{5.19}$$

where $\tilde{\mathbf{P}}$ is orthogonal then

$$\mathbf{F}^H\mathbf{F} = \hat{\mathbf{R}}^H\hat{\mathbf{R}}$$

as required. The form of $\mathbf{F}$ is typified by the case where $n = 6$ as

$$\mathbf{F} = \begin{pmatrix} x & x & x & x & x \\ & x & x & x & x \\ & & x & x & x \\ & & & x & x \\ & & & & x \\ x & x & x & x & x \end{pmatrix}$$

and $\mathbf{F}$ may be transformed to upper triangular form by a sequence of plane rotations in planes $(1, n)$, $(2, n)$, ..., $(n-1, n)$, where the rotation in the (i, n) plane is chosen to annihilate the element in the (n, i) position of $\mathbf{F}$.

We note that although we have assumed $\mathbf{B}$, and hence $\tilde{\mathbf{R}}$, to be of full rank and that $m \geqslant n$, these restrictions can in fact be relaxed by defining α as

$$\alpha = \text{sign}\,(\tilde{r}_{11})[-(\lambda_1 + \bar{\lambda}_1)]^{\frac{1}{2}}, \qquad \text{sign}\,(a) = \begin{cases} a/|a|, & a \neq 0 \\ 1, & a = 0, \end{cases} \tag{5.20}$$

which merely extends the definition of α to allow for the case where $\tilde{r}_{11} = 0$. Equations (5.13)–(5.16) can now be used to determine $\tilde{\mathbf{U}}$.

This concludes the description of the basic method of obtaining the Cholesky factor of $\mathbf{X}$.

6. The Real Non-negative Definite Case

We noted at the end of Section 2 that when **A** and **C** are both real then we can work entirely in real arithmetic by using the factorization of Equation (2.8). For this case, corresponding to Equations (2.9)–(2.11), Equations (5.13)–(5.16) are replaced by the equations

$$s_{11}^{T}(\tilde{u}_{11}^{T}\tilde{u}_{11})+(\tilde{u}_{11}^{T}\tilde{u}_{11})s_{11} = -\tilde{r}_{11}^{T}r_{11} \tag{6.1}$$

$$\mathbf{S}_1^{T}\tilde{\mathbf{u}}+\tilde{\mathbf{u}}(\tilde{u}_{11}s_{11}\tilde{u}_{11}^{-1}) = -\tilde{\mathbf{r}}\alpha-\mathbf{s}\tilde{u}_{11}^{T} \tag{6.2}$$

$$\mathbf{S}_1^{T}(\mathbf{U}_1^{T}\mathbf{U}_1)+(\mathbf{U}_1^{T}\mathbf{U}_1)\mathbf{S}_1 = -\hat{\mathbf{R}}^{T}\hat{\mathbf{R}} \tag{6.3}$$

where

$$\hat{\mathbf{R}}^{T}\hat{\mathbf{R}} = \mathbf{R}_1^{T}\mathbf{R}_1+\mathbf{y}\mathbf{y}^{T}, \qquad \mathbf{y} = \tilde{\mathbf{r}}-\tilde{\mathbf{u}}\alpha^{T}, \qquad \alpha = \tilde{r}_{11}\tilde{u}_{11}^{-1}. \tag{6.4}$$

s_{11}, $\tilde{u}_{11}$ and $\tilde{r}_{11}$ are either scalars or two by two matrices and $\mathbf{s}$, $\tilde{\mathbf{u}}$ and $\tilde{\mathbf{r}}$ are either vectors or matrices with two columns. Of course, in the scalar case $\tilde{u}_{11}s_{11}\tilde{u}_{11}^{-1} = s_{11}$. The matrix **F** of Equation (5.17) is replaced by

$$\mathbf{F} = \begin{pmatrix}\mathbf{R}_1\\ \mathbf{y}^{T}\end{pmatrix} \tag{6.5}$$

and in the two by two case $\mathbf{y}^{T}$ contains two rows so that there are two subdiagonal elements per column to be annihilated in restoring **F** to upper triangular form.

There are some computational pitfalls to be avoided in the two by two case and the remainder of this section is devoted to the tedious but important details needed for this case. In the scalar case Equation (6.1) gives

$$\tilde{u}_{11} = \tilde{r}_{11}/(-2s_{11})^{\frac{1}{2}}, \quad \text{scalar case,} \tag{6.6}$$

but in the two by two case some care is needed because we naturally wish to avoid forming $\tilde{r}_{11}^{T}\tilde{r}_{11}$ explicitly and we wish to avoid finding $\tilde{u}_{11}$ from $\tilde{x}_{11} = \tilde{u}_{11}^{T}\tilde{u}_{11}$. The most satisfactory solution seems to be to regard Equation (6.1) as a special case of Equation (3.6) and hence reduce Equation (6.1) to the form of Equation (3.7). This approach also allows us to handle Equation (6.2) in a satisfactory manner, even when $\tilde{u}_{11}$ is singular.

The reduction of s_{11} to upper triangular form can be achieved with a single plane rotation. If we denote s_{11} as

$$s_{11} = \begin{pmatrix} s_1 & s_2\\ s_3 & s_4\end{pmatrix}, \tag{6.7}$$

let λ be an eigenvalue of s_{11} and let $\hat{\mathbf{Q}}^{H}$ be the plane rotation matrix that annihilates the second element of the vector **z** given by

$$\mathbf{z} = \begin{pmatrix} s_1-\lambda\\ s_3\end{pmatrix}, \tag{6.8}$$

then the matrix $\hat{s}_{11} = \hat{\mathbf{Q}}^{H}s_{11}\hat{\mathbf{Q}}$ will be upper triangular. Since this case corresponds to a complex conjugate pair of eigenvalues, $\hat{s}_{11}$ will have the form

$$\hat{s}_{11} = \begin{pmatrix}\beta & \bar{t}\\ 0 & \bar{\beta}\end{pmatrix}, \quad \text{where } \beta = \lambda \text{ or } \beta = \bar{\lambda}. \tag{6.9}$$

Now let the QU-factorizations of $\tilde{u}_{11}Q$ and $\tilde{r}_{11}Q$ be

$$\tilde{u}_{11}\hat{Q} = \hat{T}\tilde{v}_{11} \quad \text{and} \quad \tilde{r}_{11}\hat{Q} = \hat{P}\tilde{p}_{11}, \tag{6.10}$$

where $\hat{T}$ and $\hat{P}$ are unitary and

$$\tilde{v}_{11} = \begin{pmatrix} v_1 & \bar{v}_2 \\ 0 & v_3 \end{pmatrix}, \qquad \tilde{p}_{11} = \begin{pmatrix} p_1 & \bar{p}_2 \\ 0 & p_3 \end{pmatrix}. \tag{6.11}$$

