AA242B: MECHANICAL VIBRATIONS

Solution Methods for the Generalized Eigenvalue Problem

These slides are based on the recommended textbook: M. Géradin and D. Rixen, "Mechanical Vibrations: Theory and Applications to Structural Dynamics," Second Edition, Wiley, John & Sons, Incorporated, ISBN-13:9780471975465

Outline

1 Eigenvector Iteration Methods

2 Subspace Construction Methods

- Generalized eigenvalue problem associated with an n-degree-of-freedom system

$$\boxed{\mathbf{K}\mathbf{x} = \omega^2 \mathbf{M}\mathbf{x}}$$

  - n eigenvalues

$$0 \leq \omega_1^2 \leq \omega_2^2 \leq \cdots \leq \omega_n^2$$

    and n associated eigenvectors

$$\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n$$

- Dynamic flexibility matrix
  - consider a system that has no rigid-body modes $\Rightarrow$ $\mathbf{K}$ is non-singular
  - construct the dynamic flexibility matrix defined as

$$\mathbf{D} = \mathbf{K}^{-1}\mathbf{M}$$

  - the associated eigenvalue problem is

$$\mathbf{K}\mathbf{x} = \omega^2\mathbf{M}\mathbf{x} \quad \Rightarrow \quad \mathbf{D}\mathbf{x} = \lambda\mathbf{x} \qquad \text{with} \qquad \lambda = \frac{1}{\omega^2}$$

    - eigenvectors are same as for the generalized eigenproblem $\mathbf{K}\mathbf{x} = \omega^2\mathbf{M}\mathbf{x}$
    - eigenvalues are given by

$$\lambda_i = \frac{1}{\omega_i^2} \Rightarrow \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$$

  - drawbacks of working with $\mathbf{D}$
    - building cost is $\mathcal{O}(n^3)$
    - non-symmetric matrix and as such must be stored completely
  - these drawbacks can be dealt with (for example, see next the inverse iteration form of the power iteration algorithm)

- Determination of the fundamental eigenmode
  - an iterative solution to $\mathbf{D}\mathbf{x} = \lambda\mathbf{x}$ can be obtained by considering the iteration

$$\boxed{\mathbf{z}_{p+1} = \mathbf{D}\mathbf{z}_p}$$

  - proof: expand $\mathbf{z}_0$ in the basis of the eigenmodes of the system

$$\mathbf{z}_0 = \sum_{i=1}^{n} \alpha_i \mathbf{x}_i$$

Then, at the p-th iteration,

$$\mathbf{z}_p = \mathbf{D}^p \mathbf{z}_0 = \sum_{i=1}^{n} \alpha_i \mathbf{D}^p \mathbf{x}_i = \sum_{i=1}^{n} \alpha_i \lambda_i^p \mathbf{x}_i = \lambda_1^p \left(\alpha_1 \mathbf{x}_1 + \sum_{i=2}^{n} \alpha_i \left(\frac{\lambda_i}{\lambda_1} \right)^p \mathbf{x}_i \right)$$

Recall that $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ and assume that $\lambda_1 > \lambda_2$

$$\Longrightarrow \qquad \mathbf{z}_p \to \lambda_1^p \alpha_1 \mathbf{x}_1 \Rightarrow \frac{\|\mathbf{z}_{p+1}\|}{\|\mathbf{z}_p\|} \to \lambda_1 \text{ as } p \to \infty$$

$\Longrightarrow$ the power iteration algorithm converges to the **lowest** eigenvalue ω_1^2

- Convergence analysis of the power iteration algorithm
  - $\forall p$, $\mathbf{z}_p = \lambda_1^p \mathbf{x}_1 + \lambda_2^p \mathbf{x}_2 + \cdots$ (recall that $\mathbf{x}_i$ is defined up to a multiplicative constant)

