Identification of Large Space Structures: A Factorization Approach

Trevor Williams*
Kingston Polytechnic, Kingston upon Thames, England

Reliable techniques for on-orbit identification of the vibrational dynamics of large space structures are important chiefly for two reasons. First, the complexity of such vehicles makes accurate finite-element modeling difficult; second, extremely limited ground testing possible for these structures prevents correction of modeling errors before flight. This paper describes a new algorithm that uses in-flight vibration data to estimate directly a factored form for the positive definite symmetric mass matrix of the structure. Applications of matrix factorization methods to, for instance, Kalman filtering and least-squares problems suggest greatly improved accuracy over methods that estimate the mass matrix itself. The properties of the new identification algorithm are analyzed in detail here and illustrated by means of examples.

Introduction

TWO factors combine to make essential the reliable on-orbit identification of the vibrational dynamics of large space structures (LSS). One is the inherent complexity of such vehicles, making accurate finite-element modeling of them difficult. The second is the extremely limited ground testing possible with these structures, which prevents the correction of modeling errors before flight.

This identification problem is complicated further because the typical LSS has many closely spaced, little-known natural frequencies. It is thus difficult to apply forces to the structure at a frequency that excites just one specified structural resonance at a time, so any method based on mode-by-mode identification cannot be used here with any success. This rules out many of the standard structural identification techniques studied extensively in the past.

Consequently, various authors have recently proposed[1-4] methods that identify the parameters of some mathematical model for the dynamics of the LSS considered, rather than identifying its modal description directly. (The corresponding natural frequencies, damping ratios, and mode shapes are then obtainable as the solution of an eigenstructure problem.) In the case of the Eigensystem Realization Algorithm,[1] the model considered is a state space (i.e., first-order) one, whereas, in Refs. 2-4, it is a standard second-order structural model given in terms of mass, stiffness, and damping matrices. All these methods use some form of least-squares estimation to identify the relevant matrices from measured vibration data.

The object of this paper is to show how *matrix factorization* methods (sometimes referred to as *matrix square root* methods) can be applied to the LSS identification problem in a novel way, yielding a new algorithm with significantly improved accuracy. Specifically, the new technique directly estimates a square root of the mass matrix rather than this matrix itself. The great value of matrix factorization methods has been amply demonstrated in aerospace applications such as the Apollo project navigation software and the Jet Propulsion Laboratory orbit ephemeris correction software employed on planetary missions. In fact, use of factorization techniques for the Kalman filtering and least-squares estimation problems encountered there yielded algorithms typically[5] as accurate when run in single precision as standard methods run in double precision. Similar gains in accuracy motivated the matrix square root Lyapunov equation solver.[6] These accuracy improvements, essentially due to the reduced dynamic range of matrix square roots, are especially important in onboard computations, where the computer word length available tends to be limited.

A further advantage of the matrix factorization identification method studied here is that it is guaranteed to result in a symmetric estimate for the mass matrix: by contrast, sensor measurement noise and/or computer rounding error generally prevent a direct estimate for this matrix from being symmetric.

In this paper, details are given of an implementation of the square root identification algorithm that uses the numerically reliable method of *QR decomposition*[8,12] for the relevant least-squares estimation problem. The properties of the new algorithm are analyzed and described in detail and are illustrated by means of examples.

Problem Formulation

Consider the fully general n-mode model

$$M\ddot{q}+(C+G)\dot{q}+(K+H)q=u \tag{1}$$

for the structural dynamics of an LSS, where q is the vector of generalized coordinates, u that of generalized applied forces, and the matrices shown have the following structure and terminology:[7]

$$\begin{aligned} &M=M^T>0, && \text{the mass matrix}\\ &C=C^T\geq 0, && \text{the damping matrix}\\ &G=-G^T, && \text{the gyroscopic matrix}\\ &K=K^T, && \text{the stiffness matrix}\\ &H=-H^T, && \text{the circulatory matrix} \end{aligned} \tag{2}$$

In what follows, we simplify the notation by writing Eq. (1) as

$$M\ddot{q}+\bar{C}\dot{q}+\bar{K}q=u \tag{3}$$

where $\bar{C}$ and $\bar{K}$ have no particular structure unless we restrict ourselves to special cases, for instance nongyroscopic systems with $G=0$. For any system though, $C=\frac{1}{2}(\bar{C}+\bar{C}^T)$, $G=\frac{1}{2}(\bar{C}-\bar{C}^T)$, etc.

Received Feb. 19, 1986; revision received Jan. 5, 1987.

*Senior Research Associate, School of Computing. Member AIAA.