$\hat{T}$ and $\hat{P}$ can be chosen so that v_1, v_3, p_1 and p_3 are real and non-negative and for simplicity we shall assume this to be the case. Then with the transformations of Equations (6.9) and (6.10) Equation (6.1) becomes

$$\hat{s}_{11}^{H}(\tilde{v}_{11}^{H}\tilde{v}_{11}) + (\tilde{v}_{11}^{H}\tilde{v}_{11})\hat{s}_{11} = -\tilde{p}_{11}^{H}\tilde{p}_{11} \tag{6.12}$$

from which we find that

$$v_1 = p_1/\tilde{\alpha}, \quad \tilde{\alpha} = [-(\beta+\bar{\beta})]^{\frac{1}{2}} \tag{6.13}$$

$$v_2 = -(\tilde{\alpha}p_2 + v_1 t)/(2\beta) \tag{6.14}$$

$$v_3 = (p_3^2 + |y|^2)^{\frac{1}{2}}/\tilde{\alpha}, \quad y = p_2 - \tilde{\alpha}v_2. \tag{6.15}$$

Having found $\tilde{v}_{11}$ we then obtain $\tilde{u}_{11}$ from the QU-factorization of $\tilde{v}_{11}\hat{Q}^{H}$. Choosing the diagonal elements of $\tilde{u}_{11}$ to be real ensures *a priori* that the remaining element is also real.

We now turn our attention to Equation (6.2) and in particular to the computation of $(\tilde{u}_{11}s_{11}\tilde{u}_{11}^{-1})$ and α. First we note that

$$\tilde{u}_{11}s_{11}\tilde{u}_{11}^{-1} = \hat{T}(\tilde{v}_{11}\hat{s}_{11}\tilde{v}_{11}^{-1})\hat{T}^{H} \quad \text{and} \quad \alpha = \hat{P}(\tilde{p}_{11}\tilde{v}_{11}^{-1})\hat{T}^{H}. \tag{6.16}$$

Now

$$\tilde{v}_{11}\hat{s}_{11}\tilde{v}_{11}^{-1} = \begin{pmatrix} \beta & [\bar{v}_2(\bar{\beta}-\beta)+v_1\bar{t}]/v_3 \\ 0 & \bar{\beta} \end{pmatrix} \tag{6.17}$$

and

$$\tilde{p}_{11}\tilde{v}_{11}^{-1} = \begin{pmatrix} \tilde{\alpha} & (\bar{p}_2 - \tilde{\alpha}\bar{v}_2)/v_3 \\ 0 & p_3/v_3 \end{pmatrix}. \tag{6.18}$$

A bit of algebraic manipulation shows that

$$\bar{v}_2(\bar{\beta}-\beta) + v_1\bar{t} = \tilde{\alpha}(\tilde{\alpha}\bar{v}_2 - \bar{p}_2) \tag{6.19}$$

and hence

$$\tilde{v}_{11}\hat{s}_{11}\tilde{v}_{11}^{-1} = \begin{pmatrix} \beta & \bar{\gamma} \\ 0 & \bar{\beta} \end{pmatrix}, \qquad \tilde{p}_{11}\tilde{v}_{11}^{-1} = \begin{pmatrix} \tilde{\alpha} & \bar{\delta} \\ 0 & \eta \end{pmatrix}, \tag{6.20}$$

where

$$\delta = y/v_3, \qquad \eta = p_3/v_3, \qquad \gamma = -\tilde{\alpha}\delta. \tag{6.21}$$

In our NPL routine for solving the non-negative definite Lyapunov equation we choose $\delta = 0$ and $\eta = \tilde{\alpha}$ whenever $y = 0$ which allows us to handle the case where $y = 0$ and $p_3 = 0$ quite satisfactorily and we then have for any y and p_3 that

$$|\delta| \leqslant \tilde{\alpha}, \quad 0 \leqslant \eta \leqslant \tilde{\alpha} \quad \text{and} \quad |\gamma| \leqslant \tilde{\alpha}^2. \tag{6.22}$$

Having found $\tilde{v}_{11}\hat{s}_{11}\tilde{v}_{11}^{-1}$ and $\tilde{p}_{11}\tilde{v}_{11}^{-1}$ we can then recover $\tilde{u}_{11}s_{11}\tilde{u}_{11}^{-1}$ and α from Equation (6.16).

Although, as we have shown, $\tilde{u}_{11}s_{11}\tilde{u}_{11}^{-1}$ and α are rather special it is imperative that they are computed as described above in order to avoid the serious effects of rounding errors when $\tilde{u}_{11}$ is ill-conditioned. (A discussion of such effects in a rather different context is given in Hammarling & Wilkinson, 1980.)

An example of a Lyapunov equation which gives rise to an almost singular two by two matrix $\tilde{u}_{11}$ is given by

$$A = \begin{pmatrix} 2 & -(3+\varepsilon) & 6 & 7 \\ 3 & -4 & 4 & 5 \\ 0 & 0 & 2 & -3 \\ 0 & 0 & 3 & -4 \end{pmatrix}, \quad \varepsilon > 0 \text{ and small,}$$

$$B = \begin{pmatrix} 1 & -1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

for which

$$\tilde{u}_{11} = 1/\{2[(1+3\varepsilon)(2+3\varepsilon)]^{\frac{1}{2}}\} \begin{pmatrix} 2+3\varepsilon & -(2+4\varepsilon) \\ 0 & \varepsilon(1+3\varepsilon)^{\frac{1}{2}} \end{pmatrix}.$$

7. The Case Where A is Normal

If $\mathbf{A}$ is a normal matrix so that

$$\mathbf{A}^H\mathbf{A} = \mathbf{A}\mathbf{A}^H \tag{7.1}$$

then considerable simplification takes place because the Schur factorization of Equation (2.1) can be replaced by the spectral factorization (Wilkinson, 1965), given by

$$\mathbf{A} = \mathbf{Q}\mathbf{D}\mathbf{Q}^H, \tag{7.2}$$

where $\mathbf{D}$ is the diagonal matrix

$$\mathbf{D} = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}. \tag{7.3}$$

For the Bartels–Stewart algorithm Equations (2.5)–(2.7) become

$$\tilde{x}_{11} = -\tilde{c}_{11}/(\lambda_1 + \bar{\lambda}_1) \tag{7.4}$$

$$(\mathbf{D}_1 + \lambda_1 \mathbf{I})\tilde{\mathbf{x}} = -\tilde{\mathbf{c}} \tag{7.5}$$

$$\mathbf{D}_1^H \mathbf{X}_1 + \mathbf{X}_1 \mathbf{D}_1 = -\mathbf{C}_1. \tag{7.6}$$

