Mechanical Systems and Signal Processing 23 (2009) 1759–1772

Contents lists available at ScienceDirect

Mechanical Systems and Signal Processing

journal homepage: www.elsevier.com/locate/jnlabr/ymssp

An optimization technique for damped model updating with measured data satisfying quadratic orthogonality constraint

B.N. Datta [a,1], S. Deng [a], V.O. Sokolov [a,*], D.R. Sarkissian [b]

[a] Northern Illinois University, DeKalb, IL, USA
[b] Nizhny Novgorod State University of Architecture and Civil Engineering, Russia

ARTICLE INFO

Article history:
Received 14 March 2008
Received in revised form
16 July 2008
Accepted 23 July 2008
Available online 14 August 2008

Keywords:
Model updating
Numerical optimization
Vibration

ABSTRACT

A two-stage optimization procedure for matrix updating of a damped finite element model in structural dynamics is proposed. In Stage I, the measured data is updated to satisfy a recently established orthogonality relation between the eigenvectors of a quadratic matrix pencil associated with the model. This updated data is then used in Stage II to update the stiffness matrix so that (i) the updated model reproduces the measured data, and (ii) the symmetry of the original model is preserved. The results of the paper generalize some well known and recent results in several ways. Our contribution also includes mathematically proved results to demonstrate that Stage I is essential to obtain a symmetric updated matrix K in Stage II, unless, of course, the measured eigenvector matrix is such that the orthogonality constraint has been already satisfied, which is extremely unlikely to happen in practice.

1. Introduction

Finite element model updating problem concerns updating an analytical symmetric finite element generated model using measured data from a real-life or an experimental structure. The updating needs to be done in such a way that the symmetry of the model is preserved and the updated model contains some of the desirable physical and structural properties of the original finite element model. The problem routinely arises in vibration industries, including automobile, and air- and space-crafts.

There now exists a large amount of work on direct matrix updating methods. Most of this work prior to 1995 is contained in the book by Friswell and Mottershead [1] and references therein. Some of the more recent results can be found in [2–9]. Most existing methods concern updating of an undamped analytical model of the form:

$$M_a\ddot{x}(t) + K_a x(t) = 0$$

where M_a, K_a are, respectively, the analytical mass and stiffness matrices. For the sake of convenience, we will denote this model simply by (M_a, K_a). Similar notations will be used for the updated model; that is, (M_u, K_u) will stand for the undamped updated model and (M_u, C_u, K_u) will stand for the *damped updated model*

$$M_u\ddot{x} + C_u\dot{x} + K_u x = 0$$

The analytical and updated eigenvalues and eigenvectors will also be denoted in a similar way.

* Corresponding author. Tel.: +18157931428.
E-mail addresses: dattab@math.niu.edu (B.N. Datta), deng@math.niu.edu (S. Deng), sokolov@math.niu.edu (V.O. Sokolov), sarkiss@math.niu.edu (D.R. Sarkissian).
[1] The research of the author has been supported by NSF Grant # DMS-0505784.

0888-3270/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ymssp.2008.07.017

A standard practice is to formulate the updating problem in an optimization setting such that the undamped updated model satisfies the following basic properties of the original model [1]:

(i) $M_u = M_u^T$;
(ii) $K_u = K_u^T$ (symmetry);
(iii) $\Phi_u^T M_u \Phi_u = I$ (orthogonality);
(iv) $K_u \Phi_u = M_u \Phi_u \Lambda_u$ (eigenvalue–eigenvector relation).

Maintaining symmetry and reproduction of the measured data are the basic requirements for model updating. However, satisfaction of the orthogonality relation by the measured data is also of prime importance, because the measured data, which comes from an experiment or a real-life structure, very often fails to satisfy the orthogonality constraint (iii).

Besides these basic requirements, there are also other engineering and computational challenges associated with the updating problem. These include: (i) dealing with incompleteness of the measured data, and (ii) complex measured data versus real analytical data, etc. For details, see book by Friswell and Mottershead [1]. In this paper, we deal only with real representations of the data and assume that either modal expansion or model reduction has been performed to deal with the issue of the incomplete measured data.

There are two types of updating procedures. The first type of methods, assuming the mass matrix as the reference matrix, update first the measured data so that it satisfies the mass-orthogonality constraint (iii). This is then followed by updating the stiffness matrix so as to satisfy the constraints (ii) and (iv). The others update, either separately or simultaneously, the mass and stiffness matrices, satisfying the constraints (i)–(iv) [8,10,11]. There also now exists a method which updates the stiffness matrix first satisfying the constraints (ii) and (iv) and then computes the missing entries of the measured modes in a computational setting such that computed data satisfies the mass orthogonality constraint [2]. The method proposed in [2] has the additional important feature that the eigenvalues and eigenvectors which are not updated remain unchanged by the updating procedure. This guarantees that "no spurious modes appear in the frequency range of interest".

In this paper, we propose a new method of the first type for a damped model. Thus, our method consists of two stages. In Stage I, we update the measured eigenvectors so that they satisfy a quadratic orthogonality relation proved in Corollary 4 of this paper. This result is a real-form generalization of the three orthogonality relations, proved earlier in [12]. In Stage II, the updated measured eigenvectors from Stage I are used to update the stiffness matrix so that it remains symmetric after updating and the measured eigenvalues and eigenvectors are reproduced by the updated model. Thus, our method generalizes methods for undamped models of the first type to a damped model. The results of numerical experiments on some case studies are presented to show the accuracy of the proposed method. Our contribution also includes mathematically established results to show that *satisfaction of the orthogonality relation by the measured data is necessary and sufficient for the solution of Stage II to be symmetric*.

It is to be noted that there are other methods, including the method of Friswell et al. [6], the algorithmic implementation of this method [13], and several control-theoretic methods (e.g. [14–16]) that consider updating of the damped models. For details of these control theoretic methods, see Friswell and Mottershead [1, p. 154]. However those methods do not explicitly update the measured data so as to satisfy any orthogonality constraints. On the other hand, as noted before, our mathematical results of Theorems 2 and 3 demonstrate that Stage I must be performed explicitly or it is to be implicitly assumed that the measured data already satisfies an appropriate orthogonality constraint, before performing Stage II. Otherwise, the feasibility set of Stage II problem might be empty.

In our case the problems in both stages are nonlinear optimization problems. The Stage I problem is a nonconvex minimization problem with equality relations. This is a difficult optimization problem to solve. An augmented Lagrangian method is proposed to deal with this problem. Some convergence properties of this method are discussed.

The Stage II problem is a convex quadratic problem. This is a rather nice optimization problem to deal with and there are several excellent numerical methods for such problems in the literature (see [17]).

Implementations in optimization settings of Stages I and II require that the appropriate gradient formulas must be computed in terms of the known quantities only, which are, in our case, just a few measured eigenvalues and eigenvectors and the corresponding sets from the analytical model. Such gradient formulas have been mathematically derived in the paper.

The following notations are used throughout this paper.

- $\mathbb{R}]^{p\times q}$—the set of all real $p \times q$ matrices;
- $\langle A,B\rangle = \mathrm{trace}(A^TB) = \mathrm{trace}(AB^T) = \sum_{i=1}^{p}\sum_{j=1}^{q} A_{ij}B_{ij}$; where $A,B \in \mathbb{R}^{p\times q}$;
- $\langle A,A\rangle = \|A\|_F^2$;
- A^T—the transpose of A;
- diag(A)—the diagonal matrix with the same diagonal elements as A;
- M_a, C_a, K_a—real symmetric analytical mass, damping, and stiffness matrices, respectively;
- M_u, C_u, K_u—updated real symmetric mass, damping, and stiffness matrices, respectively;
- Λ, Φ—eigenvalue and eigenvector matrices (possibly complex);
- $\mathscr{T}_M, X_M$—real-form representations of the measured eigenvalue and eigenvector matrices.