The identification problem as formulated in Refs. 2-4 is to estimate M, $\bar{C}$, and $\bar{K}$ from measurements of forces applied to the structure and the vibrations that result. In summary, all three methods define the data matrices

$$X=\begin{bmatrix} \ddot{q}^T(t_1) & \dot{q}^T(t_1) & q^T(t_1) \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \ddot{q}^T(t_k) & \dot{q}^T(t_k) & q^T(t_k) \end{bmatrix} \text{ and } U=\begin{bmatrix} u^T(t_1) \\ \cdot \\ \cdot \\ \cdot \\ u^T(t_k) \end{bmatrix} \quad (4)$$

from measurements taken at $k \geq 3n$ times $\{t_i: i=1,\ldots,k\}$. Transposing Eq. (3) then yields the overdetermined system of linear equations

$$X\begin{bmatrix} M^T \\ \bar{C}^T \\ \bar{K}^T \end{bmatrix} = U \quad (5)$$

which can be solved a column at a time (giving M, $\bar{C}$, and $\bar{K}$ row by row) by least-squares estimation. Note that this solution could in theory be found by the classical method of the normal equations as $(X^TX)^{-1}X^TU$; this method suffers, however, from severe and well-documented[8,12] accuracy problems. To be specific, its computed value for the least-squares solution generally contains errors proportional to the error amplification factor encountered in inverting X^TX, i.e., its spectral condition number $K_2(X^TX)$, whereas the inherent sensitivity of the least-squares problem to rounding errors in the data is only proportional to the condition number of X. Since $K_2(X^TX)=K_2(X)^2$, unacceptable errors are introduced into the solution as computed using the normal equations.[8] Numerically reliable least-squares techniques do exist; they are based on QR decomposition and the Singular Value Decomposition.[8,12] The former, which is less expensive computationally, is used here; it is based on the fact that, for any X, there exists an orthogonal Q (i.e., a real matrix such that $Q^TQ=I$) for which $Q^TX=R$, with R upper triangular. Applying this decomposition to Eq. (5) gives

$$R\begin{bmatrix} M^T \\ \bar{C}^T \\ \bar{K}^T \end{bmatrix} = Q^TU \quad (6)$$

the solution to which is easily obtained by a process of back-substitution. The errors in this computed least-squares estimate can be shown to be proportional to $K_2(R)=K_2(X)$, i.e., as small as can be expected, given the inherent sensitivity of the problem.

The preceding comments show the importance of generating data for which X is not only of full rank but also has columns as nearly orthogonal as possible, i.e., for which $K_2(X)$ is low. This requirement has important practical implications for the choices of excitation profile $u(t)$ and sampling instants $\{t_i\}$; see Refs. 2-4 for more detailed discussions.

Matrix Factorization Methods

In the general case, the only matrix in Eq. (3) with any specified structure is M, which is positive definite symmetric. This structure guarantees the existence of various factorizations for it: in particular,

$$M=\bar{L}\bar{L}^T \quad (7)$$

where the lower triangular $\bar{L}$ is the Cholesky factor (or square root) of M, and

$$M=LDL^T \quad (8)$$

where L is unit lower triangular and $D=\text{diag}\,(d_i)$ with all d_i positive. The Cholesky decomposition (7) is less efficient to compute than the L-D factorization (8), involving as it does the evaluation of n square roots. This follows from the relation $\bar{L}=L\,\text{diag}(\sqrt{d_i})$.

The aim of what follows is to reformulate the identification scheme (6) in such a way as to generate estimates for L and D (or $\bar{L}$) directly rather than for M. Applications such as Kalman filtering[5] and solving the continuous Lyapunov equation[6] in stability theory, both problems in which the desired solution is a positive definite symmetric matrix Y, have amply demonstrated the far greater accuracy obtainable by operating exclusively on the Cholesky or L-D factorization of Y instead of on Y itself. Similarly, the QR decomposition approach to the least-squares problem outlined previously can itself be seen to be a matrix-square-root technique: R^T is just the Cholesky factor of X^TX.

A general rule is that the results obtained using such matrix factorization methods are as accurate when calculated in single precision as standard ones (based on Y) in double precision. This is essentially due to the lower dynamic range of matrix factors: if Y has elements with sizes in the range 10^{-a}-10^{+a}, its Cholesky factor will have elements in roughly the range $10^{-a/2}$-$10^{+a/2}$. (Very similar accuracy considerations motivate the extremely important problem of matrix scaling, or preconditioning, where the aim is to reduce the dynamic range of a given matrix as much as possible before, for instance, calculating its eigenstructure. See Ref. 9 for a recent survey of this subject.) These accuracy improvements are particularly important when operating on a computer with quite limited word length, as is typically the case in onboard calculations.

A simple example will serve to illustrate these dynamic range considerations. Let

$$Y=\begin{pmatrix} 1 & 10^4 \\ 10^4 & 10^8+1 \end{pmatrix}$$

with Cholesky factor

$$\bar{L}=\begin{pmatrix} 1 & 0 \\ 10^4 & 1 \end{pmatrix}$$

and L-D factorization $L=\bar{L}$, $D=I_2$.

Clearly, although the elements of Y vary in size by a factor of (10^8+1), those of $\bar{L}$ and L vary by only 10^4. The practical importance of this is apparent if we consider storing these matrices in single precision on a typical machine, i.e., to seven-figure accuracy. Either of the factored forms would then be stored exactly, while Y would be rounded to the matrix

$$\tilde{Y}=\begin{pmatrix} 1 & 10^4 \\ 10^4 & 10^8 \end{pmatrix}$$

which is, unlike Y, exactly singular.

This example also illustrates another general point, namely, that finding Y and then calculating its Cholesky or L-D factorization does not achieve the accuracy gains of a true (i.e., direct) matrix-square-root method. For, if Y were calculated in seven-figure arithmetic and its Cholesky factor then found, the factor obtained would actually be that of the rounded $\tilde{Y}$, i.e.,

$$\tilde{L}=\begin{pmatrix} 1 & 0 \\ 10^4 & 0 \end{pmatrix}$$

The numerical damage has already been done.