Equation (7.5) gives immediately that

$$\tilde{x}_{i1} = -\tilde{c}_{i1}/(\bar{\lambda}_i + \lambda_1)$$

and hence from Equations (7.4) and (7.6) the elements of $\tilde{\mathbf{X}}$ are given by

$$\tilde{x}_{ij} = -\tilde{c}_{ij}/(\bar{\lambda}_i + \lambda_j). \tag{7.7}$$

For the modified algorithm designed to determine $\tilde{\mathbf{U}}$ Equations (5.13)–(5.16) give

$$\tilde{u}_{11} = |\tilde{r}_{11}|/[-(\lambda_1 + \bar{\lambda}_1)]^{\frac{1}{2}} \tag{7.8}$$

$$\tilde{u}_{i1} = -(\alpha\tilde{r}_{i1})/(\bar{\lambda}_i + \lambda_1), \quad i = 2, 3, \ldots, n \tag{7.9}$$

$$\mathbf{D}_1^{\mathrm{H}}(\mathbf{U}_1^{\mathrm{H}}\mathbf{U}_1) + (\mathbf{U}_1^{\mathrm{H}}\mathbf{U}_1)\mathbf{D}_1 = -\hat{\mathbf{R}}^{\mathrm{H}}\hat{\mathbf{R}}, \tag{7.10}$$

where $\hat{\mathbf{R}}^{\mathrm{H}}\hat{\mathbf{R}}$ is as given by Equation (5.16) and so we still have to update $\mathbf{R}_1$ to give $\hat{\mathbf{R}}$ at each step.

8. The Kronecker Product Form of the Lyapunov Equation

Let $\mathbf{A} \otimes \mathbf{B}$ denote the Kronecker product of an m by n matrix $\mathbf{A}$ with a matrix $\mathbf{B}$ given by

$$\mathbf{A} \otimes \mathbf{B} = \begin{pmatrix} a_{11}\mathbf{B} & a_{12}\mathbf{B} & \ldots & a_{1n}\mathbf{B} \\ a_{21}\mathbf{B} & a_{22}\mathbf{B} & \ldots & a_{2n}\mathbf{B} \\ \vdots & \vdots & & \vdots \\ a_{m1}\mathbf{B} & a_{m2}\mathbf{B} & \ldots & a_{mn}\mathbf{B} \end{pmatrix}, \tag{8.1}$$

let $\mathbf{v}_j$ denote the jth column of a matrix $\mathbf{V}$ containing n columns and let $\underline{\mathbf{v}} = \mathrm{vec}\,(\mathbf{V})$ be defined as

$$\underline{\mathbf{v}} = \mathrm{vec}\,(\mathbf{V}) = \begin{pmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \\ \vdots \\ \mathbf{v}_n \end{pmatrix}. \tag{8.2}$$

Note that amongst the properties satisfied by Kronecker products are the following (Barnett, 1975):

$$(\mathbf{A} \otimes \mathbf{B})(\mathbf{C} \otimes \mathbf{D}) = \mathbf{AC} \otimes \mathbf{BD}, \qquad (\mathbf{A} \otimes \mathbf{B})^{\mathrm{H}} = \mathbf{A}^{\mathrm{H}} \otimes \mathbf{B}^{\mathrm{H}},$$

$$\mathbf{A} \otimes (\mathbf{B} + \mathbf{C}) = \mathbf{A} \otimes \mathbf{B} + \mathbf{A} \otimes \mathbf{C}, \qquad (\mathbf{A} + \mathbf{B}) \otimes \mathbf{C} = \mathbf{A} \otimes \mathbf{C} + \mathbf{B} \otimes \mathbf{C}.$$

If $\mathbf{Ax} = \lambda\mathbf{x}$ and $\mathbf{By} = \beta\mathbf{y}$ then

$$(\mathbf{A} \otimes \mathbf{B})(\mathbf{x} \otimes \mathbf{y}) = \lambda\beta(\mathbf{x} \otimes \mathbf{y}), \qquad (\mathbf{I} \otimes \mathbf{A} + \mathbf{B} \otimes \mathbf{I})(\mathbf{x} \otimes \mathbf{y}) = (\lambda + \beta)(\mathbf{x} \otimes \mathbf{y}).$$

Note that it follows from the first of these two results that

$$\|\mathbf{I} \otimes \mathbf{A}\|_2 = \|\mathbf{A}\|_2.$$

If we now consider the jth column of the Lyapunov equation

$$\mathbf{A}^{\mathrm{H}}\mathbf{X} + \mathbf{XA} = -\mathbf{C} \tag{8.3}$$

we get

$$\mathbf{A}^H\mathbf{x}_j + a_{1j}\mathbf{x}_1 + a_{2j}\mathbf{x}_2 + \ldots + a_{jj}\mathbf{x}_j + \ldots + a_{nj}\mathbf{x}_n = -\mathbf{c}_j,$$

and hence the solution of Equation (8.3) satisfies the equation

$$(\mathbf{I} \otimes \mathbf{A}^H + \mathbf{A}^T \otimes \mathbf{I})\underset{\sim}{\mathbf{x}} = -\underset{\sim}{\mathbf{c}}. \tag{8.4}$$

This is the Kronecker product form of the Lyapunov equation. Putting

$$\mathbf{K} = \mathbf{I} \otimes \mathbf{A}^H + \mathbf{A}^T \otimes \mathbf{I} \tag{8.5}$$

the matrix $\mathbf{K}$ is given by

$$\mathbf{K} = \begin{pmatrix} \mathbf{A}^H + a_{11}\mathbf{I} & a_{21}\mathbf{I} & \ldots & a_{n1}\mathbf{I} \\ a_{12}\mathbf{I} & \mathbf{A}^H + a_{22}\mathbf{I} & \ldots & a_{n2}\mathbf{I} \\ \vdots & \vdots & & \vdots \\ a_{1n}\mathbf{I} & a_{2n}\mathbf{I} & \ldots & \mathbf{A}^H + a_{nn}\mathbf{I} \end{pmatrix} \tag{8.6}$$

so that although $\mathbf{K}$ is an n^2 by n^2 matrix it has a rather special form. Now let $\mathbf{Q}^C$ denote the matrix whose elements are the complex conjugates of those of $\mathbf{Q}$ and consider the matrix $\mathbf{N}$ given by

$$\mathbf{N} = \mathbf{Q}^C \otimes \mathbf{Q}, \tag{8.7}$$

where $\mathbf{Q}$ is the matrix of Equation (2.1). For this matrix

$$\mathbf{N}^H = \mathbf{Q}^T \otimes \mathbf{Q}^H \quad \text{and} \quad \mathbf{N}^H\mathbf{N} = (\mathbf{Q}^T\mathbf{Q}^C) \otimes (\mathbf{Q}^H\mathbf{Q}) = \mathbf{I} \otimes \mathbf{I} = \mathbf{I}$$

so that $\mathbf{N}$ is unitary. We also have that

$$\begin{aligned} \mathbf{N}^H\mathbf{K}\mathbf{N} &= (\mathbf{Q}^T \otimes \mathbf{Q}^H)(\mathbf{I} \otimes \mathbf{A}^H + \mathbf{A}^T \otimes \mathbf{I})(\mathbf{Q}^C \otimes \mathbf{Q}) \\ &= \mathbf{Q}^T\mathbf{Q}^C \otimes \mathbf{Q}^H\mathbf{A}^H\mathbf{Q} + \mathbf{Q}^T\mathbf{A}^T\mathbf{Q}^C \otimes \mathbf{Q}^H\mathbf{Q} \\ &= \mathbf{I} \otimes \mathbf{S}^H + \mathbf{S}^T \otimes \mathbf{I}. \end{aligned} \tag{8.8}$$