2. The orthogonality relations for the quadratic pencil

In this section we state three orthogonality relations of the eigenvectors of the quadratic matrix pencil:

$$P(\lambda) = \lambda^2 M + \lambda C + K \tag{1}$$

recently obtained in [12,18]. These relations generalize the well-known orthogonality relation of a symmetric matrix and that of a symmetric positive definite linear pencil (see [19,20]). The real-version of the last relation will be proved in the next section and be used in our work in this paper.

Theorem 1 (*Orthogonality relations of the eigenvectors of quadratic matrix pencil*). *Let all the eigenvalues of* $P(\lambda) = \lambda^2 M + \lambda C + K$, *where* $M = M^T$, $C = C^T$, $K = K^T$, *be distinct and nonzero. Let* $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_{2n})$ *be the matrix of eigenvalues and* $\Phi = (\phi_1, \ldots, \phi_{2n})$ *be the matrix of corresponding eigenvectors of the pencil* $P(\lambda)$, *i.e.* $\det(P(\lambda_i)) = 0,\ P(\lambda_i)\phi_i = 0,\ \forall i$, *then there exist diagonal matrices* D_1, D_2, *and* D_3 *such that*

$$\Lambda \Phi^T M \Phi \Lambda - \Phi^T K \Phi = D_1 \tag{2}$$

$$\Lambda \Phi^T C \Phi \Lambda + \Lambda \Phi^T K \Phi + \Phi^T K \Phi \Lambda = D_2 \tag{3}$$

$$\Lambda \Phi^T M \Phi + \Phi^T M \Phi \Lambda + \Phi^T C \Phi = D_3 \tag{4}$$

Note: (i) When $C = 0$, relation (4) reduces to the well-known orthogonality relation for linear pencil, namely, $\Phi^T M \Phi = D$, where D is a diagonal matrix (possibly complex). If the linear pencil is symmetric and positive definite, then D can be chosen as the real identity matrix.

(ii) The pair (Φ, Λ), $\Phi \in \mathbb{C}^{n\times k}$, $\Lambda \in \mathbb{C}^{k\times k}$ will be called a *matrix eigenpair* of the pencil $P(\lambda)$. It can be either full set of eigenvectors-eigenvalues ($k = 2n$) or just a portion of the spectral data ($k < 2n$). Note that the pair (Φ, Λ) satisfies the following relation:

$$M\Phi\Lambda^2 + C\Phi\Lambda + K\Phi = 0 \tag{5}$$

2.1. Real-form representations of complex numbers, and eigenvalues and eigenvectors

Given a pair of complex conjugate numbers $\alpha_j \pm i\beta_j$, we can associate with it a 2×2 matrix T_j, given by

$$T_j = \begin{pmatrix} \alpha_j & \beta_j \\ -\beta_j & \alpha_j \end{pmatrix}$$

Thus, given a set of k numbers of which l are complex conjugate pairs of the above form, and remaining $k - 2l$ are real, we can associate with these numbers a real block-diagonal matrix $\mathcal{T}$ of the form

$$\mathcal{T} = \mathrm{diag}(T_1, \ldots, T_l, T_{2l+1}, \ldots, T_k) \tag{6}$$

where T_{2l+1} through T_k are real scalars. $\mathcal{T}$ is then a *real-form matrix representation* of these k numbers.

Analogously, for a set of k vectors a real-form matrix representation is given by

$$X = [u_1, v_1, \ldots, u_l, v_l, \phi_{2l+1}, \ldots, \phi_k] \tag{7}$$

Indeed, defining the block diagonal matrix

$$S = \mathrm{diag}(S_1, \ldots, S_l, S_{2l+1}, \ldots, S_k),$$

with

$$S_j = \begin{cases} \dfrac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix}, & j = 1, \ldots, l \\ 1, & j = 2l+1, \ldots, k \end{cases} \tag{8}$$

we can write

$$T = S\Lambda S^{-1}$$

and, similarly, $X = \Phi S^{-1}$, where $\Lambda = \mathrm{diag}(\alpha_1 + i\beta_1, \alpha_1 - i\beta_1, \ldots, \alpha_l + i\beta_l, \alpha_l - i\beta_l, \lambda_{2l+1}, \ldots, \lambda_k)$ and $\Phi = (u_1 + iv_1, u_1 - iv_1, \ldots, u_l + iv_l, u_l - iv_l, \phi_{2l+1}, \ldots, \phi_k)$. Pair $(X, \mathcal{T})$ will be called a real matrix eigenpair of $P(\lambda)$.

Thus if $(X, \mathcal{T})$ is a real matrix eigenpair of $P(\lambda)$, then we have

$$MX\mathcal{T}^2 + CX\mathcal{T} + KX = 0 \tag{9}$$

Also, the distinctiveness of the eigenvalues in Λ implies, that

$$\begin{pmatrix} \Phi \\ \Phi\Lambda \end{pmatrix} \quad \text{is of full rank} \tag{10}$$

This, in turn means, that

$$\begin{pmatrix} X \\ X\mathcal{T} \end{pmatrix} \quad \text{is of full rank} \tag{11}$$

3. Existence of symmetric solution of the model updating problem

In this section, we establish mathematical results to demonstrate the fact that the updated measured data must satisfy orthogonality condition for the existence of a symmetric solution to the Stage II. We consider both cases: undamped and damped models.

3.1. Linear case (undamped model)

Consider the generalized symmetric eigenvalue problem $K\phi = \lambda M\phi$ and the associated real-form representation of the complex matrix eigenpair $(X, \mathcal{T})$.

Theorem 2. *Given $M = M^T$, $\mathcal{T} \in \mathbb{R}^{k\times k}$, a block-diagonal matrix of the form* (6), *$X \in \mathbb{R}^{n\times k}$ of the form* (7) *with full rank. There exists a symmetric non-zero matrix K such that $KX = MX\mathcal{T}$ if and only if $X^T MX = B$, where B is a block-diagonal matrix of the form*:

$$B = \operatorname{diag}(B_1, \ldots, B_l, B_{2l+1}, \ldots, B_k), \quad B_j = \begin{cases} \begin{pmatrix} a_j & b_j \\ b_j & -a_j \end{pmatrix}, & j = 1, \ldots, l \\ b_j, & j = 2l+1, \ldots, k \end{cases} \tag{12}$$

Proof. ($\Leftarrow$) *Sufficiency*: Since X has full rank, the matrix equation $KX = MX\mathcal{T}$ has a nonzero solution. To prove that there exists a symmetric solution K to the equation, we consider an extension of $(X, \mathcal{T})$ in the form $X_{ext} = [X\ \hat{X}] \in \mathbb{R}^{n\times n}$, $\mathcal{T}_{ext} = \operatorname{diag}(\mathcal{T}, \hat{\mathcal{T}}) \in \mathbb{R}^{n\times n}$ such that $X_{ext}^T M X_{ext} = B_{ext} = \operatorname{diag}(B, \hat{B})$, where $\hat{B}$ is a block-diagonal matrix, X_{ext} is of full rank, and $\hat{\mathcal{T}}$ is a block-diagonal matrix. Now define

$$K = X_{ext}^{-T} B_{ext} \mathcal{T}_{ext} X_{ext}^{-1}$$

Then, obviously, $KX = MX\mathcal{T}$, moreover since $B_{ext}\mathcal{T}_{ext}$ is a symmetric matrix, K is also symmetric. Different choices of $\hat{X}$ and $\hat{\mathcal{T}}$ will produce different symmetric solutions to the above equation.