$$\Longrightarrow \frac{\mathbf{z}_{p+1}^T \mathbf{e}_j}{\mathbf{z}_p^T \mathbf{e}_j} = \frac{\lambda_1^{p+1}\mathbf{x}_1^T\mathbf{e}_j + \lambda_2^{p+1}\mathbf{x}_2^T\mathbf{e}_j + \cdots}{\lambda_1^p\mathbf{x}_1^T\mathbf{e}_j + \lambda_2^p\mathbf{x}_2^T\mathbf{e}_j + \cdots} = \frac{\lambda_1^{p+1}(\mathbf{x}_1 + r^{p+1}\mathbf{x}_2 + \cdots)^T\mathbf{e}_j}{\lambda_1^p(\mathbf{x}_1 + r^p\mathbf{x}_2 + \cdots)^T\mathbf{e}_j}$$

$$= \lambda_1 \frac{(\mathbf{x}_1 + r^{p+1}\mathbf{x}_2 + \cdots)^T\mathbf{e}_j}{(\mathbf{x}_1 + r^p\mathbf{x}_2 + \cdots)^T\mathbf{e}_j}$$

where $r = \dfrac{\lambda_2}{\lambda_1}$

  - define $a_j = \dfrac{\mathbf{x}_2^T\mathbf{e}_j}{\mathbf{x}_1^T\mathbf{e}_j}$ and assume that $|\lambda_2| < |\lambda_1| \Rightarrow r < 1$

$$\lambda_{1_p} = \frac{\mathbf{z}_{p+1}^T\mathbf{e}_j}{\mathbf{z}_p^T\mathbf{e}_j} = \frac{\lambda_1(1 + a_j r^{p+1} + \cdots)}{1 + a_j r^p + \cdots} \approx_{p\to\infty} \lambda_1(1 + a_j r^{p+1})(1 + a_j r^p)^{-1}$$

$$\approx_{p\to\infty} \lambda_1\left(1 + a_j r^p (r-1)\right)$$

  - number of iterations N after which λ_{1_N} is stabilized with t significant digits

$$\left|\frac{\lambda_{1_{N+1}} - \lambda_{1_N}}{\lambda_{1_N}}\right| \leq 10^{-t} \Rightarrow a_j r^N \approx 10^{-t} \Rightarrow N \approx \frac{t}{\log\left(\frac{\lambda_1}{\lambda_2}\right)}$$

$\Longrightarrow N$ — and therefore the convergence rate of the power algorithm — depends only on the ratio of the first two eigenvalues

- Determination of the higher modes by orthogonal deflation
  - once $\mathbf{x}_1$ is determined, the following orthogonal projection operator can be formed (orthogonal projection $\Leftrightarrow \mathbf{P}_1^2 = \mathbf{P}_1$ and $Ker(\mathbf{P}_1)$ is $\mathbf{M}$-orthogonal to span $(\mathbf{P}_1)$)

$$\mathbf{P}_1 = \mathbf{I} - \frac{\mathbf{x}_1 \mathbf{x}_1^T \mathbf{M}}{\mathbf{x}_1^T \mathbf{M} \mathbf{x}_1} \qquad (\Rightarrow \mathbf{P}_1 \mathbf{x}_1 = \mathbf{0})$$

  - next, the following starting vector can be constructed

$$\mathbf{z}_0^* = \mathbf{P}_1 \mathbf{z}_0 = \sum_{i=2}^{n} \alpha_i \mathbf{x}_i$$

    this vector does not have any component in the $\mathbf{x}_1$ "direction" $\Rightarrow$ the power method equipped with the starting vector $\mathbf{z}_0^\star$ converges **in principle** to $(\lambda_2, \mathbf{x}_2)$
  - **in practice**, the mode $\mathbf{x}_1$ reappears in the iterations because of roundoff errors; to remediate this issue, the following algorithm can be used instead

$$\begin{cases} \mathbf{z}_p^* & = & \mathbf{P}_1 \mathbf{z}_p \\ \mathbf{z}_{p+1} & = & \mathbf{D} \mathbf{z}_p^* \end{cases}$$

  - this approach can be generalized to extract higher modes

- Alternatively, one can construct a power iteration algorithm based on

$$\overline{\mathbf{D}} = \mathbf{M}^{-1}\mathbf{K}$$

in which case

- the associated eigenvalue problem is

$$\mathbf{K}\mathbf{x} = \omega^2\mathbf{M}\mathbf{x} \quad \Rightarrow \quad \overline{\mathbf{D}}\mathbf{x} = \omega^2\mathbf{x}$$