Putting

$$\mathbf{L} = \mathbf{I} \otimes \mathbf{S}^H + \mathbf{S}^T \otimes \mathbf{I} \tag{8.9}$$

we can see that

$$\mathbf{L}\tilde{\underset{\sim}{\mathbf{x}}} = -\tilde{\underset{\sim}{\mathbf{c}}} \tag{8.10}$$

is the Kronecker product form of the Lyapunov Equation (2.3). The matrix $\mathbf{L}$ is lower triangular and is given by

$$\mathbf{L} = \begin{pmatrix} \mathbf{S}^H + \lambda_1\mathbf{I} & 0 & 0 & \ldots & 0 \\ s_{12}\mathbf{I} & \mathbf{S}^H + \lambda_2\mathbf{I} & 0 & \ldots & 0 \\ s_{13}\mathbf{I} & s_{23}\mathbf{I} & \mathbf{S}^H + \lambda_3\mathbf{I} & \ldots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ s_{1n}\mathbf{I} & s_{2n}\mathbf{I} & s_{3n}\mathbf{I} & \ldots & \mathbf{S}^H + \lambda_n\mathbf{I} \end{pmatrix} \tag{8.11}$$

so that Equation (8.10) can be solved by block forward substitution, the jth step being to solve the n by n lower triangular system

$$(\mathbf{S}^H + \lambda_j\mathbf{I})\tilde{\mathbf{x}}_j = -\tilde{\mathbf{c}}_j - s_{1j}\tilde{\mathbf{x}}_1 - s_{2j}\tilde{\mathbf{x}}_2 - \ldots - s_{j-1,j}\tilde{\mathbf{x}}_{j-1}, \quad j = 1, 2, \ldots, n. \tag{8.12}$$

Since $\tilde{\mathbf{X}}$ is symmetric we, in fact, need only compute the final $(n-j+1)$ elements of $\tilde{\mathbf{x}}_j$, the first $(j-1)$ elements having been obtained on previous steps.

This is the Kronecker product form of the Bartels–Stewart algorithm.

9. Sensitivity of the Solution of the Lyapunov Equation

Laub (1979, 1980) has given a measure for the sensitivity of the solution of the Lyapunov equation to changes in the data for the case where A is a normal matrix. If we consider just the case where A is perturbed and hence consider the equation

$$(\mathbf{A}+\mathbf{E})^H(\mathbf{X}+\mathbf{F})+(\mathbf{X}+\mathbf{F})(\mathbf{A}+\mathbf{E}) = -\mathbf{C} \tag{9.1}$$

then

$$\mathbf{A}^H\mathbf{F}+\mathbf{F}\mathbf{A} = -[\mathbf{E}^H(\mathbf{X}+\mathbf{F})+(\mathbf{X}+\mathbf{F})\mathbf{E}] \tag{9.2}$$

and with the factorization of Equation (2.1) this becomes

$$\mathbf{S}^H\tilde{\mathbf{F}}+\tilde{\mathbf{F}}\mathbf{S} = -\mathbf{G}, \tag{9.3}$$

where

$$\mathbf{G} = \tilde{\mathbf{E}}^H(\tilde{\mathbf{X}}+\tilde{\mathbf{F}})+(\tilde{\mathbf{X}}+\tilde{\mathbf{F}})\tilde{\mathbf{E}}, \qquad \tilde{\mathbf{F}} = \mathbf{Q}^H\mathbf{F}\mathbf{Q}, \qquad \tilde{\mathbf{E}} = \mathbf{Q}^H\mathbf{E}\mathbf{Q}. \tag{9.4}$$

The spectral norm of G satisfies

$$\|\mathbf{G}\|_2 \leq 2\|\tilde{\mathbf{E}}\|_2\,\|\tilde{\mathbf{X}}+\tilde{\mathbf{F}}\|_2 = 2\|\mathbf{E}\|_2\,\|\mathbf{X}+\mathbf{F}\|_2. \tag{9.5}$$

Now, in the case where A is normal, Equation (7.7) gives

$$\tilde{f}_{ij} = -g_{ij}/(\bar{\lambda}_i+\lambda_j), \tag{9.6}$$

so that

$$\begin{aligned}\|\tilde{\mathbf{F}}\|_2 = \|\mathbf{F}\|_2 &\leq n^{\frac{1}{2}}\|\mathbf{G}\|_2/\min_{ij}|\bar{\lambda}_i+\lambda_j| \\ &\leq 2n^{\frac{1}{2}}\left(\max_i|\lambda_i|/\min_{ij}|\bar{\lambda}_i+\lambda_j|\right)(\|\mathbf{E}\|_2/\|\mathbf{A}\|_2)\|\mathbf{X}+\mathbf{F}\|_2\end{aligned}$$

and hence

$$\|\mathbf{F}\|_2/\|\mathbf{X}+\mathbf{F}\|_2 \leq 2n^{\frac{1}{2}}\left(\max_i|\lambda_i|/\min_{ij}|\bar{\lambda}_i+\lambda_j|\right)(\|\mathbf{E}\|_2/\|\mathbf{A}\|_2). \tag{9.7}$$

Thus we can regard the value

$$\max_i|\lambda_i|/\min_{ij}|\bar{\lambda}_i+\lambda_j|$$

as a condition number for this problem, which is essentially equivalent to the result quoted by Laub.

Obtaining an equivalent result to Equation (9.7) when A is not normal seems unlikely to be straightforward, but Golub *et al.* (1979) have given an analysis of the sensitivity of the real Sylvester equation in terms of the Kronecker product and this leads to a useful practical method of measuring the sensitivity. Here we give an analysis only in terms of the Lyapunov equation.