($\Rightarrow$) *Necessity*: Since M is symmetric, there exists a matrix Φ, such that $\Phi^T M\Phi = D$, where D is a diagonal matrix, see [19].

Setting $X = \Phi S^{-1}$, where S is defined as in Eq. (8), we have, $X^T MX = S^{-T}\Phi^T M\Phi S^{-1} = S^{-T} D S^{-1} = B$. Thus, a 2×2 block of B is of the following form:

$$\sqrt{2}\begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix}^{-T} \begin{pmatrix} a+ib & 0 \\ 0 & a-ib \end{pmatrix} \sqrt{2}\begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix}^{-1} = \begin{pmatrix} a & b \\ b & -a \end{pmatrix} \qquad \square$$

3.2. Quadratic case (damped model)

In this section, we prove an analogous result for the quadratic pencil $P(\lambda)$. In this case there are $2n$ eigenvalues and eigenvectors.

The real-form representation of the matrix eigenpair (Φ, Λ) of $P(\lambda)$ is denoted by $(X, \mathcal{T})$.

The pair $(X, \mathcal{T})$ satisfies relation (9), which can be written as

$$\begin{pmatrix} -K & 0 \\ 0 & M \end{pmatrix} \begin{pmatrix} X \\ X\mathcal{T} \end{pmatrix} = \begin{pmatrix} C & M \\ M & 0 \end{pmatrix} \begin{pmatrix} X \\ X\mathcal{T} \end{pmatrix} \mathcal{T} \tag{13}$$

This shows that the matrix

$$\begin{pmatrix} X \\ X\mathcal{T} \end{pmatrix}$$

is the real-form representation of the eigenvector matrix of the $2n \times 2n$ linear pencil

$$\left(\begin{pmatrix} -K & 0 \\ 0 & M \end{pmatrix}, \begin{pmatrix} C & M \\ M & 0 \end{pmatrix} \right)$$

Since M, C, K are symmetric, by Theorem 2, therefore, we have

$$\begin{pmatrix} X \\ X\mathscr{T} \end{pmatrix}^T \begin{pmatrix} C & M \\ M & 0 \end{pmatrix} \begin{pmatrix} X \\ X\mathscr{T} \end{pmatrix} = B(X, \mathscr{T}) \tag{14}$$

where $B = \mathrm{diag}(B_1, \ldots, B_l, B_{2l+1}, \ldots, B_k)$ is a block-diagonal matrix with blocks defined as in Eq. (12). Note relation (14) is equivalent to the orthogonality relation (4) for the complex eigenpair.

To solve the inverse problem, i.e. to find a symmetric K which will satisfy the eigenvalue-eigenvector relation (9) for given $M = M^T, C = C^T$, and the rank condition $(X, \mathscr{T})$ satisfying Eq. (11), we find an extension of the matrices X and $\mathscr{T}$, $X_{ext} = [X\, \hat{X}] \in \mathbb{R}^{n\times n}$, $\mathscr{T}_{ext} = \mathrm{diag}(\mathscr{T}, \hat{\mathscr{T}}) \in \mathbb{R}^{n\times n}$, such that

$$\begin{pmatrix} X_{ext} \\ X_{ext}\mathscr{T}_{ext} \end{pmatrix} \quad \text{is of full rank and}$$

$$\begin{pmatrix} X_{ext} \\ X_{ext}\mathscr{T}_{ext} \end{pmatrix}^T \begin{pmatrix} C & M \\ M & 0 \end{pmatrix} \begin{pmatrix} X_{ext} \\ X_{ext}\mathscr{T}_{ext} \end{pmatrix} = [B(X, \mathscr{T}), \hat{B}] = B_{ext}(X_{ext}, \mathscr{T}_{ext})$$

Here $\hat{\mathscr{T}}, \hat{B}$ are real block-diagonal matrices. Then we can take K as the solution to the following linear system

$$\begin{pmatrix} X_{ext} \\ X_{ext}\mathscr{T}_{ext} \end{pmatrix}^T \begin{pmatrix} -K & 0 \\ 0 & M \end{pmatrix} \begin{pmatrix} X_{ext} \\ X_{ext}\mathscr{T}_{ext} \end{pmatrix} = B_{ext}\mathscr{T}_{ext}$$

i.e. $K = X_{ext}^{-T}(\mathscr{T}_{ext}^T X_{ext}^T M X_{ext}\mathscr{T}_{ext} - B_{ext}\mathscr{T}_{ext})X_{ext}^{-1}$. This is a real symmetric matrix (note, that $B_{ext}\mathscr{T}_{ext}$ is a symmetric matrix).

The above discussion leads to the following theorem.

Theorem 3. *Given $M = M^T \in \mathbb{R}^{n\times n}, C = C^T \in \mathbb{R}^{n\times n}$; and $\mathscr{T} \in \mathbb{R}^{k\times k}, X \in \mathbb{R}^{n\times k}$, $k<n$ matrices of the form* (6), (7), *respectively. Let $(X, \mathscr{T})$ satisfy condition* (11). *Then there is a real symmetric matrix K such that $MX\mathscr{T}^2 + CX\mathscr{T} + KX = 0$ if and only if*

$$\begin{pmatrix} X \\ X\mathscr{T} \end{pmatrix}^T \begin{pmatrix} C & M \\ M & 0 \end{pmatrix} \begin{pmatrix} X \\ X\mathscr{T} \end{pmatrix} = B(X, \mathscr{T}) \tag{15}$$

where B is some block-diagonal matrix with blocks of the form (12).

Corollary 4. *Assume that $\mathscr{T}$ is a real-form representation of an eigenvalue matrix of the pencil (M, C, K) and all the diagonal blocks of $\mathscr{T}$ are distinct, then B is of form* (12) *if and only if $B\mathscr{T} = \mathscr{T}^T B$.*

Proof. Consider the matrix D_3 from Eq. (4), and note that $B = S^{-T}D_3S^{-1}$. Then, $B\mathscr{T} = S^{-T}D_3S^{-1}S\Lambda S^{-1} = S^{-T}D_3\Lambda S^{-1}$ and $\mathscr{T}^T B = S^{-T}\Lambda S^T S^{-T}D_3S^{-1} = S^{-T}\Lambda D_3 S^{-1}$. Now, $B = S^{-T}D_3S^{-1}$ is of the form (12) if and only if $D = S^TBS$ is a diagonal matrix, which is equivalent, to

$$\Lambda D_3 = D_3\Lambda$$

This implies that $B\mathscr{T} = \mathscr{T}^T B$. □

4. A two-stage model updating scheme

In this section, we introduce our two-stage model updating scheme for FEMU. Throughout, we assume that M_a, C_a, K_a, $X_M = \Phi_M S^{-1}$, and $\mathscr{T}_M = S\Lambda_M S^{-1}$ are given.

Stage I: In this stage, the real-form representation of the measured eigenvector matrix X_M is updated so that it becomes as close as possible to the analytical data in the sense that a weighted distance between them is minimized. Furthermore, an orthogonality constraint stated in Corollary 4 is enforced. Mathematically, the problem may be stated as follows:

$$\begin{array}{lll} (\mathscr{P}) & \min & \frac{1}{2}\|W_1^{-1/2}(X - X_M)W_1^{-1/2}\|_F^2 \\ & \text{s.t.} & H(X) = 0 \\ & & X \in \mathbb{R}^{n\times k} \end{array}$$

where

$$H(X) = B(X, \mathscr{T}_M)\mathscr{T}_M - \mathscr{T}_M^T B(X, \mathscr{T}_M) \tag{16}$$

with $B(X, \mathscr{T}_M)$ given by Eq. (14) and W_1 is some positive-definite weighting matrix. A solution to the problem will be denoted by X_u.