A further advantage of matrix factorization methods is that the estimates they yield for the matrix Y are guaranteed symmetric, which follows from the obvious symmetry of the products $\bar{L}\bar{L}^T$ and LDL^T. By contrast, rounding and/or measurement errors generally prevent Y from being symmetric, or indeed positive definite, if it is estimated directly. (The behavior of matrix-square-root identification techniques in

the face of errors in the data matrices X and U is quite different from that of standard methods; this will be treated in greater detail later in the paper.)

Matrix factorization techniques can produce striking accuracy gains when applied to problems involving the inverse of the factored matrix. An example of such a problem in the LSS context is the very important one of finding the natural frequencies and damping ratios of Eq. (3). These are given by the eigenvalues of [7]

$$\begin{pmatrix} -M^{-1}\bar{C} & -M^{-1}\bar{K} \\ I & 0 \end{pmatrix}$$

or

$$\begin{pmatrix} -\bar{L}^{-1}\bar{C}\bar{L}^{-T} & -\bar{L}^{-1}\bar{K}\bar{L}^{-T} \\ I & 0 \end{pmatrix}$$

These formulations are equivalent mathematically, but the second tends to yield far more accurate results than the first in digital computation, even if M and $\bar{L}$ are both known exactly. The main reason is that the condition number of M, i.e., the error amplification factor expected when inverting this matrix, is precisely[8] the square of that of $\bar{L}$ (cf. the remarks of the last section on the normal equations). As a rule then, inverting M corrupts twice as many figures in the computed matrix as inverting $\bar{L}$ does. Even this may underestimate the superiority of the matrix factor formulation though, as Wilkinson[10] has shown that, triangular matrices such as $\bar{L}$ can often be inverted very accurately, even when they have moderately large condition numbers. (The LSS eigenstructure problem will be dealt with in greater detail later in the paper.)

Identification Algorithm

The method we derive here is based on the L-D factorization (8) for M rather than on its Cholesky factorization, so as to take advantage of the superior efficiency of this square-root-free form. Premultiplying Eq. (3) by L^{-1} gives

$$DL^T\ddot{q}+L^{-1}\bar{C}\dot{q}+L^{-1}\bar{K}q=L^{-1}u \tag{9}$$

so that the least-sqaures problem (5) becomes

$$X\begin{bmatrix} LD \\ \bar{C}^TL^{-T} \\ \bar{K}^TL^{-T} \end{bmatrix}=UL^{-T} \tag{10}$$

with X and U as defined in Eq. (4). (Note that determining the product LD specifies L and D uniquely since L is unit-triangular.) We again use the numerically reliable method based on the QR decomposition of X to solve this least-squares problem, giving

$$R\begin{bmatrix} LD \\ \bar{C}^TL^{-T} \\ \bar{K}^TL^{-T} \end{bmatrix}=VL^{-T} \tag{11}$$

where $V=Q^TU$. Because $R=Q^TX$ is upper-triangular, the solution is easily found by a process of back substitution.

There is, however, an obvious difficulty that must be overcome before we can use Eq. (11) to estimate L, D, etc.: that the right-hand side itself depends on the unknown L. Fortunately, the unit-lower-triangular structure of L can, if sufficient care is taken, be exploited to overcome this problem and can allow LD, $\bar{C}^TL^{-T}$, and $\bar{K}^TL^{-T}$ to be estimated column by column. To demonstrate this, write the right-hand side of Eq. (11) as $\hat{V}$, the solution of the triangular system of linear equations

$$\hat{V}L^T=V \tag{12}$$

Now, assume that we have been able to find the first $(j-1)$ columns of $\hat{V}$ and L. If we can calculate column j of $\hat{V}$ from these known columns and the initially specified V, we can use it in Eq. (11) to estimate the jth column of LD, etc., thus allowing the identification algorithm to proceed. (The unit-upper-triangular structure of L^T guarantees that the first column of $\hat{V}$ is precisely equal to that of V, so that identification of the first column of LD, etc., is certainly possible.)

Consider then $j\geq 2$, and partition V, $\hat{V}$, and L as

$$V=(V_{jj},v_j,x), \qquad \hat{V}=(\hat{V}_{jj},\hat{v}_j,x)$$

and

$$L=\begin{bmatrix} L_{jj} & \mathbf{0} & 0 \\ l_j & 1 & \mathbf{0} \\ x & x & x \end{bmatrix} \tag{13}$$

where V_{jj}, $\hat{V}_{jj}$, and L_{jj} have $(j-1)$ columns and x denotes an irrelevant and/or unknown quantity. Substituting into Eq. (12) and considering column j then gives $\hat{V}_{jj}l_j+\hat{v}_j=v_j$, or

$$\hat{v}_j=v_j-\hat{V}_{jj}l_j \tag{14}$$

This simple expression for the next column of $\hat{V}$, essentially a process of forward elimination,[8] is the key to being able to identify the factorized system parameters L, D, $\bar{C}$, and $\bar{K}$ as desired.