If A is perturbed as in Equation (9.1) then

$$\mathbf{I}\otimes(\mathbf{A}+\mathbf{E})^H+(\mathbf{A}+\mathbf{E})^T\otimes\mathbf{I} = \mathbf{K}+(\mathbf{I}\otimes\mathbf{E}^H+\mathbf{E}^T\otimes\mathbf{I})$$

and corresponding to Equation (9.1) we have

$$(\mathbf{K}+\mathbf{G})(\underline{x}+\underline{f}) = -\underline{c}, \qquad \mathbf{G} = \mathbf{I}\otimes\mathbf{E}^{H}+\mathbf{E}^{T}\otimes\mathbf{I}, \qquad \underline{f} = \mathrm{vec}\,(\mathbf{F}). \tag{9.8}$$

Using a standard result on the sensitivity of linear equations (see, for example, Forsythe & Moler, 1967) this gives

$$\|\underline{f}\|_2/\|\underline{x}+\underline{f}\|_2 \leqslant \|\mathbf{K}^{-1}\|_2\,\|\mathbf{G}\|_2 \tag{9.9}$$

so that denoting the Euclidean (Frobenius) norm of $\mathbf{A}$ by $\|\mathbf{A}\|_E$ we get

$$\|\mathbf{F}\|_E/\|\mathbf{X}+\mathbf{F}\|_E \leqslant (2\|\mathbf{K}^{-1}\|_2\,\|\mathbf{A}\|_2)(\|\mathbf{E}\|_2/\|\mathbf{A}\|_2) \tag{9.10}$$

and hence we can regard the value $\|\mathbf{K}^{-1}\|_2\,\|\mathbf{A}\|_2$ as a condition number for this problem. This is essentially a special case of the result of Golub *et al.* (1979). To obtain the above result we have used the inequality $\|\mathbf{G}\|_2 \leqslant 2\|\mathbf{E}\|_2$, but it should be noted that when $\mathbf{E}$ is complex this can be a substantial overestimate. For example when

$$\mathbf{E} = \varepsilon\begin{pmatrix} i & 0 \\ 0 & i \end{pmatrix}$$

we find that $\mathbf{G} = 0$ so that $\|\mathbf{G}\|_2 = 0$, but $\|\mathbf{E}\|_2 = |\varepsilon|$.

In order to obtain a means of estimating $\|\mathbf{K}^{-1}\|_2$ we note that, since the matrix $\mathbf{L}$ of Equation (8.9) is unitarily similar to $\mathbf{K}$, we have

$$\|\mathbf{K}^{-1}\|_2 = \|\mathbf{L}^{-1}\|_2$$

and hence

$$\|\mathbf{F}\|_E/\|\mathbf{X}+\mathbf{F}\|_E \leqslant (2\|\mathbf{L}^{-1}\|_2\,\|\mathbf{A}\|_2)(\|\mathbf{E}\|_2/\|\mathbf{A}\|_2). \tag{9.11}$$

The value $\|\mathbf{L}^{-1}\|$ can be estimated, as $\|\mathbf{z}\|_2/\|\mathbf{y}\|_2$, by solving the two sets of equations $\mathbf{L}^T\mathbf{y} = \mathbf{b}$ and $\mathbf{L}\mathbf{z} = \mathbf{y}$, where the vector $\mathbf{b}$ is chosen to promote as large a value of $\|\mathbf{y}\|_2$ as possible using the technique described in Cline, Moler, Stewart & Wilkinson (1979) (see also Dongarra *et al.*, 1979; O'Leary, 1980). Although this estimate requires about $2n^3$ multiplications this is not a great increase in cost relative to the solution of the Lyapunov equation itself. The biggest disadvantage is the extra n^2 storage locations needed for the estimator.

Note that when $\mathbf{A}$ is normal $\mathbf{S}$ is diagonal so that $\mathbf{L}$ is also diagonal with diagonal elements $(\bar{\lambda}_i+\lambda_j)$ and hence in this case

$$\|\mathbf{L}^{-1}\|_2\,\|\mathbf{A}\|_2 = \max_i |\lambda_i|/\min_{ij} |\bar{\lambda}_i+\lambda_j|$$

so that Equation (9.11) is consistent with Equation (9.7).

We also note that the matrix $\mathbf{L}$ will be ill-conditioned if the matrix $\mathbf{A}$ has one or more eigenvalues close to the imaginary axis relative to any of the other eigenvalues and from practical considerations one would expect this to be the case because $\mathbf{A}$ is only just stable (Laub, 1979; Bucy, 1975). The matrix $\mathbf{L}$ will also be ill-conditioned if $\mathbf{A}$ is close to a matrix with one or more eigenvalues close to the imaginary axis relative to any of the other eigenvalues, but unless we investigate the sensitivity of the eigenvalues of $\mathbf{A}$ (Wilkinson, 1965), or estimate $\|\mathbf{L}^{-1}\|$, such a matrix may not be easy to detect. For example, if we consider the case where A and C are given by

$$\mathbf{A} = \begin{pmatrix} -0.5 & 1 & \dots & 1 \\ 0 & -0.5 & \dots & 1 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & -0.5 \end{pmatrix}, \qquad \mathbf{C} = \mathbf{I}$$

so that

$$\mathbf{A} = \mathbf{S}, \qquad \mathbf{C} = \tilde{\mathbf{C}} \quad \text{and} \quad \mathbf{X} = \tilde{\mathbf{X}},$$

then we find that

$$\mathbf{S}_1^H + \lambda_1 \mathbf{I} = \begin{pmatrix} -1 & 0 & \dots & 0 \\ 1 & -1 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 1 & 1 & \dots & -1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} x_{11} \\ \mathbf{x} \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 2 \\ \vdots \\ 2^{n-2} \end{pmatrix}.$$

As can be seen from the growth in the elements of $\mathbf{x}$, the matrix $(\mathbf{S}_1^H + \lambda_1 \mathbf{I})$ is very ill-conditioned when n is not small (Wilkinson, 1977). To see that, despite appearances, A is only just stable we note that the matrix $(\mathbf{A}+\mathbf{E})$, where

$$\mathbf{E} = \begin{pmatrix} 0 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 0 \\ \varepsilon & 0 & \dots & 0 \end{pmatrix}$$

has a zero eigenvalue when $\varepsilon = 1/(4.3^{n-2})$.

The additional question of whether or not the Cholesky factor, U, of X is less sensitive to perturbations than X itself also seems to be a difficult question to answer. Comparing Equation (5.13) with (2.6) suggests that at worst U is no more sensitive to perturbations in A than X and it is not difficult to construct examples where U is very much less sensitive. For example, if we consider the case where

$$\mathbf{A} = \begin{pmatrix} -\varepsilon & 0 \\ 1-\varepsilon & -1 \end{pmatrix}, \qquad \mathbf{B} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \qquad \varepsilon > 0,$$

then the solution of the equation $\mathbf{A}^T\mathbf{X} + \mathbf{X}\mathbf{A} = -\mathbf{B}^T\mathbf{B}$ is

$$\mathbf{X} = \frac{1}{2\varepsilon} \begin{pmatrix} 1 & 1 \\ 1 & 1+\varepsilon \end{pmatrix}$$

and the Cholesky factor of X is

$$\mathbf{U} = \frac{1}{(2\varepsilon)^{\frac{1}{2}}} \begin{pmatrix} 1 & 1 \\ 0 & \varepsilon^{\frac{1}{2}} \end{pmatrix}$$

so that $\|\mathbf{X}\|_2$ varies as ε^{-1}, but $\|\mathbf{U}\|_2$ only varies as $\varepsilon^{-\frac{1}{2}}$.