Stage II: Let X_u be a solution from Stage I. In this stage, we would like to update the stiffness matrix K so that it becomes as close as possible to K_a in the sense that a weighted distance between K and K_a is minimized. In addition, constraints on symmetry for K and eigenvalue-eigenvector relation of Eq. (9) are enforced. Mathematically, this amounts to solving the following minimization problem:

$$
\begin{aligned}
(\mathcal{Q}) \quad \min \quad & \tfrac{1}{2}\|W_2^{-1/2}(K-K_a)W_2^{-1/2}\|_F^2 \\
\text{s.t.} \quad & M_a X_u \mathcal{T}_M^2 + C_a X_u \mathcal{T}_M + K X_u = 0 \\
& K^T = K \ (\text{symmetry}) \\
& K \in \mathbb{R}^{n\times n}
\end{aligned}
$$

where $W_2 > 0$ is a positive-definite weighting matrix. The solution to the problem will be denoted by K_u.

Problem $(\mathcal{Q})$ is a convex quadratic programming problem with a unique solution. There exists an analytical expression [6] and a computational algorithm [13] based on numerical linear algebra techniques. Since $(\mathcal{Q})$ is a simple convex quadratic programming problem, we can also solve it numerically by existing optimization techniques.

Remarks. It is also possible to update both the stiffness and damping matrices satisfying the orthogonality relation of Stage I. This will require reformulation of the problem. Such reformulation is currently being investigated.

5. A solution method and its convergence properties

We now focus on how to solve $(\mathcal{P})$. As noted before Problem $(\mathcal{Q})$ is a convex quadratic programming problem for which there exist excellent numerical methods. However, in our numerical experiment, we use the same method developed for $(\mathcal{P})$ to solve $(\mathcal{Q})$.

To simplify the presentation, set $W = I$, and

$$f(X) = \tfrac{1}{2}\|X - X_M\|_F^2$$

Then the Lagrangian function for $(\mathcal{P})$ is

$$L(X,Y) = f(X) + \langle Y, H(X)\rangle$$

where $Y \in \mathbb{R}^{k\times k}$. Some remarks about our Lagrange function L are in order. By definition, $H(X)^T = -H(X)$. Hence the system $H(X) = 0$ defines $k(k-1)/2$ constraints.

The necessary optimality conditions for $(\mathcal{P})$ can now be stated as follows: find a pair (X_*, Y_*) such that

$$\nabla_X L(X_*, Y_*) = 0, \quad H(X_*) = 0 \tag{17}$$

Problem $(\mathcal{P})$ has a convex quadratic objective function, and polynomial equality constraints. So it is a polynomial programming problem. But the feasible region defined by the polynomial equality constraints are nonconvex in general. Hence we are dealing with a nonconvex minimization problem with equality constraints.

Generally speaking, if X_* is an optimal solution for $(\mathcal{P})$, and the constraint system satisfies certain regularity condition at X_*, then there is a $Y_* \in \mathbb{R}^{k\times k}$ such that (X_*, Y_*) satisfies Eq. (17). Elements of Y_* are usually called Lagrangian multipliers.

Optimization techniques for constrained problems such as $(\mathcal{P})$ have been well developed in the past fifty years. There are many efficient methods, including augmented Lagrangian methods, to solve a nonlinear programming problem with equality constraints. The first augmented Lagrangian method was independently proposed by Hestenes [21] and Powell [22] by adding a quadratic penalty term to its Lagrangian function $L(X,Y)$. Because of its attractive features, such as ease to implement, it has emerged as an important method for handling constrained optimization problems. The literature on augmented Lagrangian methods is vast. We refer the reader to [23,24] for a thorough treatment on this class of methods and its convergence theory. Following Hestenes and Powell, we propose an augmented Lagrangian method to solve $(\mathcal{P})$. To this end, we introduce the following parameterized family of the augmented Lagrangian functions:

$$L_\rho(X,Y) = L(X,Y) + \frac{\rho}{2}\|H(X)\|_F^2 \tag{18}$$

where ρ is a positive constant.

We now discuss an issue associated with our proposed augmented Lagrangian method.

The existence of a global minimizer in Step 3: for $\mathcal{P}(i)$, the following proposition guarantees that there is a solution.

Proposition 5. *Let $\rho > 0$ and $Y \in \mathbb{R}^{k\times k}$. Then $\arg\min_X L_\rho(X,Y)$ is non-empty.*

Proof. We will prove that $L_\rho(\cdot, Y)$ is level-bounded [24, Definition 1.8]; that is, for each real μ, the set $\{X | L_\rho(X,Y) \leqslant \mu\}$ is bounded. Once this is done, the non-emptiness of $\arg\min_X L_\rho(X,Y)$ follows from [24, Theorem 1.9].

Let $\{X_i\}$ be a sequence such that $\|X_i\|_F \to \infty$ as $i \to \infty$. Then

$$L_\rho(X_i, Y) \to \infty \quad \text{as } i \to \infty$$

since, by Cauchy-Schwartz inequality $\langle Y, H(X_i)\rangle \geqslant -\|Y\|_F \|H(X_i)\|_F$,

$$\langle Y, H(X_i)\rangle + \rho/2\|H(X_i)\|_F^2 \geqslant (\rho/2\|H(X_i)\|_F - \|Y\|_F)\|H(X_i)\|_F$$

and $f(X_i) \to \infty$ as $i \to \infty$. If $\{X : L_\rho(X, Y) \leqslant \mu\}$ were not a bounded set for some μ, then there would exist a sequence $\{X_i\}$ sequence such that $\|X_i\|_F \to \infty$ as $i \to \infty$. This would imply that

$$\mu \geqslant f(X_i) + \langle Y, H(X_i)\rangle + \rho/2\|H(X_i)\|_F^2 \to +\infty$$

as $i \to \infty$ by the above argument. The contradiction proves the level-boundedness of $L_\rho(\cdot, Y)$. □

Algorithm 1. The augmented Lagrangian method for $(\mathscr{P})$.

INPUT: $X_0, Y_0, \rho_0 > 0, 0 < \beta < 1$, and $\varepsilon > 0$
OUTPUT: Solution to $(\mathscr{P})$
1: **for** $i = 0, 1, \ldots$ **do**
2: Stop if $\|\nabla_X L(X_i, Y_i)\| \leqslant \varepsilon$, and $\|H(X_i)\| \leqslant \varepsilon$
3: Solve the unconstrained optimization subproblem:

$$(\mathscr{P}(i)) \quad \min_X L_{\rho_i}(X, Y_i)$$

with the stoping criteria, $\|\nabla_X L_{\rho_i}(X, Y_i)\|_F < \beta^i$
Let X_{i+1} be the solution of $(\mathscr{P}(i))$
4: Update the multiplier matrix:

$$Y_{i+1} = Y_i + \rho_i H(X_{i+1}).$$

Then choose $\rho_{i+1} > \rho_i$
5: **end for**

5.1. The convergence of the proposed method

As we have already pointed out before, $(\mathscr{P})$ is a nonconvex programming problem. It is well known in optimization literature that finding a global minimizer for a nonconvex programming problem is a very challenging task. A practical way for solving a sequence of nonconvex programming problems in Step 3 is to find a sequence of critical points instead. The following theorem, which is embedded somewhere in the general convergence theory in [23,24], ensures that such a sequence of critical points will have a convergent subsequence.