The identification algorithm can be made significantly more efficient by exploiting the fact that the upper-triangular part of LD is known to be exactly zero and so need not be identified. Examination of Eq. (11) shows that setting the first $(j-1)$ elements of column j of LD equal to zero in this manner amounts to discarding the first $(j-1)$ columns of R when estimating this vector. This results in the upper trapezoidal system of equations

$$\begin{bmatrix} x \\ \hat{R}_{jj} \end{bmatrix}\begin{bmatrix} \hat{l}_jd_j \\ \hat{c}_j \\ \hat{k}_j \end{bmatrix}=\hat{v}_j \tag{15}$$

where $\hat{l}_j$ denotes the nontrivial part of column j of L, $\hat{c}_j$ and $\hat{k}_j$ are the jth columns of $\bar{C}^TL^{-T}$ and $\bar{K}^TL^{-T}$, respectively, and $\hat{R}_{jj}$ is the trailing principal submatrix of R obtained by omitting the first $(j-1)$ rows and columns. Clearly, the equations represented by $\hat{R}_{jj}$ suffice to fix the solution to Eq. (15). The remaining equations, denoted by x in Eq. (15), are then satisfied automatically—unless the data is noisy (see the next section). Since each column of the solution is estimated from a progressively smaller system of equations, a total operations count reduction of about $3/2\ n^3$ is obtained, which can be considerable for the typical LSS case of large n.

In summary, the new identification algorithm proceeds as follows, once the data matrices X and U have been generated.

Step 1: Find orthogonal Q for which $Q^TX=R$, upper-triangular. Calculate $V=Q^TU$.

Step 2: $j=1,2,\ldots,n$:

a) Find $\hat{v}_j=v_j-\hat{V}_{jj}l_j$, the jth column of $\hat{V}=VL^{-T}$, with notation as in Eq. (13). (Note that $\hat{v}_1=v_1$.)

b) The triangular system

$$\hat{R}_{jj}\begin{bmatrix} \hat{l}_jd_j \\ \hat{c}_j \\ \hat{k}_j \end{bmatrix}=\hat{v}_{j2}$$

yields the jth solution column, where $\hat{v}_{j2}$ consists of elements j through $3n$ of $\hat{v}_j$ and all other notation is as in Eq. (15).

Finally,

Step 3: Calculate $\bar{C}=L(\bar{C}^TL^{-T})^T$, $\bar{K}=L(\bar{K}^TL^{-T})^T$. Note that it is often preferable on the grounds of accuracy not actually to perform Step 3 but instead to deal exclusively with the modified system description $[DL^T, L^{-1}\bar{C}, L^{-1}\bar{K}]$ of Eq. (9); this will be elaborated on in the next section.

The following simple example illustrates the operation of this algorithm. Consider a two-mode nongyroscopic, noncirculatory system with closely spaced natural frequencies $\omega_1=2$ rad/s and $\omega_2=2.5$ rad/s, and low modal damping given by $\xi_1=\xi_2=\frac{1}{2}\%$. Suppose data X and U have been generated and then transformed, as in Step 1, to give

$$R=Q^TX=\begin{bmatrix}1&-1&1&-1&1&-1\\0&1&-1&1&-1&1\\0&0&1&-1&1&-1\\0&0&0&1&-1&1\\0&0&0&0&1&-1\\0&0&0&0&0&1\end{bmatrix}$$

and

$$V=Q^TU=\begin{bmatrix}0.5530&-1.6354\\0.4470&2.4354\\0.3530&-1.4354\\-0.3305&1.4549\\0.3500&-1.4300\\4.7750&6.2050\end{bmatrix}$$

(Note that $K_2(X)=K_2(R)=8.0552$ is quite low, as desired for good vibration data.) The remainder of the algorithm then proceeds as follows, where all results are tabulated exactly.

Step 2: $j=1$:

a) $\hat{v}_1=v_1$.

b) $\hat{R}_{11}=R$, $\hat{v}_{12}=v_1$, so that the first column of the solution is

$$\begin{bmatrix}l_1d_1\\\hat{c}_1\\\hat{k}_1\end{bmatrix}=\left[\begin{array}{c}1.0000\\0.8000\\\hline0.0225\\0.0195\\\hline5.1250\\4.7750\end{array}\right]$$

Clearly then, $d_1=1$, $l_1=(1,0.8)^T$.

$j=2$:

a) $$\hat{v}_2=v_2-\hat{v}_1l_{21}=\begin{bmatrix}-2.0778\\2.0778\\-1.7178\\1.7193\\-1.7100\\2.3850\end{bmatrix}$$

b) $$\left[\begin{array}{c|cc|cc}1&-1&1&-1&1\\\hline0&1&-1&1&-1\\0&0&1&-1&1\\\hline0&0&0&1&-1\\0&0&0&0&1\end{array}\right]\begin{bmatrix}\hat{l}_2d_2\\\hat{c}_2\\\hat{k}_2\end{bmatrix}=\left[\begin{array}{c}2.0778\\\hline-1.7178\\1.7193\\\hline-1.7100\\2.3850\end{array}\right]$$

has solution

$$\left[\begin{array}{c}0.3600\\\hline0.0015\\0.0093\\\hline0.6750\\2.3850\end{array}\right]$$

so $d_2=0.3600$, $l_2=(0,1)^T$, etc.

Step 3:

$$L=\begin{bmatrix}1&0\\0.8&1\end{bmatrix}$$

so

$$\bar{C}=L\begin{bmatrix}0.0225&0.0015\\0.0195&0.0093\end{bmatrix}^T=\begin{bmatrix}0.0225&0.0195\\0.0195&0.0249\end{bmatrix}$$

and

$$\bar{K}=L\begin{bmatrix}5.1250&0.6750\\4.7750&2.3850\end{bmatrix}^T=\begin{bmatrix}5.1250&4.7750\\4.7750&0.2050\end{bmatrix}$$

It is easy to verify that the identified matrices $\bar{C}$ and $\bar{K}$, together with

$$M=LDL^T=\begin{bmatrix}1&0.8\\0.8&1\end{bmatrix}$$

are indeed the correct values for the given data.