10. The Discrete-time Lyapunov Equation

Consider the continuous-time Lyapunov equation

$$\mathbf{A}_1^H\mathbf{X}+\mathbf{X}\mathbf{A}_1 = -\mathbf{C}_1, \qquad \mathbf{C}_1 = \mathbf{C}_1^H. \tag{10.1}$$

Suppose that $(\mathbf{I}-\mathbf{A}_1)^{-1}$ exists and that we define the matrix $\mathbf{A}$ as

$$\mathbf{A} = (\mathbf{I}-\mathbf{A}_1)^{-1}(\mathbf{I}+\mathbf{A}_1) \tag{10.2}$$

so that

$$\mathbf{A}_1 = (\mathbf{A}-\mathbf{I})(\mathbf{A}+\mathbf{I})^{-1}.$$

Under this transformation Equation (10.1) becomes

$$(\mathbf{A}+\mathbf{I})^{-H}(\mathbf{A}^H-\mathbf{I})\mathbf{X}+\mathbf{X}(\mathbf{A}-\mathbf{I})(\mathbf{A}+\mathbf{I})^{-1} = -\mathbf{C}_1$$

which can be re-arranged to give

$$\mathbf{A}^H\mathbf{X}\mathbf{A}-\mathbf{X} = -\tfrac{1}{2}(\mathbf{A}+\mathbf{I})^H\mathbf{C}_1(\mathbf{A}+\mathbf{I})$$

and putting

$$\mathbf{C} = \tfrac{1}{2}(\mathbf{A}+\mathbf{I})^H\mathbf{C}_1(\mathbf{A}+\mathbf{I}) \tag{10.3}$$

we have

$$\mathbf{A}^H\mathbf{X}\mathbf{A}-\mathbf{X} = -\mathbf{C}. \tag{10.4}$$

This is the discrete-time Lyapunov equation and plays the corresponding role for discrete-time systems as Equation (1.1) for continuous-time systems (Barnett, 1975).

If λ_i is an eigenvalue of $\mathbf{A}$ and β_i the corresponding eigenvalue of $\mathbf{A}_1$ then

$$\lambda_i = (1+\beta_i)/(1-\beta_i) \quad \text{and} \quad \beta_i = (\lambda_i-1)/(\lambda_i+1) \tag{10.5}$$

and it follows that Equation (10.4) has a unique Hermitian solution if and only if $\beta_i+\bar{\beta}_j \neq 0$ for all i and j, that is $\lambda_i\bar{\lambda}_j \neq 1$ for all i and j. In particular, corresponding to the case where $\mathbf{A}_1$ is stable and $\mathbf{C}_1$ is non-negative definite, if $|\lambda_i| < 1$ for all i, so that $\mathbf{A}$ is convergent, and $\mathbf{C}$ is non-negative definite then $\mathbf{X}$ is also non-negative definite.

We note also that if the Schur factorization of $\mathbf{A}_1$ is

$$\mathbf{A}_1 = \mathbf{Q}\mathbf{S}_1\mathbf{Q}^H \tag{10.6}$$

then the Schur factorization of $\mathbf{A}$ is given by

$$\mathbf{A} = \mathbf{Q}(\mathbf{I}-\mathbf{S}_1)^{-1}(\mathbf{I}+\mathbf{S}_1)\mathbf{Q}^H. \tag{10.7}$$

The numerical solution of the discrete-time Lyapunov equation follows similar lines to that of the continuous-time equation and we give just a brief description of the main results. We shall use the same notation as for the continuous-time equation.

From the Schur factorization of A and corresponding to Equation (2.3), Equation (10.4) gives

$$\mathbf{S}^H\tilde{\mathbf{X}}\mathbf{S}-\tilde{\mathbf{X}} = -\tilde{\mathbf{C}} \tag{10.8}$$

and hence corresponding to Equations (2.5)–(2.7) we have

$$\tilde{x}_{11} = \tilde{c}_{11}/[(1-|\lambda_1|)(1+|\lambda_1|)] \tag{10.9}$$

$$(\lambda_1\mathbf{S}_1^H-\mathbf{I})\tilde{\mathbf{x}} = -\tilde{\mathbf{c}}-\lambda_1\tilde{x}_{11}\mathbf{s} \tag{10.10}$$

$$S_1^H X_1 S_1 - X_1 = -C_1 - \tilde{x}_{11} s s^H - [s(S_1^H \tilde{x})^H + (S_1^H \tilde{x}) s^H]. \qquad (10.11)$$

For the real Schur factorization of Equation (2.10) these equations are replaced by

$$s_{11}^T \tilde{x}_{11} s_{11} - \tilde{x}_{11} = -\tilde{c}_{11} \qquad (10.12)$$

$$S_1^T \tilde{x} s_{11} - \tilde{x} = -\tilde{c} - s\tilde{x}_{11} s_{11} \qquad (10.13)$$

$$S_1^T X_1 S_1 - X_1 = -C_1 - s\tilde{x}_{11} s^T - [s(S_1^T \tilde{x})^T + (S_1^T \tilde{x}) s^T]. \qquad (10.14)$$

For the positive definite case we assume that the eigenvalues of A are such that $|\lambda_i| < 1$ for all i, so that A is convergent. Then with the substitutions of Equations (5.2) and (5.3) and with some algebraic manipulation, Equations (10.9)–(10.11) become

$$\tilde{u}_{11} = |\tilde{r}_{11}|/[(1-|\lambda_1|)(1+|\lambda_1|)]^{\frac{1}{2}} \qquad (10.15)$$

$$(\lambda_1 S_1^H - I)\tilde{u} = -\alpha\tilde{r} - \lambda_1 \tilde{u}_{11} s, \qquad \alpha = \tilde{r}_{11}/\tilde{u}_{11} \qquad (10.16)$$

$$S_1^H(U_1^H U_1) S_1 - U_1^H U_1 = -\hat{R}^H \hat{R}, \qquad (10.17)$$

where

$$\hat{R}^H \hat{R} = R_1^H R_1 + y y^H, \qquad y = \bar{\alpha}(S_1^H \tilde{u} + \tilde{u}_{11} s) - \bar{\lambda}_1 r. \qquad (10.18)$$