To make our presentation self-contained and for the reader's convenience, we include a proof of this theorem in the next section after computable gradient formulas are derived. Let us set

$$F(X) = [H_{12}(X), \ldots, H_{1k}(X), H_{23}(X), \ldots, H_{2k}(X), \ldots, H_{(k-1)k}(X)]^T$$

and

$$\hat{Y} = [Y_{12}, \ldots, Y_{1k}, Y_{23}, \ldots, Y_{2k}, \ldots, Y_{(k-1)k}]^T \tag{19}$$

Then $F : \mathbb{R}^{k\times k} \to \mathbb{R}^{(k-1)k/2}$, and it is easy to see that $F(X) = 0$ if and only if $H(X) = 0$. Also, simple calculations show that

$$\langle Y, H(X)\rangle = 2\langle \hat{Y}, F(X)\rangle$$

and

$$\|H(X)\|_F^2 = 2\|F(X)\|_F^2$$

Our Lagrangian functions then become

$$L(X, Y) = L(X, \hat{Y}) = f(X) + 2\langle \hat{Y}, F(X)\rangle$$

and

$$L_\rho(X, Y) = L_\rho(X, \hat{Y}) = L(X, \hat{Y}) + \rho\|F(X)\|_F^2$$

Theorem 6. *Suppose $\|\nabla_X L_{\rho_j}(X_j, \hat{Y}_j)\|_F < \beta^j$, $0 < \beta < 1$ for $j = 0, 1, 2, \ldots$, with $\{\|\hat{Y}_j\|_F\}$ bounded, and $\rho_j < \rho_{j+1}$ and $\rho_j \to \infty$ as $j \to \infty$. If there is a convergent subsequence $\{X_{j_i}\}$ of $\{X_j\}$ with $X_{j_i} \to X_*$ such that $\nabla F(X_*)$ maps $\mathbb{R}^{n\times k}$ onto $\mathbb{R}^{[(k-1)\times k/2]\times[n\times k]}$, then there is a $\hat{Y}_*$ such that*

$$\nabla_X L(X_*, \hat{Y}_*) = 0, \quad F(X_*) = 0 \tag{20}$$

Remarks. (a) In view of the relationships between F and H, $\hat{Y}$ and Y, it is easy to see that $\nabla_X L(X_*, \hat{Y}_*) = 0$ and $F(X_*) = 0$ if and only if there is some Y_* (by Eq. (19)) such that the pair (X_*, Y_*) satisfies Eq. (17). (b) It is also clear from Eq. (19) that $\{\|Y_j\|_F\}$ bounded if and only if $\{\|\hat{Y}_j\|_F\}$ is bounded. The following simple example shows that if the conditions listed in Theorem 6 are not satisfied, Lagrangian multipliers may not exist.

Example 7. Consider $\min_{x\in R} x$ subject to $x^2 = 0$. Then it is trivial to see that $x = 0$ is an optimal solution. However there is no Lagrangian multiplier for the minimization problem.

6. Computable gradient formulas

To implement the proposed method effectively, we need to have gradient formulas that can be computed in terms of the given quantities. In this section, we will show that gradients of functions associated with problems ($\mathscr{P}$) and ($\mathscr{Q}$) can be computed in terms of their associated matrices. This is particularly important since our numerical experiments are conducted in MATLAB environment.

The basic idea for deriving gradient formulas comes from operator theory on adjoint operators as has been used in [25]. It is easy and elementary to see that

$$\nabla f(X) = X - X_M$$

Let

$$h(X) = \tfrac{1}{2}\|H(X)\|_F^2 = \tfrac{1}{2}\langle H(X), H(X)\rangle$$
$$g_Y(X) = \langle Y, H(X)\rangle$$

Then

$$\nabla h(X) = 2(M_a X G(X)\mathscr{T}_M^T + (C_a X + M_a X\mathscr{T}_M)G(X)) \tag{21}$$

where $G(X) = H(X)\mathscr{T}_M^T - \mathscr{T}_M H(X)$ and $H(X)$ is given by (16)

$$\nabla g_Y(X) = 2(M_a X(Y\mathscr{T}_M^T - \mathscr{T}_M Y)\mathscr{T}_M^T + (C_a X + M_a X\mathscr{T}_M)(Y\mathscr{T}_M^T - \mathscr{T}_M Y)) \tag{22}$$

The details of the derivation of formula (21) are given in Appendix A.

With the above gradient formulas, the gradient formulas for L and L_ρ with respect to X can be written as

$$\nabla_X L(X,Y) = \nabla f(X) + \nabla g_Y(X)$$
$$\nabla_X L_\rho(X,Y) = \nabla_X L(X,Y) + \rho\nabla h(X)$$

6.1. Gradient formulas for problem $\mathscr{Q}$

Gradient functions in Problem ($\mathscr{Q}$) are much simpler, and can be written down as follows. For $\bar{h}(K) = \frac{1}{2}\|K - K_a\|_F^2$, we have

$$\nabla\bar{h}(K) = K - K_a$$

For $\bar{f}(K) = \frac{1}{2}\|M_a X_u\mathscr{T}_M^2 + C_a X_u\mathscr{T}_M + KX_u\|_F^2$, we have

$$\nabla\bar{f}(K) = (M_a X_u\mathscr{T}_M^2 + C_a X_u\mathscr{T}_M + KX_u)X_u^T$$

We will now use the above gradient formulas to obtain the necessary optimality conditions for Stage I.

Necessary optimality conditions in matrix form for ($\mathscr{P}$): Find X and Y such that

$$\nabla f(X) + \nabla g_Y(X) = 0$$
$$H(X) = 0$$

where $H(X)$ is given by Eq. (16), and $\nabla g_Y(X)$ is given by Eq. (22). The above necessary optimality conditions expressed in terms of given matrices are significant. It not only opens up the possibility of solving ($\mathscr{P}$) by solving the above systems of equations (such as by Newton's method for nonlinear equations), but also forms the basis for sensitivity analysis when the problem data undergoes small changes.

We conclude this section by including a proof of Theorem 6. We begin with a well-known lemma on the invertibility of a matrix (see e.g. [19, p. 319]).

Lemma 8. *Let p, q be positive integers such that $p \geqslant q$. Suppose that A is an $p \times q$ matrix with rank q. Then $[A^T A]^{-1}$ exists.*

Proof. The proof is by contradiction. Let u be a nonzero $q \times 1$ vector such that $A^T A u = 0$. Then $(Au)^T(Au) = 0$. It follows that $Au = 0$. The rank condition on A implies that $u = 0$, which is a contradiction. This proves $A^T A$ is invertible. □

Proof of Theorem 6. Without any loss of generality, suppose that $X_j \to X_*$ as $j \to \infty$. For any fixed X, since $\nabla F(X)$ is a $[(k-1)\times k/2]\times[n\times k]$ matrix, $\nabla F(X_*)$ has rank $(k-1)\times k/2$, and $\nabla F(\cdot)$ is continuous. We may further assume that $\nabla F(X_j)$ has rank $(k-1)\times k/2$ for all j. Set $A_j = \nabla F(X_j)^T$ and $A_* = \nabla F(X_*)^T$. By Lemma 8, $A_j^T A_j$ is invertible. The continuity of $\nabla F(\cdot)$ implies that $[A_j^T A_j]^{-1} \to [A_*^T A_*]^{-1}$. By

$$\nabla_X L_{\rho_j}(X_j, \hat{Y}_j) = \nabla f(X_j) + 2\nabla F(X_j)^T[\hat{Y}_j + \rho_j F(X_j)] \tag{23}$$

and multiplying both sides of Eq. (23) by $\nabla F(X_j) = A_j^T$, we have

$$2(\hat{Y}_j + \rho_j F(X_j)) = [A_j^T A_j]^{-1} A_j^T[\nabla_X L_{\rho_j}(X_j, \hat{Y}_j) - \nabla f(X_j)]$$