Algorithm Properties

We now give a description of some of the attributes of the new identification algorithm, so as to examine further how it compares with more standard methods.

Accuracy

The improvement in accuracy resulting from the reduced dynamic range of matrix factors is illustrated by the following example. Consider a circulatory (i.e., $\bar{K}_0^T\neq\bar{K}$), but undamped and nongyroscopic, system with $n=3$. Since $\bar{C}=0$, we can omit the quantities $\{\dot{q}^T(t_i)\}$ from X as defined in Eq. (4), reducing the dimensions of this matrix to $(k\times 2n)$. With this simplification, suppose that data X and U have been generated and then transformed as in Step 1 to give

$$R=Q^TX=\begin{bmatrix}-10^4&1&0&0&1&1\\0&1&-10^4&-1&0&0\\0&0&1&0&0&0\\0&0&0&1&-1&0\\0&0&0&0&3&1\\0&0&0&0&0&1\end{bmatrix}$$

$$V = Q^T U = \begin{bmatrix} 1 & 2 & 2 \\ -2 & 1 & -10^4 \\ 1 & 10^4 & 2 \\ 1 & -1 & 0 \\ 3 & 3 & 2 \\ 0 & 0 & 2 \end{bmatrix}$$

Assuming that all calculations are done in seven-figure arithmetic, representative of single precision on many computers, the new algorithm produces estimates

$$L = \begin{bmatrix} 1 & 0 & 0 \\ 10^4 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

$$D = I_3 \text{ and } L^{-1}\bar{K} = \begin{bmatrix} 2 & -2\times 10^4 & -2 \\ 1 & 1-10^4 & -1 \\ 0 & 0 & 2 \end{bmatrix}^T$$

It is easy to verify that these identified values for L, D, and

$$\bar{K} = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

are precisely the true values for the given data. By contrast, if a non-square-root algorithm based on Eq. (6) is applied to the same R and V, and if all calculations are again in seven-digit arithmetic, the resulting direct estimate for M is

$$\begin{bmatrix} 1 & 10^4 & 1 \\ 10^4 & 10^8 & 10^4 \\ 1 & 10^4 & 2 \end{bmatrix}$$

This singular matrix is not equal to the (nonsingular) true mass matrix

$$\begin{bmatrix} 1 & 10^4 & 1 \\ 10^4 & 10^8+1 & 10^4 \\ 1 & 10^4 & 2 \end{bmatrix}$$

The corruption results from rounding terms of the form 10^8+1 to 10^8 when estimating column 2 of M^T. In fact, the same rounding occurs if L and D are found exactly and $M=LDL^T$ is then formed in seven-digit arithmetic. This accuracy loss whenever dealing with M explicitly is the reason for the earlier observation that it is usually preferable to operate exclusively on the modified system description $\{LD, L^{-1}\bar{C}, L^{-1}\bar{K}\}$ of Eq. (9) rather than ever actually forming $\{M, \bar{C}, \bar{K}\}$.

As a specific illustration of this point, consider the problem of calculating the eigenvalues of the example system. These can clearly be expressed as $\pm\sqrt{\lambda_i}$, where $\{\lambda_i\}$ are the eigenvalues of $-M^{-1}\bar{K}$. Other mathematically equivalent descriptions for the $\{\lambda_i\}$ are the eigenvalues of $-L^{-1}\bar{K}L^{-T}$, the generalized eigenvalues[8,12] of $\{-\bar{K}, M\}$ (i.e., the solutions of $-\bar{K}x=\lambda Mx$), and the generalized eigenvalues of $\{-L^{-1}\bar{K}, L^T\}$. Since this last factored form minimizes the dynamic range of the matrices considered, it should be expected to give the most accurate results of any of the alternative formulations; that this is actually true is seen from Table 1, where all columns but the last are calculated in seven-figure arithmetic.

The results of the first three methods are clearly all in error to varying extents; most notably, the eigenvalues of $-M^{-1}\bar{K}$ yield a very quickly diverging unstable mode with time constant $1/\sqrt{7.917\times 10^7} = 1.124\times 10^{-4}$ s, whereas all true modes are purely oscillatory. By contrast, the generalized eigenvalues of $\{-L^{-1}\bar{K}, L^T\}$ are all correct. Note that the corresponding preferred formulation for a damped or gyroscopic system is the generalized eigenvalue problem

$$\left\{ \begin{pmatrix} -L^{-1}\bar{C} & -L^{-1}\bar{K} \\ I & 0 \end{pmatrix}, \begin{pmatrix} DL^T & 0 \\ 0 & I \end{pmatrix} \right\}$$

This is expressed entirely in terms of the matrices output directly from the new identification algorithm.

Efficiency

This is measured by the approximate number of floating-point operations required to estimate the $(n\times n)$ matrices L, D, $\bar{C}$, and $\bar{K}$ from k data samples. The operation counts for each stage of the new algorithm are tabulated below.