- the eigenvectors and eigenvalues of this problem are the same as for the generalized eigenproblem $\mathbf{K}\mathbf{x} = \omega^2\mathbf{M}\mathbf{x}$
- the eigenvalues can be sorted and labeled as follows

$$\omega_1^2 \leq \omega_2^2 \leq \cdots \leq \omega_n^2$$

- the power iteration algorithm converges to ω_n^2 – that is, to the **highest** eigenvalue
- the drawbacks of working with $\overline{\mathbf{D}}$ are identical to those of working with $\mathbf{D}$ and can be dealt with in a similar way

- Problem: it is not desirable to construct the dynamical matrix $\mathbf{D} = \mathbf{K}^{-1}\mathbf{M}$ (or $\overline{\mathbf{D}} = \mathbf{M}^{-1}\mathbf{K}$) at the base of the power iteration algorithm
- Solution: power iteration approach using the more computationally efficient iterate $\mathbf{z}_{p+1} = \mathbf{K}^{-1}(\mathbf{M}\mathbf{z}_p)$, which can be organized in two steps as follows

$$\begin{cases} \mathbf{y}_p & = & \mathbf{M}\mathbf{z}_p \\ \mathbf{K}\mathbf{z}_{p+1} & = & \mathbf{y}_p \end{cases}$$

  - incurs the solution of an algebraic problem of the form

$$\mathbf{K}\mathbf{q} = \mathbf{g}$$

  - factorization-based algorithm
    1. factorization of $\mathbf{K}$ (once)
    2. forward and backward backward substitutions (repeated)
- Recall that convergence to the highest λ is convergence to the lowest ω^2 ($\lambda = \dfrac{1}{\omega^2}$)

- Spectral shifting (by μ) of the eigenvalue problem

$$(\mathbf{K} - \mu\mathbf{M})\mathbf{x} = (\omega^2 - \mu)\mathbf{M}\mathbf{x} \quad \Rightarrow \quad (\mathbf{K} - \mu\mathbf{M})^{-1}\mathbf{M}\mathbf{x} = \frac{1}{\omega^2 - \mu}\mathbf{x}$$

  - associated inverse iteration scheme

$$\begin{cases} \mathbf{y}_p & = & \mathbf{M}\mathbf{z}_p \\ (\mathbf{K} - \mu\mathbf{M})\mathbf{z}_{p+1} & = & \mathbf{y}_p \end{cases}$$

  - iterations

$$\mathbf{z}_0 = \sum_{i=1}^{n} \alpha_i \mathbf{x}_i \Rightarrow \mathbf{z}_p = \sum_{i=1}^{n} \alpha_i \left((\mathbf{K} - \mu\mathbf{M})^{-1}\mathbf{M} \right)^p \mathbf{x}_i = \sum_{i=1}^{n} \frac{\alpha_i}{(\omega_i^2 - \mu)^p} \mathbf{x}_i$$

converge toward the eigenmode $\mathbf{x}_r$ whose associated eigenvalue ω_r^2 satisfies

$$|\omega_r^2 - \mu| = \min_j \left\{ |\omega_j^2 - \mu| \right\}$$

and therefore is closest to μ, and this convergence accelerates as μ approaches ω_r^2

- The subspace concept
  - let $\mathbf{X}$ denote the $\mathbf{M}$-normalized solution of the generalized eigenvalue problem $\mathbf{K}\mathbf{x} = \omega^2\mathbf{M}\mathbf{x}$

$$\mathbf{X}^T\mathbf{M}\mathbf{X} = \mathbf{I}_n, \quad \mathbf{X}^T\mathbf{K}\mathbf{X} = \mathbf{\Lambda} = \mathrm{diag}(\omega_1^2, \cdots, \omega_n^2)$$

    then $\mathrm{span}(\mathbf{X}) = \mathbb{E}$ of dimension n
  - consider the first $m \ll n$ eigenmodes

$$\mathbf{X}^\star = \begin{bmatrix} \mathbf{x}_1 & \cdots & \mathbf{x}_m \end{bmatrix} \in \mathbb{R}^{n\times m}$$