Since $\|\nabla_X L_{\rho_j}(X_j, \hat{Y}_j)\|_F \to 0$ and $[A_j^T A_j]^{-1} \to [A_*^T A_*]^{-1}$ as $j \to \infty$, we get

$$2(\hat{Y}_j + \rho_j F(X_j)) \to -[A_*^T A_*]^{-1} A_*^T \nabla f(X_*) \quad \text{as } j \to \infty$$

Set $2\hat{Y}_* = -[A_*^T A_*]^{-1} A_*^T \nabla f(X_*)$. By taking the limit on both sides of Eq. (23), we have that $\nabla_X L(X_*, \hat{Y}_*) = 0$. Since $2(\hat{Y}_j + \rho_j F(X_j)) \to 2\hat{Y}_*$, and $\{\|\hat{Y}_j\|_F\}$ is bounded, the sequence $\{\rho_j F(X_j)\}$ is bounded. By $\rho_j \to \infty$ as $j \to \infty$, we conclude that

$$\lim_{j\to\infty} F(X_j) = F(X_*) = 0$$

So the pair $(X_*, \hat{Y}_*)$ satisfies Eq. (20). This completes the proof. □

7. Case studies

In this section, we present the results of our numerical experiments on

- A spring–mass system of 10 degree of freedom (DoF) [1].
- A vibrating beam.

The data for our experiments are set up as follows:

- The matrices M_a, C_a are kept fixed.
- To simulate the measured data $(X_M, \mathcal{T}_M)$, we add a random noise with some Gaussian distribution to the eigendata of the analytical model.
- The weighting matrix was taken as $W = I$.

We used MATLAB with double arithmetics to run numerical experiments. As a solver of an unconstrained optimization problem MATLAB optimization toolbox routine fminunc which implements BFGS quasi-Newton method has been used.

7.1. A mass–spring system of 10 DoF

Consider the example of a mass–spring system of 10 DoF, as depicted in Fig. 1. In this example all rigid bodies have a mass of 1 kg, and all springs have stiffness 1 kN/m. The analytical model is given by

$$M_a = I$$

$$C_a = \begin{pmatrix} 0.48100 & -8.3809 & & & & & & & & \\ -8.3809 & 8.3809 & -1.0254 & & & & & & & \\ & -1.0254 & 1.0254 & -7.2827 & & & & & & \\ & & -7.2827 & 7.2827 & -4.4050 & & & & & \\ & & & -4.4050 & 4.4050 & -9.9719 & & & & \\ & & & & -9.9719 & 9.9719 & -5.6247 & & & \\ & & & & & -5.6247 & 5.6247 & -4.6585 & & \\ & & & & & & -4.6585 & 4.6585 & -4.1901 & \\ & & & & & & & -4.1901 & 4.1901 & -2.1160 \\ & & & & & & & & -2.1160 & 2.1160 \end{pmatrix}$$

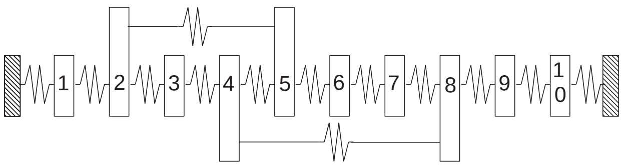

Fig. 1. Mass–spring system.

$$K_a = \begin{pmatrix} 2000 & -1000 & & & & & & & & \\ -1000 & 3000 & -1000 & & -1000 & & & & & \\ & -1000 & 2000 & -1000 & & & & & & \\ & & -1000 & 3000 & -1000 & & & -1000 & & \\ & -1000 & & -1000 & 3000 & -1000 & & & & \\ & & & & -1000 & 2000 & -1000 & & & \\ & & & & & -1000 & 2000 & -1000 & & \\ & & & -1000 & & & -1000 & 3000 & -1000 & \\ & & & & & & & -1000 & 2000 & -1000 \\ & & & & & & & & -1000 & 2000 \end{pmatrix}$$

The measured data for our experiment was simulated by reducing stiffness of the spring between masses 2 and 5 to 600 N/m and adding Gaussian noise with $\sigma = 2\%$.

The analytical eigenvalue and eigenvector matrices are:

$$\mathcal{T}_a = \begin{pmatrix} -6.23 & 71.1 & 0 & 0 \\ -71.1 & -6.23 & 0 & 0 \\ 0 & 0 & -3.67 & 65.9 \\ 0 & 0 & -65.9 & -3.67 \end{pmatrix}, \quad X_a = \begin{pmatrix} 0.142 & 0.001 & -0.161 & -0.001 \\ -0.438 & 0.020 & 0.372 & -0.050 \\ 0.288 & 0.065 & -0.056 & 0.031 \\ -0.502 & -0.206 & -0.191 & 0.087 \\ 0.479 & 0.148 & -0.296 & -0.034 \\ -0.136 & -0.011 & 0.263 & 0.091 \\ -0.066 & 0.011 & -0.346 & -0.145 \\ 0.339 & 0.003 & 0.599 & 0.063 \\ -0.122 & -0.007 & -0.296 & -0.093 \\ 0.040 & 0.010 & 0.115 & 0.075 \end{pmatrix}$$

The measured eigenvalue and eigenvector matrices are:

$$\mathcal{T}_M = \begin{pmatrix} -6.16 & 69.8 & 0 & 0 \\ -69.8 & -6.16 & 0 & 0 \\ 0 & 0 & -4.7 & 64.9 \\ 0 & 0 & -64.9 & -4.7 \end{pmatrix}, \quad X_M = \begin{pmatrix} 0.102 & 0.026 & -0.172 & -0.023 \\ -0.283 & -0.061 & 0.401 & 0.023 \\ 0.282 & 0.115 & -0.195 & -0.005 \\ -0.579 & -0.240 & 0.074 & 0.068 \\ 0.341 & 0.242 & -0.354 & -0.202 \\ -0.067 & -0.054 & 0.286 & 0.242 \\ -0.168 & 0.036 & -0.249 & -0.340 \\ 0.508 & -0.042 & 0.362 & 0.382 \\ -0.207 & 0.009 & -0.183 & -0.246 \\ 0.077 & 0.011 & 0.060 & 0.130 \end{pmatrix}$$

Results of Stage I: Orthogonality condition was not satisfied with the measured eigenvector matrix as shown by the residual:

$$\|H(X_M)\|_F = 17.7$$

However, with preprocessing of the measured data using Stage I we obtained

$$X_u = \begin{pmatrix} 0.086 & 0.128 & 0.139 & 0.021 \\ -0.270 & -0.226 & -0.324 & 0.055 \\ 0.266 & 0.176 & 0.233 & -0.131 \\ -0.437 & -0.239 & -0.222 & 0.264 \\ 0.227 & 0.387 & 0.343 & 0.041 \\ 0.031 & -0.261 & -0.218 & -0.232 \\ -0.147 & 0.205 & 0.183 & 0.312 \\ 0.395 & -0.239 & -0.226 & -0.544 \\ -0.145 & 0.142 & 0.149 & 0.244 \\ 0.051 & -0.061 & -0.064 & -0.096 \end{pmatrix}$$

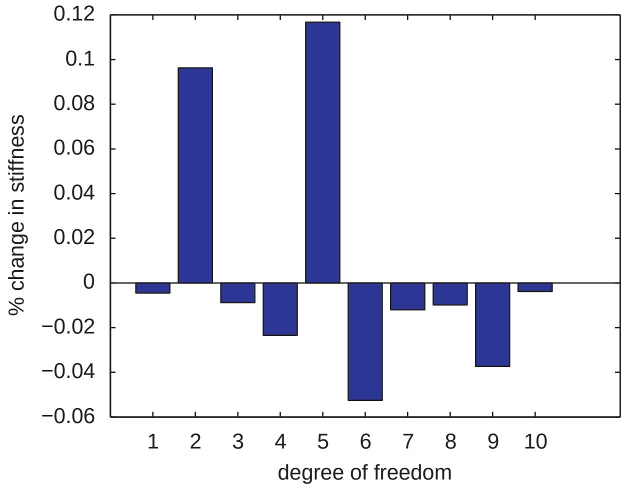

Fig. 2. The percentage change in the diagonal elements of the stiffness matrix for mass–spring system.