QR decomposition[8] of X:	$9n^2(k-n)$
Calculation of $V=Q^TU$:	$n^2(6k-9n/2)$
Calculation of $\hat{V}$ by Eq. (14):	$\frac{1}{2}n^2k$
Finding least-squares solution by Eq. (15):	$3n^3$
Calculation of $\bar{C}$ and $\bar{K}$:	n^3
Total:	$\frac{1}{2}n^2(31k-19n)$

(Note that the operation count quoted for calculating V is that obtained if we premultiply U in turn by each of the individual Householder transformations making up Q. Forming the product Q explicitly and then calculating Q^TU would be considerably less efficient; see Ref. 8 for details.)

To provide a meaningful comparison with the new algorithm, we consider the number of operations involved in calculating a matrix square root of M in a different way, namely, estimating M, $\bar{C}$, and $\bar{K}$ from the standard least-squares equation (6) and then computing the Cholesky factor of M. (Of course, the examples already given illustrate that this pseudofactorization method does not attain the accuracy of the new algorithm.) The operation counts for this procedure follow.

QR decomposition of X:	$9n^2(k-n)$
Calculation of $V=Q^TU$:	$n^2(6k-9n/2)$
Finding least-squares solution by Eq. (6):	$9n^3/2$
Cholesky factorization of M:	$n^3/6$
Total:	$n^2(15k-53n/6$

Table 1 $\{\lambda_i\}$ as computed by various methods

ev of $-M^{-1}\bar{K}$	ev of $-L^{-1}\bar{K}L^{-T}$	gev of $\{-\bar{K},M\}$	gev of $\{-L^{-1}\bar{K}, L^T$	Exact
-1.123×10^{-4}	-16.00	-1.000×10^{-8}	-1.000×10^{-8}	-1.000×10^{-8}
-2.000	-2.000	-2.000	-2.000	-2.000
$+7.917\times 10^7$	-2.000×10^8	-2.306×10^{22}	-2.000×10^8	-2.000×10^8

The new algorithm can be seen to require only 2.3% more operations for $k=3n$, and 3.1% more if $=k10n$. Thus, its improved accuracy is not obtained at the expense of any appreciable increase in calculation time. Indeed, if the modified system description $\{DL^T, L^{-1}\bar{C}, L^{-1}\bar{K}\}$ is all that is required, making Step 3 unnecessary, the new algorithm actually becomes more efficient than the standard one for $k<(10/3)n$.

Robustness

Another topic of great practical interest is the effect of data perturbations on the system model estimates produced by the new and standard algorithms. These perturbations in X and U may be due to sensor or actuator noise; they may also be due to the fact that, as in all digital computation, the exact data are rounded when they are stored in the finite word length of a computer. Thus, instead of solving the true least-squares problem (5), we are actually solving a perturbed problem

$$(X+\delta X)\begin{bmatrix}(M+\delta M)^T\\(\bar{C}+\delta\bar{C})^T\\(\bar{K}+\delta\bar{K})^T\end{bmatrix}=U+\delta U \tag{16}$$

Now, while the true mass matrix M is known to be symmetric, there is no reason why the perturbed solution $(M+\delta M)$ should also be, for arbitrary δX and δU. Thus, if we find the least-squares solution to Eq. (16) using the standard QR decomposition method, i.e., by solving

$$(R+\delta R)\begin{bmatrix}(M+\delta M)^T\\(\bar{C}+\delta\bar{C})^T\\(\bar{K}+\delta\bar{K})^T\end{bmatrix}=V+\delta V \tag{17}$$

with R and δR upper-triangular, the resulting mass matrix estimate $M_1=M+\delta M$ is generally not symmetric.

By contrast, the estimated L-D factorization generated by the new algorithm guarantees that the corresponding perturbed mass matrix $M_2=(L+\delta L)(D+\delta D)(L+\delta L)^T$ is symmetric, no matter what the corruption in the data X and U. Clearly then, the locations in M_1 and M_2 that are affected by any particular δR or δV must be quite different. Basically, the perturbations in the symmetric M_2 are restricted to smaller submatrices than those in M_1 although, the perturbations in the estimated factors $(L+\delta L)$ and $(D+\delta D)$ need not be so highly structured. Consider, for instance, the effects of a nonzero δv_{ij}: the back-substitution process used to solve Eq. (17) in the standard method means that this data error will corrupt elements 1-i of column j of M_1^T. In the new algorithm, however, since columns 1 through $(j-1)$ have already been found correctly and M_2 is guaranteed symmetric, δv_{ij} can only corrupt elements j-i of column j (i.e., none, if $i<j$). It can similarly be shown that an error in row i of R will corrupt elements 1-i of all columns of M_1^T, but only these elements of columns 1 through i of M_2. On the other hand, the upper-triangular structure of $(R+\delta R)$ results in the fact that both algorithms always produce the same estimates, $(\bar{C}+\delta\bar{C})$ and $(\bar{K}+\delta\bar{K})$, for $\bar{C}$ and $\bar{K}$ regardless of any differences in their mass matrix estimates.