    $\mathrm{span}(\mathbf{X}^\star) = \mathbb{E}^\star$ of dimension m; $\mathbb{E}^\star \subset \mathbb{E}$
  - property 1: $\forall \mathbf{a} \in \mathbb{R}^m$

$$\mathbf{z} = \mathbf{X}^\star\mathbf{a} \in \mathbb{E}^\star$$

  - property 2: if $\mathbf{A} \in \mathbb{R}^{m\times m}$ and $\det(\mathbf{A}) \neq 0$

$$\mathrm{span}(\mathbf{X}^\star\mathbf{A}) = \mathrm{span}(\mathbf{X}^\star)$$

- Interaction problem
  - once any set $\mathbf{Z}^\star$ of $m \ll n$ linearly independent vectors in $\mathbb{E}^\star$ is found ($\mathbf{Z}^\star \in \mathbb{R}^{n\times m}$)
    - the first m eigensolutions are determined by forming and solving the **reduced** eigenproblem
      $$\mathbf{K}^\star \mathbf{y} = \omega^2 \mathbf{M}^\star \mathbf{y}$$
      where $\mathbf{K}^\star = \mathbf{Z}^{\star T}\mathbf{K}\mathbf{Z}^\star \in \mathbb{R}^{m\times m}$ and $\mathbf{M}^\star = \mathbf{Z}^{\star T}\mathbf{M}\mathbf{Z}^\star \in \mathbb{R}^{m\times m}$
    - the matrix $\mathbf{Y} = \begin{bmatrix} \mathbf{y}_1 & \cdots & \mathbf{y}_m \end{bmatrix} \in \mathbb{R}^{m\times m}$ is assembled
    - the matrix $\mathbf{X}^\star = \mathbf{Z}^\star \mathbf{Y} \in \mathbb{R}^{n\times m}$, which has the same span as the matrix $\mathbf{Z}^\star$, is constructed

- Using the inverse iteration algorithm, the subspace method generates a sequence of matrices $\mathbf{Z}_0, \mathbf{Z}_1, \cdots, \mathbf{Z}_k \in \mathbb{R}^{n\times m}$ spanning the subspaces $\mathbb{E}_0^\star, \mathbb{E}_1^\star, \cdots, \mathbb{E}_k^\star$ that converge toward $\mathbb{E}_\infty^\star = \mathbb{E}^\star = \text{span}(\mathbf{X}^\star)$ (iterative method on the Grassmann manifold $\mathcal{G}(m,n)$)
- Algorithm
  - iteration kernel
$$\begin{cases} \mathbf{V}_k & = & \mathbf{M}\mathbf{Z}_k \\ \mathbf{K}\mathbf{Z}_{k+1} & = & \mathbf{V}_k \end{cases}$$
  - aim is convergence toward $\mathbb{E}^\star$ of the subspace as a whole, and not of the individual vectors in $\mathbf{Z}_k$
  - to make sure that the reduced-order bases $\mathbf{Z}_i$ span subspaces of dimension m, orthogonality of the column vectors is ensured by one of two ways
    - Gram-Schmidt orthogonalization
    - re-orthogonalization with respect to the $\mathbf{M}$-induced norm

- Convergence analysis
  - the subspaces spanned by the matrices $\mathbf{Z}_k$ converge toward the subspace spanned by $\mathbf{X}^\star$, provided that $\mathbf{Z}_0$ is not orthogonal to any of the first m eigenvectors
  - global convergence of the subspace iteration algorithm guarantees that subspace $\mathbb{E}_k^\star \subset \mathbb{E}$ converges toward $\mathbb{E}^\star$
  - the convergence rate toward each mode $\mathbf{x}_j$ is of the order of $\mathcal{O}(r_j^k)$ where
    $$r_j = \frac{\omega_j^2}{\omega_{m+1}^2}$$
    measures the ratio of the eigenvalue ω_j^2 and that which is immediately outside the subspace $\mathbb{E}^\star$. This indicates that a higher convergence rate is obtained by using $q > m$ iteration vectors (which increases computational cost however). A practical trade-off suggests to use $q = \min(2m, m+8)$.
  - considering the convergence rate mentioned above, multiple eigenvalues do not decrease the rate of convergence provided that $\omega_{q+1}^2 > \omega_m^2$