For the real Schur factorization these equations can be replaced by

$$s_{11}^T(\tilde{u}_{11}^T \tilde{u}_{11}) s_{11} - \tilde{u}_{11}^T \tilde{u}_{11} = -\tilde{r}_{11}^T r_{11} \qquad (10.19)$$

$$S_1^T \tilde{u}(\tilde{u}_{11} s_{11} \tilde{u}_{11}^{-1}) - \tilde{u} = -\tilde{r}\alpha - s\tilde{u}_{11}^T(\tilde{u}_{11} s_{11} \tilde{u}_{11}^{-1}), \qquad \alpha = \tilde{r}_{11}\tilde{u}_{11}^{-1} \qquad (10.20)$$

$$S_1^T(U_1^T U_1) S_1 - U_1^T U = -\hat{R}^T \hat{R} \qquad (10.21)$$

where

$$\hat{R}^T \hat{R} = R_1^T R_1 + y y^T, \qquad y = [\tilde{r} S_1^T \tilde{u} + s\tilde{u}_{11}^T]. \qquad (10.22)$$

In the scalar case Equation (10.22) can be replaced by Equation (10.18), with $\bar{\alpha} = \alpha$, $\bar{\lambda}_1 = \lambda_1$ and $S_1^H = S_1^T$, but in the two by two case y consists of four columns rather than the hoped-for two columns. It is an open question as to whether or not a matrix y containing just two columns can be found.

As with the continuous-time case care is needed in the solution of Equations (10.19) and (10.20). Here, corresponding to Equations (6.13)–(6.14), we have

$$v_1 = p_1/\tilde{\alpha}, \quad \tilde{\alpha} = [(1-|\beta|)(1+|\beta|)]^{\frac{1}{2}} \qquad (10.23)$$

$$v_2 = (\tilde{\alpha} p_2 + \beta v_1 t)/[(1-\beta)(1+\beta)] \qquad (10.24)$$

$$v_3 = (p_3^2 + |y|^2)^{\frac{1}{2}}/\tilde{\alpha}, \qquad (10.25)$$

where

$$y = [(\bar{\beta} - \beta) p_2 - p_1 t]/[(1-\beta)(1+\beta)],$$

and corresponding to Equation (6.21) we have

$$\delta = \beta y/v_3, \qquad \eta = p_3/v_3, \qquad \gamma = -\tilde{\alpha} y/v_3. \qquad (10.26)$$

11. The Implicit Lyapunov Equation

Finally, we note that similar methods may be applied to the continuous-time implicit Lyapunov equation

$$\mathbf{A}^H\mathbf{X}\mathbf{B}+\mathbf{B}^H\mathbf{X}\mathbf{A} = -\mathbf{C}, \qquad \mathbf{C} = \mathbf{C}^H, \tag{11.1}$$

where **A** and **B** are square. The solution of the more general implicit Sylvester equation $\mathbf{AXD}+\mathbf{BXC} = \mathbf{E}$ is discussed in Golub *et al.* (1979) and in Epton (1980). It can be shown that Equation (11.1) has a unique Hermitian solution, **X**, if and only if $\lambda_i+\bar{\lambda}_j \neq 0$ for all i and j, where λ_i is an eigenvalue of the generalized eigenvalue problem

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{B}\mathbf{x}. \tag{11.2}$$

When **B** is non-singular Equation (11.1) can be transformed into the explicit equation

$$(\mathbf{A}\mathbf{B}^{-1})^H\mathbf{X}+\mathbf{X}(\mathbf{A}\mathbf{B}^{-1}) = -\mathbf{B}^{-H}\mathbf{C}\mathbf{B}^{-1}, \tag{11.3}$$

but, as with the eigenvalue problem of Equation (11.2), for numerical stability it is generally better to use the implicit form.

In place of the Schur factorization we use the Stewart factorization of A and B (Stewart, 1972; Moler & Stewart, 1973) given by

$$\mathrm{A} = \mathrm{QSZ}^H, \qquad \mathrm{B} = \mathrm{QTZ}^H, \tag{11.4}$$

where Q and Z are unitary and S and T are upper triangular. If we put

$$\tilde{\mathbf{X}} = \mathbf{Q}^H\mathbf{X}\mathbf{Q} \quad \text{and} \quad \tilde{\mathbf{C}} = \mathbf{Z}^H\mathbf{C}\mathbf{Z} \tag{11.5}$$

then Equation (11.1) becomes

$$\mathbf{S}^H\tilde{\mathbf{X}}\mathbf{T}+\mathbf{T}^H\tilde{\mathbf{X}}\mathbf{S} = -\tilde{\mathbf{C}} \tag{11.6}$$

which corresponds to Equation (2.3) for the explicit Lyapunov equation. Partitioning S and T as

$$\mathbf{S} = \begin{pmatrix} \alpha_1 & \mathbf{s}^H \\ 0 & \mathbf{S}_1 \end{pmatrix}, \qquad \mathbf{T} = \begin{pmatrix} \beta_1 & \mathbf{t}^H \\ 0 & \mathbf{T}_1 \end{pmatrix} \tag{11.7}$$

then, corresponding to Equations (2.5)–(2.7), we get

$$\tilde{x}_{11} = -\tilde{c}_{11}/(\alpha_1\bar{\beta}_1+\bar{\alpha}_1\beta_1) \tag{11.8}$$

$$(\beta_1\mathbf{S}_1^H+\alpha_1\mathbf{T}_1^H)\tilde{\mathbf{x}} = -\tilde{\mathbf{c}}-\tilde{x}_{11}(\alpha_1\mathbf{t}+\beta_1\mathbf{s}) \tag{11.9}$$

$$\mathbf{S}_1^H\mathbf{X}_1\mathbf{T}_1+\mathbf{T}_1^H\mathbf{X}_1\mathbf{S}_1 = -\mathbf{C}_1-\tilde{x}_{11}(\mathbf{s}\mathbf{t}^H+\mathbf{t}\mathbf{s}^H)-(\mathbf{S}_1^H\tilde{\mathbf{x}}\mathbf{t}^H+\mathbf{t}\tilde{\mathbf{x}}\mathbf{S}_1)-(\mathbf{T}_1^H\tilde{\mathbf{x}}\mathbf{s}^H+\mathbf{s}\tilde{\mathbf{x}}^H\mathbf{T}_1) \tag{11.10}$$

and corresponding to Equations (5.13)–(5.16) for the positive definite case we get

$$\tilde{u}_{11} = |\tilde{r}_{11}|/[-(\alpha_1\bar{\beta}_1+\bar{\alpha}_1\beta_1)]^{\frac{1}{2}} \tag{11.11}$$

$$(\beta_1\mathbf{S}_1^H+\alpha_1\mathbf{T}_1^H)\tilde{\mathbf{u}} = \alpha\tilde{\mathbf{r}}-\tilde{u}_{11}(\alpha_1\mathbf{t}+\beta_1\mathbf{s}), \quad \alpha = \tilde{r}_{11}/\tilde{u}_{11} \tag{11.12}$$

$$\mathbf{S}_1^H(\mathbf{U}_1^H\mathbf{U}_1)\mathbf{T}_1+\mathbf{T}_1^H(\mathbf{U}_1^H\mathbf{U}_1)\mathbf{S}_1 = -(\mathbf{R}_1^H\mathbf{R}_1+\mathbf{y}\mathbf{y}^H), \tag{11.13}$$

where

$$\mathbf{y} = (\bar{\beta}_1\mathbf{v}-\bar{\alpha}_1\mathbf{w})/|\alpha|, \qquad \mathbf{v} = \mathbf{S}_1^H\tilde{\mathbf{u}}+\tilde{u}_{11}\mathbf{s}, \qquad \mathbf{w} = \mathbf{T}_1^H\tilde{\mathbf{u}}+\tilde{u}_{11}\mathbf{t}.$$

Notice that Equation (11.13) is still only a rank-one update.