The following simple example illustrates these points. Consider the data

$$R=\begin{bmatrix}1&-1&-1&0&0&0\\0&1&-1&-1&0&0\\0&0&1&-1&-1&0\\0&0&0&1&-1&-1\\0&0&0&0&1&-1\\0&0&0&0&0&1\end{bmatrix}$$

$$\text{and } V=\begin{bmatrix}1&-1\\0&0\\-2&2\\1&0\\2&-2\\-1&1\end{bmatrix}$$

with true solution

$$M=\begin{pmatrix}2&1\\1&1\end{pmatrix},\ \bar{C}=\begin{pmatrix}0&1\\1&0\end{pmatrix},\ \bar{K}=\begin{pmatrix}1&-1\\-1&1\end{pmatrix}$$

If element 1, 5 of R is now perturbed to ϵ and element 3,2 of V to $(2+\delta)$, the above analysis shows that M_1 should include errors of the form

$$\epsilon\begin{pmatrix}x&0\\x&0\end{pmatrix}+\delta\begin{pmatrix}0&0\\x&x\end{pmatrix}$$

while the symmetric M_2 should only have errors of the form

$$\epsilon\begin{pmatrix}x&0\\0&0\end{pmatrix}+\delta\begin{pmatrix}0&0\\0&x\end{pmatrix}$$

Both algorithms should yield the same estimate for $\bar{C}$, with error

$$\delta\begin{pmatrix}0&0\\x&0\end{pmatrix}$$

and should give $\bar{K}$ exactly.

The actual mass matrix estimate produced by the standard algorithm here is

$$\begin{pmatrix}2-\epsilon&1\\1+\epsilon+2\delta&1+\delta\end{pmatrix}$$

indeed not symmetric, and with errors as predicted. The new algorithm generates estimated factors

$$L+\delta L=\begin{bmatrix}1&0\\\dfrac{1}{2-\epsilon}&1\end{bmatrix},\quad D+\delta D=\begin{bmatrix}2-\epsilon&0\\0&1+\delta-\dfrac{1}{2-\epsilon}\end{bmatrix}$$

These have no obvious structure, but the resulting

$$M_2=\begin{bmatrix}2-\epsilon&1\\1&1+\delta\end{bmatrix}$$

does, with all errors being limited to the diagonal elements as expected. Similarly, both algorithms give

$$\delta\bar{C}=\begin{bmatrix}0&0\\\delta&0\end{bmatrix},\ \delta\bar{K}=0$$

for these data, also as anticipated.

Use of A Priori Data

The identification algorithm as analyzed so far assumes no a priori knowledge of the dynamics of the system considered, i.e., it treats the matrices L, D, $\bar{C}$, and $\bar{K}$ as initially completely unknown. If we do, in fact, know something of the structure of these matrices before launch, this information can be used to simplify some aspect of the on-orbit identification problem addressed by the algorithm. Two specific ways in which this can be done are now described.

The first is based on the observation that the mass matrix of a realistic LSS will often be sparse, with relatively few nonzero

elements, and these in some highly structured pattern. If the corresponding matrix factor L were also sparse, the efficiency of the identification algorithm could be improved by exploiting this fact, i.e., by ignoring all known "hard" zeros in L whenever operating on this matrix. We therefore wish to determine under what conditions the factor L corresponding to a sparse M is itself sparse. Full results are well beyond the scope of this paper, forming as they do the subject of a large part of the book by George and Liu,[11] but one fundamental relation can be given: if M is banded with bandwidth p (i.e., $m_{ij}=0$ for all i, j such that $|i-j|>p$), then[8,11] L is also banded, again with bandwidth p. If M is sparse but not banded, L need not be sparse; however, there exists in general a permutation P (corresponding here to a reordering of sensor and actuator positions) for which $\tilde{M}=PMP$ is banded and so has a banded factorization.

The widely used *reverse Cuthill-McKee algorithm*[11] can be used to calculate P if the positions (but not necessarily the values) of the nonzero elements of M are known. Thus, if we have a priori knowledge of the sparse structure of M, the new identification algorithm can be applied to reordered data to give a banded, and so sparse, L-D factorization for $\tilde{M}$. The approximate operation counts when $\tilde{M}$ and $\tilde{L}$ have bandwidth, p, and this structure is exploited are as follows:

QR decomposition of X:	$9n^2(k-n)$
Calculation of V:	$n^2(6k-9n/2)$
Calculation of $\hat{V}$:	nkp
Finding least-squares solution:	$2n^2(n+p)$
Calculation of $\bar{C}$ and $\bar{K}$:	$2n^2p$
Total:	$n(15kn-23/n^2/2+kp+4np)$

Comparing these with the counts for the general formulation of the algorithm shows a savings of 7.6% for the typical case $p=n/10$ and $k=3n$, which is certainly worthwhile.

The remaining application of a priori knowledge to be considered here allows us to use a reduced number of actuators to excite the structure when generating its on-orbit vibration data. This is a factored equivalent of the method proposed in Ref. 2 and proceeds as follows. It has already been noted that X must be of full rank, $3n$, if the least-squares problem (5) or (10) is to be solvable: by Sylvester's inequality, this implies that rank (U) = rank $(M,\bar{C},\bar{K})=n$, as M is nonsingular. Thus, for full system identification we must use n independent actuators, and the excitation profiles applied to these must be chosen with care so as to ensure that both X and U are of full rank. However, if, as in Ref. 2, a submatrix of M is assumed to be initially known (e.g., to be zero if M is block diagonal, a special case of a sparse structure) and the system is known to be nongyroscopic and noncirculatory, the number of actuators can be reduced accordingly.