The orthogonality condition is now satisfied with this updated eigenvector matrix as shown by the following residual:

$$\|H(X_u)\|_F = 1.97 \times 10^{-8}$$

Results of Stage II: With the updated eigenvector matrix X_u from Stage I:

- The updated matrix K_u was symmetric, as shown by the residual norm:

$$\|K_u - K_u^T\|_F = 9.293 \times 10^{-9}$$

- The measured eigenvalues and corrected measured eigenvectors were reproduced accurately by the updated model, as shown by the following residual:

$$\|R(K_u)\|_F = 1.78143 \times 10^{-6}$$

where $R(K) = M_a X_u \mathscr{T}_M^2 + C_a X_u \mathscr{T}_M + K X_u$.

Note: It is clear from Fig. 2 that the largest changes correspond to degrees of freedom 2 and 5. The changes corresponding to the other degrees of freedom are reasonably small.

7.2. Vibrating beam

Consider a discrete spring–mass model of a vibrating beam [26], which consists of $n+2$ masses $\{m_i\}_{i=-1}^n$, linked by massless rigid rods of length $\{l_i\}_{i=0}^n$ which are themselves connected by n rotational springs of stiffness $\{k_i\}_{i=1}^n$. This model corresponds to a finite difference approximation of a beam with distributed parameters. The vibration of the beam with clamped left hand end and with no force applied at the free end is governed by

$$M\ddot{x} + Kx = 0$$

where

$$K = EL^{-1}E\hat{K}E^TL^{-1}E^T$$

$\hat{K} = \text{diag}(k_1, \ldots, k_n)$, $L = \text{diag}(l_1, \ldots, l_n)$, $M = \text{diag}(m_1, \ldots, m_n)$,

$$E = \begin{pmatrix} 1 & -1 & 0 & \ldots & 0 \\ 0 & 1 & -1 & \ldots & 0 \\ & \ddots & \ddots & \ddots & \\ 0 & \ldots & 0 & 1 & -1 \\ 0 & \ldots & 0 & 0 & 1 \end{pmatrix}$$

The simulated beam has 16 rods of length $\frac{1}{16}$ m each, and all masses are 0.1 kg.

Results of Stage I: The measured data was obtained from the analytical data in the same way as the previous example. To simulate the measured data, the coefficients k_3, k_5, k_9 were reduced by 40%, 50%, 30%, respectively, and gaussian noises with $\sigma = 2\%$ were added. The simulated measured eigenvector matrix became

$$X_M = \begin{pmatrix} -1.2528\times10^{-5} & -0.00029345 & -0.0028605 \\ 7.9745\times10^{-5} & 0.0015719 & 0.012768 \\ -0.00039017 & -0.0064281 & -0.043326 \\ 0.0015785 & 0.021449 & 0.11279 \\ -0.0054324 & -0.058119 & -0.23438 \\ 0.016097 & 0.13210 & 0.38359 \\ -0.041476 & -0.24960 & -0.47393 \\ 0.091780 & 0.39290 & 0.39510 \\ -0.17778 & -0.49152 & -0.10852 \\ 0.29897 & 0.45559 & -0.24409 \\ -0.44078 & -0.23769 & 0.37962 \\ 0.53922 & -0.10088 & -0.16311 \\ -0.52197 & 0.36497 & -0.22086 \\ 0.34568 & -0.34533 & 0.33064 \\ -0.10101 & 0.11494 & -0.12827 \end{pmatrix}$$

Without application of Stage I, the matrix X_M did not satisfy the orthogonality constraint, as shown by the residual:

$$\|H(X_M)\|_F = 3.668\times10^{7}$$

Application of Stage I yielded:

$$X_u = \begin{pmatrix} -1.6363\times10^{-5} & -0.00029433 & -0.0028606 \\ 9.6687\times10^{-5} & 0.0015758 & 0.012768 \\ -0.00044692 & -0.0064414 & -0.043328 \\ 0.0017231 & 0.021483 & 0.11280 \\ -0.0057238 & -0.058189 & -0.23441 \\ 0.016549 & 0.13220 & 0.38365 \\ -0.041977 & -0.24972 & -0.47406 \\ 0.092072 & 0.39297 & 0.39535 \\ -0.17760 & -0.49144 & -0.10893 \\ 0.29832 & 0.45532 & -0.24353 \\ -0.44009 & -0.23728 & 0.37893 \\ 0.53906 & -0.10129 & -0.16238 \\ -0.52253 & 0.36525 & -0.22148 \\ 0.34638 & -0.34546 & 0.33102 \\ -0.10127 & 0.11497 & -0.12838 \end{pmatrix}$$

The updated matrix X_u did satisfy the orthogonality constraint, as shown by the residual:

$$\|H(X_u)\|_F = 9.09143\times10^{-6}$$

Results of Stage II: With the updated eigenvector matrix X_u from Stage I:

- The updated matrix K_u was symmetric, as shown by the residual norm:

  $$\|K_u - K_u^T\|_F = 9.897\times10^{-8}$$

- The measured eigenvalues and the corrected measured eigenvectors were reproduced accurately by the updated model, as shown by the following residual:

  $$\|R(K_u)\|_F = 5.795\times10^{-5}$$

8. Summary and conclusions

An important criterion for acceptability of a set of measured data (especially the eigenvectors) for a model updating procedure is that it must satisfy an appropriate orthogonality condition. For symmetric positive definite linear model which is often discussed in the literature, the orthogonality condition is the mass-normalization property of the eigenvector matrix. Until recently such orthogonality relations for a damped model were not known. Recently three orthogonality relations for a symmetric quadratic pencil have been derived. One such relation, involving the mass and damping matrices, has been used in this paper to update a damped model. This consideration gives rise to a two-stage optimization procedure. In Stage I, a set of measured eigenvector matrices are updated to satisfy the orthogonality constraint and in Stage II, the stiffness matrix is updated with the constraints on the symmetry of the stiffness matrix and reproduction of the measured eigenvalues and updated measured eigenvectors by the updated model. The orthogonality constraint in Stage I being nonlinear, finding an explicit solution, and finding a global solution to the posed optimization problem, unlike the linear case, remain difficult tasks. However, the gradient formulas, needed for numerical solutions of optimization problems in both stages, have been derived in terms of only the given quantities; namely, the coefficient matrices and a partial knowledge of the analytical and measured eigenvalues and eigenvectors. It has also been shown mathematically by solving a quadratic inverse eigenvalue problem in the paper that satisfaction of the orthogonality constraint in Stage I is necessary and sufficient for the updating in Stage II to be successful. The possibility of combining the two stages to simultaneously update stiffness and damping matrices so that the constraints on orthogonality, symmetry, and reproduction of the measured eigenvalues and eigenvectors are satisfied, is currently being investigated. A further challenging task will be to devise a model updating procedure that preserves the connectivity and other physical structures of the original model. However, the following underlying inverse eigenvalue problem needs to solved first:

Given a set of k complex numbers and a set of k complex vectors ($k<2n$), both closed under complex conjugation, and two fixed matrices M and C, each of order n, find a symmetric matrix K of order n with a *specified structure* such that the spectrum of the resulting quadratic pencil (M, C, K) will contain the given numbers, and the given vectors will be the corresponding eigenvectors of the pencil.