I am grateful to D. W. Martin, M. J. D. Powell and Hilary J. Symm who made a number of helpful suggestions to improve the paper. Particular thanks go to P. B. Kronheimer who programmed and tested the real continuous-time positive definite case and offered a number of useful ideas and corrections.

References

Barnett, S. 1975 *Introduction to Mathematical Control Theory*. Oxford: Oxford University Press.

Barnett, S. & Storey, C. 1968 Some applications of the Lyapunov matrix equation. *J. Inst. Maths Applics* **4**, 33–42.

Bartels, R. H. & Stewart, G. W. 1972 Solution of the matrix equation $AX+XB = C$. *Communs Ass. comput. Mach.* **15**, 820–826.

Belanger, P. R. & McGillivray, T. P. 1976 Computational experience with the solution of the matrix Lyapunov equation. *IEEE Trans. Autom. Control* **AC-21**, 799–800.

Bucy, R. S. 1975 Structural stability for the Riccati equation. *SIAM J. Control* **13**, 749–753.

Cline, A. K., Moler, C. B., Stewart, G. W. & Wilkinson, J. H. 1979 An estimate for the condition number of a matrix. *SIAM J. Num. Analysis* **16**, 368–375.

Dongarra, J. J., Moler, C. B., Bunch, J. R. & Stewart, G. W. 1979 *Linpack Users' Guide*. Philadelphia: SIAM.

Epton, M. A. 1980 Methods for the solution of $AXD-BXC = E$ and its application in the numerical solution of implicit ordinary differential equations. *BIT* **20**, 341–345.

Forsythe, G. E. & Moler, C. B. 1967 *Computer Solution of Linear Algebraic Equations*. Englewood Cliffs, N.J.: Prentice-Hall.

Galeone, L. & Peluso, R. I. 1979 Un metodo iterativo per l'equazione matriciale $AX+XB^T = C$. *Boll. Un. Mat. Ital.* **16**, 1–8.

Gantmacher, F. R. 1959 *Applications of the Theory of Matrices*. London: Interscience.

Gill, P. E., Golub, G. H., Murray, W. & Saunders, M. A. 1974 Methods for modifying matrix factorizations. *Maths Comput.* **28**, 505–535.

Gill, P. E. & Murray, W. 1974 *Numerical Methods for Constrained Optimization*. London: Academic Press.

Gill, P. E. & Murray, W. 1977 Modification of matrix factorizations after a rank-one change. In *The State of the Art in Numerical Analysis* (Jacobs, D. A. H., Ed.). London: Academic Press.

Givens, W. 1961 Elementary divisors and some properties of the Lyapunov mapping $X \Rightarrow AX+XA^*$. *Report ANL–6456*, Argonne National Laboratory.

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

Golub, G. H., Nash, S. & Van Loan, C. 1979 A Hessenberg–Schur method for the problem $AX+XB = C$. *IEEE Trans. Autom. Control* **AC-24**, 909–913.

Hagander, P. 1972 Numerical solution of $A^TS+SA+Q = 0$. *Inform. Sci.* **4**, 35–40.

Hammarling, S. & Wilkinson, J. H. 1980 On linear systems arising from finite difference approximations to elliptic differential equations. *NPL Report DNACS* 34/80, National Physical Laboratory, Teddington, Middlesex.

Hoskins, W. D., Meek, D. S. & Walton, D. J. 1977 The numerical solution of the matrix equation $XA+AY = F$. *BIT* **17**, 184–190.

Laub, A. J. 1979 A Schur method for solving algebraic Riccati equations. *IEEE Trans. Autom. Control* **AC-24**, 913–921.

Laub, A. J. 1980 Survey of computational methods in control theory. In *Electric Power Problems: the Mathematical Challenge* (Erisman, A., *et al.*, Eds.). Philadelphia: SIAM.

Lawson, C. L. & Hanson, R. J. 1974 *Solving Least Squares Problems*. Englewood Cliffs, N.J.: Prentice-Hall.

Moler, C. B. & Stewart, G. W. 1973 An algorithm for generalized matrix eigenvalue problems. *SIAM J. Num. Analysis* **10**, 241–256.

O'Leary, D. P. 1980 Estimating conditions numbers. *SIAM J. Sci. Stat. Comput.* **1**, 205–209.

Pace, I. C. & Barnett, S. 1972 Comparison of numerical methods for solving Liapunov matrix equations. *Int. J. Control* **15**, 907–915.

Paige, C. C. 1980 Error analysis of some techniques for updating orthogonal decompositions. *Maths Comput.* **34**, 465–471.

Rothschild, D. & Jameson, A. 1970 Comparison of four numerical algorithms for solving the Lyapunov matrix equation. *Int. J. Control* **11**, 181–198.

Saunders, M. A. 1972 Large scale linear programming using the Cholesky factorization. Report *STAN-CS*-72-252, Stanford University.

Sima, V. 1980 Comparison of some algorithms for solving the Lyapunov-type equations. *Rev. Roum. Sci. Techn.-Electrotechn. Energy* **25**, 235–246.

Snyders, J. & Zakai, M. 1970 On nonnegative solutions of the equation $AD + DA' = -C$. *SIAM J. Appl. Math.* **18**, 704–714.

Stewart, G. W. 1972 On the sensitivity of the eigenvalue problem $Ax = \lambda Bx$. *SIAM J. Num. Analysis* **9**, 669–686.

Stewart, G. W. 1979 The effects of rounding error on an algorithm for downdating a Cholesky factorization. *J. Inst. Maths Applics* **23**, 203–213.

Wilkinson, J. H. 1965 *The Algebraic Eigenvalue Problem*. London: Oxford University Press.

Wilkinson, J. H. 1977 Some recent advances in numerical linear algebra. In *The State of the Art in Numerical Analysis* (Jacobs, D. A. H., Ed.). London: Academic Press.