To be specific, write

$$M=\begin{pmatrix} M_1 & M_2 \\ M_2^T & M_3 \end{pmatrix}$$

where M_2 is initially known and M_1 is of dimension $(m\times m)$, etc., partition $C=C^T$ and $K=K^T$ accordingly, and write the L-D factorization of M as

$$L=\begin{pmatrix} L_1 & 0 \\ L_2 & L_3 \end{pmatrix} \text{ and } D=\begin{pmatrix} D_1 & 0 \\ 0 & D_2 \end{pmatrix}$$

Now, if only m actuators are used to generate the data of Eq. (4), U reduces to $(U_1, 0)$ with U_1 $(k\times m)$: substituting these partitions into Eq. (10) and considering the first column block gives

$$X\begin{bmatrix} L_1D_1 \\ 0 \\ \hline C_1^TL_1^{-T} \\ C_2^TL_1^{-T} \\ \hline K_1^TL_1^{-T} \\ K_2^TL_1^{-T} \end{bmatrix} = [U_1-X]L_1^{-T} \begin{bmatrix} 0 \\ M_2^T \\ \hline 0 \\ 0 \\ \hline 0 \\ 0 \end{bmatrix} \tag{18}$$

where the known term M_2^T has been taken to the right-hand side. Comparing this with Eq. (10), as originally formulated, shows that this least-squares problem can be solved by means of the forward elimination process of Eq. (14) acting on a modified input matrix, so long as the columns of X that are active on the left-hand side are linearly independent. (Note that this condition can indeed be satisfied with m actuators if sufficient care is taken in selecting excitation profiles.) This determines the unknowns L_1, D_1, C_1, C_2, K_1, and K_2, while L_2 is given from $M_2=L_1D_1L_2^T$. The remaining unknowns L_3, D_2, C_3, and K_3 are found in similar fashion from the second column block of Eq. (10): taking all quantities that are now known to the right-hand side gives

$$X\begin{bmatrix} 0 \\ L_3D_2 \\ \hline 0 \\ C_3^TL_3^{-T} \\ \hline 0 \\ K_3^TL_3^{-T} \end{bmatrix} = [-U_1L_1^{-T}L_2^T-X \begin{bmatrix} 0 \\ 0 \\ \hline (C_2-C_1^TL_1^{-T}L_2^T) \\ -C_2^TL_1^{-T}L_2^T \\ \hline (K_2-K_1^TL_1^{-T}L_2^T) \\ -K_2^TL_1^{-T}L_2^T \end{bmatrix}]L_3^{-T} \tag{19}$$

Again, so long as U_1 has been selected so that the columns of X that are active on the left-hand side are linearly independent, this equation can be solved using the forward elimination process of Eq. (14).

Conclusions

A new algorithm has been presented for the identification of the dynamics of a flexible LSS from in-flight vibration measurements. The novel aspect of this method is that it identifies directly a square root of the mass matrix rather than this matrix itself, thus giving rise to all the well-documented accuracy improvements typical of algorithms based on matrix factorization techniques. The properties of the new algorithm were analyzed in detail and illustrated by means of examples.

Acknowledgments

The author gratefully acknowledges the financial support of the United Kingdom Science and Engineering Research Council, as well as the constructive suggestions of the Associate Editor and an anonymous reviewer.

References

[1]Juang, J.-N. and Pappa, R.S., "An Eigensystem Realization Algorithm for Modal Parameter Identification and Model Reduction," *Journal of Guidance, Control, and Dynamics*, Vol. 8, Sept.-Oct. 1975, pp. 620-527.

[2]Hendricks, S.L. et al., "Identification of Large Flexible Structures Mass/Stiffness and Damping from On-Orbit Experiments," *Journal of Guidance, Control, and Dynamics*, Vol. 7, March-Apr. 1984, pp. 244-245.

[3]Rajaram, S. and Junkins, J.L., "Identification of Vibrating Flexible Structures," *Journal of Guidance, Control, and Dynamics,* Vol. 8, July-Aug. 1985, pp. 463-470.

[4]Craig, R.R. and Blair, M.A., "A Generalized Multiple-Input, Multiple-Output Modal Parameter Estimation Algorithm," *AIAA Journal,* Vol. 23, June 1985, pp. 931-937.

[5]Bierman, G.J., *Factorization Methods for Discrete Sequential Estimation,* Academic Press, New York, 1977.

[6]Hammarling, S.J., "Numerical Solution of the Stable, Non-Negative Definite Lyapunov Equation," *IMA Journal of Numerical Analysis,* Vol. 2, July 1982, pp. 303-323.

[7]Meirovitch, L., *Computational Methods in Structural Dynamics,* Sijthoff and Noordhoff, Alphen aan den Rijn, the Netherlands, 1980.

[8]Golub, G.H. and Van Loan, C.F., *Matrix Computations,* Johns Hopkins University Press, Baltimore, MD, 1983.

[9]Williams, T.W.C., "Obtaining Minimal Gerschgorin Discs by Scaling the States," *International Journal of Control,* Vol. 42, Nov. 1985, pp. 1155-1173.

[10]Wilkinson, J.H., *Rounding Errors in Algebraic Processes,* Her Majesty's Stationery Office, London, England, 1963.

[11]George, A. and Liu, J.W., *Computer Solution of Large Sparse Positive Definite Systems,* Prentice-Hall, Englewood Cliffs, NJ, 1981.

[12]Stewart, G.W., *Introduction to Matrix Computations,* Academic Press, New York, 1973.