- Principle of the method
  - generate a subspace that includes the fundamental eigensolutions by applying the inverse iteration algorithm to a starting vector $\mathbf{z}_0$
  - construct the following Krylov sequence

$$\left\{ \mathbf{z}_0, \mathbf{K}^{-1}\mathbf{M}\mathbf{z}_0, \left(\mathbf{K}^{-1}\mathbf{M}\right)^2 \mathbf{z}_0, \cdots \right\}$$

  - orthogonalize the terms of the sequence
- Properties (for the basic form of the algorithm)
  - the search subspace is built column by column: its dimension is increased at each iteration
  - fast convergence
  - however, orthonormality rapidly deteriorates (and re-orthogonalization procedures can be computationally expensive)
  - inability to detect multiple eigenvalues

- Algorithm
  - iteration

$$\gamma_{p+1}\mathbf{x}_{p+1} = \mathbf{K}^{-1}\mathbf{M}\mathbf{x}_p - \alpha_p\mathbf{x}_p - \beta_{p-1}\mathbf{x}_{p-1}$$

where $\beta_{p-1},\ \alpha_p,\ \gamma_{p+1}$ are determined such that

$$\begin{cases} \mathbf{x}_{p+1}^T\mathbf{M}\mathbf{x}_j & = & 0 \quad j < p+1 \\ \mathbf{x}_{p+1}^T\mathbf{M}\mathbf{x}_{p+1} & = & 1 \end{cases}$$

  - requiring **M**-orthogonality leads to

$$\begin{cases} \alpha_p & = & \mathbf{x}_p^T\mathbf{M}\mathbf{K}^{-1}\mathbf{M}\mathbf{x}_p \\ \beta_{p-1} & = & \mathbf{x}_p^T\mathbf{M}\mathbf{K}^{-1}\mathbf{M}\mathbf{x}_{p-1} \\ \gamma_p & = & \beta_{p-1} \end{cases}$$

  - the reduced-order basis $\mathbf{X} = \begin{bmatrix} \mathbf{x}_0 & \cdots & \mathbf{x}_p \end{bmatrix} \in \mathbb{R}^{n\times p}$ satisfies

$$\mathbf{K}^{-1}\mathbf{M}\mathbf{X} = \mathbf{X}\mathbf{T} + \mathbf{S} = \begin{bmatrix} \mathbf{0} & \cdots & \mathbf{0} & \gamma_{p+1}\mathbf{x}_{p+1} \end{bmatrix} = \gamma_{p+1}\mathbf{x}_{p+1}\mathbf{e}_p^T$$

where $\mathbf{T} =$
$\text{tridiag}\left(\begin{bmatrix} \gamma_1 & \cdots & \gamma_p \end{bmatrix}, \begin{bmatrix} \alpha_0 & \cdots & \alpha_p \end{bmatrix}, \begin{bmatrix} \gamma_1 & \cdots & \gamma_p \end{bmatrix}\right)$

- Interaction problem
  - start with
$$\mathbf{K}^{-1}\mathbf{M}\mathbf{X} = \mathbf{X}\mathbf{T} + \mathbf{S}$$
  - premultiply by $\mathbf{Y}^T = (\mathbf{M}\mathbf{X})^T$ associated to the Lanczos vectors $\mathbf{X}$
  - note that $\mathbf{Y}^T\mathbf{S} = \mathbf{0}$
$$\Rightarrow \mathbf{Y}^T\mathbf{K}^{-1}\mathbf{M}\mathbf{X} = \mathbf{Y}^T\mathbf{X}\mathbf{T} = \mathbf{T}$$
  - hence, the tridiagonal matrix of the Lanczos coefficients results from the projection of $\mathbf{D} = \mathbf{K}^{-1}\mathbf{M}$ onto the subspace of Lanczos vectors $\mathbf{X}$. Thus, it has the same eigenvalues as the projection of $\mathbf{D}$ onto the reduced-order basis $\mathbf{X}$.

- Difficulties for basic form of the Lanczos method
  - deterioration of the orthogonalization process
  - detection of multiple eigenvalues
  - appearance of parasitic solutions
  - missing eigensolutions