Acknowledgment

The authors would like to thank the reviewers for their comments and corrections which improved the quality of the paper.

Appendix A. Derivation of the gradient formula (21)

We begin by listing basic properties of the inner product $\langle\cdot,\cdot\rangle$. Let $A \in R^{p\times q}$, $B \in R^{p\times r}$, and $C \in R^{r\times q}$. By using properties of the trace of a square matrix, we can easily verify that

$$\langle A, BC\rangle = \langle AC^T, B\rangle = \langle B^T A, C\rangle \tag{24}$$

Recall the definition of differentiability.

Definition 9. Let $(N, \|\cdot\|)$ and $(N', \|\cdot\|')$ be real finite-dimensional Euclidean spaces. Let $\Psi : N \to N'$ be a vector function. For $x \in N$, Ψ is said to be differentiable at x if there is some linear mapping T from N to N' such that

$$\lim_{\|\Delta x\|\to 0} \frac{1}{\|\Delta x\|}\|\Psi(x+\Delta x) - \Psi(x) - T(\Delta x)\|' = 0$$

We denote T by $\nabla\Psi(x)$. When $N' = R$, $\nabla\Psi(x)$ is called the gradient of Ψ at x.

Since

$$h(X) = \phi(B(X))$$

where $\phi(D) = \frac{1}{2}\|D\mathscr{T}_M - \mathscr{T}_M^T D\|_F^2$, and $B(X) = B(X, \mathscr{T}_M)$ is given by Eq. (14), the chain rule for vector functions tells us that $\nabla h(X)$ acting on ΔX is

$$\nabla h(X)(\Delta X) = \langle \nabla\phi(B(X)), \nabla B(X)(\Delta X)\rangle \tag{25}$$

By Definition 9 and Eq. (24), we can verify directly that

$$\nabla\phi(D) = (D\mathscr{T}_M - \mathscr{T}_M^T D)\mathscr{T}_M^T + \mathscr{T}_M(\mathscr{T}_M^T D - D\mathscr{T}_M)$$

In addition, by expanding $B(X+\Delta X) - B(X)$ and neglecting higher order term of ΔX, we have

$$\nabla B(X)(\Delta X) = \Delta X^T(C_a X + M_a X\mathscr{T}_M) + (X^T C_a + \mathscr{T}_M^T X^T M_a)\Delta X + \mathscr{T}_M^T(\Delta X)^T M_a X + X^T M_a \Delta X \mathscr{T}_M$$

Substituting the last expressions for $\nabla\phi(D)$ and $\nabla B(X)(\Delta X)$ into Eq. (25) and using property (24) repeatedly, we obtain the desired gradient formula (21) for $\nabla h(X)$. We leave detailed routine computations to the reader.

References

[1] M.I. Friswell, J.E. Mottershead, Finite Element Model Updating in Structural Dynamics, Kluwer Academic Publishers, Boston, Dordrecht, London, 1995.
[2] J. Carvalho, B.N. Datta, A. Gupta, M. Lagadapati, A direct method for matrix updating with incomplete measured data and without spurious modes, Mechanical Systems and Signal Processing 21 (2007) 2715–2731.

[3] J. Carvalho, B.N. Datta, W. Lin, C. Wang, Symmetry preserving eigenvalue embedding in finite element model updating of vibration structures, Journal of Sound and Vibration 290 (2006) 839–864.
[4] B. Datta, Finite element model updating, eigenstructure assignment and eigenvalue embedding techniques for vibrating systems, Mechanical Systems and Signal Processing 16 (1) (2001) 83–96.
[5] D.J. Ewins, Adjustment or updating of models, Sādhanā 25 (2000) 235–245.
[6] M.I. Friswell, D.J. Inman, D.F. Pilkey, The direct updating of damping and stiffness matrices, AIAA Journal 36 (3) (1998) 491–493.
[7] Y. Halevi, I. Bucher, Model updating via weighted reference basis with connectivity constraints, Journal of Sound and Vibration 265 (3) (2003) 561–581 URL: ⟨http://dx.doi.org/10.1016/S0022-460X(02)01628-0⟩.
[8] R. Kenigsbuch, Y. Halevi, Model updating in structural dynamics: a generalized preference basis approach, Mechanical Systems and Signal Processing 12 (1998) 75–90.
[9] M. Link, Updating analytical models by using local and global parameters and relaxed optimization requirements, Mechanical Systems and Signal Processing 12 (1) (1998) 7–22.
[10] B. Caesar, Updating system matrices using modal test data, in: Proceedings of the 4th International Modal Analysis Conference, London, England, 1987.
[11] F.-S. Wei, Structural dynamic model improvement using vibration test data, AIAA Journal 26 (9) (1990) 175–177.
[12] B.N. Datta, S. Elhay, Y.M. Ram, Orthogonality and partial pole assignment for the symmetric definite quadratic pencil, Linear Algebra and Applications 257 (1997) 29–48.
[13] W.W.L.Y.C. Kuo, S.F. Xu, A new model updating method for quadratic eigenvalue problems, Preprint (available at ⟨http://math.cts.nthu.edu.tw/Mathematics/preprints/prep2005-1-004-050221.pdf⟩).
[14] C. Minas, D.J. Inman, Matching finite element models to modal data, Transactions of ASME Journal of Applied Mechanics 112 (1990) 84–92.
[15] C. Minas, D.J. Inman, Correcting finite element models with measured modal results using eigenstructure models, in: Proceedings of the 6th International Modal Analysis Conference, Orlando, FL, 1998, pp. 583–587.
[16] D.C. Zimmerman, M. Windengren, Correcting finite element model using a symmetric eigenstructure assignment technique, AIAA Journal 28 (1990) 1670–1676.
[17] J. Nocedal, S.J. Wright, Numerical Optimization, Springer, Berlin, 1999.
[18] B.N. Datta, D.R. Sarkissian, Theory and computations of some inverse eigenvalue problems for the quadratic pencil, in: V. Olshevsky (Ed.), Structured Matrices in Operator Theory, Control, and Signal and Image Processing, vol. 280, Contemporary Mathematics, American Mathematical Society, Providence, RI, 2001, pp. 221–240.
[19] B.N. Datta, Numerical Linear Algebra and Applications, Brooks/Cole Publishing Company, Pacific Grove, CA, 1995.
[20] B.N. Datta, Numerical Methods for Linear Control Systems Design and Analysis, Elsevier Academic Press, Boston, 2003.
[21] M.R. Hestenes, Multiplier and gradient methods, Journal of Optimization Theory and Applications (1969) 303–320.
[22] M.J.D. Powell, A method for nonlinear constraints in minimization problems, in: R. Fletcher (Ed.), Optimization, Academic Press, New York, 1969, pp. 283–298.
[23] D.P. Bertsekas, Constrained Optimization and Lagrange Multiplier Methods, Academic Press, New York, NY, 1982.
[24] R.T. Rockafellar, R.J.-B. Wets, Variational Analysis, Springer, Berlin, 1998.
[25] M.T. Chu, K.R. Driessel, The projected gradient methods for least squares matrix approximations with spectral constraints, SIAM Journal on Numerical Analysis 27 (4) (1990) 1050–1060.
[26] M. Graham, L. Gladwell, Inverse Problems in Vibration, second ed., Springer, Berlin, 2004.