Dynamic Behavior of Complex Structures
CHAPTER 1 : PRELIMINARIES

Heinrich Voss

voss@tu-harburg.de

Hamburg University of Technology
Institute of Numerical Simulation

TUHH
Technische Universität Hamburg-Harburg

Sparse Eigenvalue Problems

For sparse linear eigenproblems

$$Ax = \lambda x$$

most of the standard solvers exploit projection processes in order to extract approximate eigenpairs from a given subspace.

Sparse Eigenvalue Problems

For sparse linear eigenproblems

$$Ax = \lambda x$$

most of the standard solvers exploit projection processes in order to extract approximate eigenpairs from a given subspace.

Differently from eigensolvers for dense matrices no similarity transformations are applied to the system matrix A in order to transform A to (block-) diagonal or (block-) triangular form and to obtain the eigenvalues and corresponding eigenvectors immediately.

For sparse linear eigenproblems

$$Ax = \lambda x$$

most of the standard solvers exploit projection processes in order to extract approximate eigenpairs from a given subspace.

Differently from eigensolvers for dense matrices no similarity transformations are applied to the system matrix A in order to transform A to (block-) diagonal or (block-) triangular form and to obtain the eigenvalues and corresponding eigenvectors immediately.

Typically, the explicit form of the matrix A is not needed but only a function

$$y \leftarrow Ax$$

yielding the matrix–vector product Ax for a given vector x.

Power method

Choose initial vector u^1
for $j = 1, 2, \dots$ until convergence **do**
  $u = Au^j$
  $u^{j+1} = u / \|u\|_2$
  $\mu_{j+1} = (u^{j+1})^H A u^{j+1}$
end for

Power method

1: Choose initial vector u^1
2: **for** $j = 1, 2, \dots$ until convergence **do**
3: $\quad u = Au^j$
4: $\quad u^{j+1} = u/\|u\|_2$
5: $\quad \mu_{j+1} = (u^{j+1})^H A u^{j+1}$
6: **end for**

If A is diagonalizable, $|\lambda_1| > |\lambda_j|$, $j = 2, 3, \dots, n$ are the eigenvalues of A, and $x^1, \dots, x^n$ are corresponding eigenvectors, then

$$u^1 = \sum_{i=1}^{n} \alpha_i x^i \quad \Longrightarrow \quad u^{j+1} = \xi A^j u^1 = \xi \lambda_1^j \Big(\alpha_1 x^1 + \underbrace{\sum_{i=2}^{n} \alpha_i \big(\frac{\lambda_i}{\lambda_1}\big)^j x^i}_{\to 0} \Big)$$

Power method

Choose initial vector u^1
for $j = 1, 2, \dots$ until convergence **do**
  $u = Au^j$
  $u^{j+1} = u / \|u\|_2$
  $\mu_{j+1} = (u^{j+1})^H A u^{j+1}$
end for

If A is diagonalizable, $|\lambda_1| > |\lambda_j|$, $j = 2, 3, \dots, n$ are the eigenvalues of A, and $x^1, \dots, x^n$ are corresponding eigenvectors, then

$$u^1 = \sum_{i=1}^{n} \alpha_i x^i \quad \Longrightarrow \quad u^{j+1} = \xi A^j u^1 = \xi \lambda_1^j \Big(\alpha_1 x^1 + \underbrace{\sum_{i=2}^{n} \alpha_i \big(\frac{\lambda_i}{\lambda_1}\big)^j x^i}_{\to 0} \Big)$$

Hence, if A has a dominant eigenvalue λ_1 which is simple, then a scaled version of u^j converges to an eigenvector of A corresponding to λ_1.

Eigenextraction

If $|\lambda_1| = |\lambda_2| > |\lambda_j|$, $j = 3, \dots, n$, $\lambda_1 \neq \lambda_2$, then

$$u^{j+1} = \xi A^j u^1 = \xi \lambda_1^j \Big(\alpha_1 x^1 + \alpha_2 \underbrace{\Big(\frac{\lambda_2}{\lambda_1}\Big)^j}_{\neq 1,\ |\cdot|=1} x^2 + \underbrace{\sum_{i=3}^{n} \alpha_i \Big(\frac{\lambda_i}{\lambda_1}\Big)^j x^i}_{\to 0} \Big).$$

Hence, for j large $\operatorname{span}\{u^{j+1}, u^{j+2}\}$ tends to $\operatorname{span}\{x^1, x^2\}$.

Eigenextraction

If $|\lambda_1| = |\lambda_2| > |\lambda_j|$, $j = 3, \dots, n$, $\lambda_1 \neq \lambda_2$, then

$$u^{j+1} = \xi A^j u^1 = \xi \lambda_1^j \Big(\alpha_1 x^1 + \alpha_2 \underbrace{\Big(\frac{\lambda_2}{\lambda_1}\Big)^j}_{\neq 1,\ |\cdot|=1} x^2 + \underbrace{\sum_{i=3}^n \alpha_i \Big(\frac{\lambda_i}{\lambda_1}\Big)^j x^i}_{\to 0} \Big).$$

Hence, for j large $\text{span}\{u^{j+1}, u^{j+2}\}$ tends to $\text{span}\{x^1, x^2\}$.

To extract approximate eigenvectors from a 2 dimensional subspace $\mathcal{V} := \text{span}\{v^1, v^2\}$, write them as linear combinations of v^1 and v^2

$$\tilde{u} = \eta_1 v^1 + \eta_2 v^2,$$

and determine η_1, η_2 and $\tilde{\lambda}$ from the requirement that the residual is orthogonal to v^1 and v^2:

$$A\tilde{u} - \tilde{\lambda}\tilde{u} \perp v^1, \ A\tilde{u} - \tilde{\lambda}\tilde{u} \perp v^2.$$

Eigenextraction

If $|\lambda_1| = |\lambda_2| > |\lambda_j|$, $j = 3, \dots, n$, $\lambda_1 \neq \lambda_2$, then

$$u^{j+1} = \xi A^j u^1 = \xi \lambda_1^j \Big(\alpha_1 x^1 + \alpha_2 \underbrace{\Big(\frac{\lambda_2}{\lambda_1}\Big)^j}_{\neq 1,\ |\cdot|=1} x^2 + \underbrace{\sum_{i=3}^{n} \alpha_i \Big(\frac{\lambda_i}{\lambda_1}\Big)^j x^i}_{\to 0} \Big).$$

Hence, for j large $\text{span}\{u^{j+1}, u^{j+2}\}$ tends to $\text{span}\{x^1, x^2\}$.

To extract approximate eigenvectors from a 2 dimensional subspace $\mathcal{V} := \text{span}\{v^1, v^2\}$, write them as linear combinations of v^1 and v^2

$$\tilde{u} = \eta_1 v^1 + \eta_2 v^2,$$

and determine η_1, η_2 and $\tilde{\lambda}$ from the requirement that the residual is orthogonal to v^1 and v^2:

$$A\tilde{u} - \tilde{\lambda}\tilde{u} \perp v^1, \; A\tilde{u} - \tilde{\lambda}\tilde{u} \perp v^2.$$

With $V = [v^1, v^2]$ and $y = (\eta_1, \eta_2)^T$ we have $\tilde{u} = Vy$, and the last condition reads

$$V^H(A - \tilde{\lambda} I)Vy = 0, \quad \text{i.e.} \quad V^H A V y = \tilde{\lambda} V^H V y,$$

which is a generalized 2×2 eigenvalue problem.

A projection method consists of approximating an eigenvector u by a vector $\tilde{u}$ belonging to some subspace $\mathcal{V}$ (the subspace of approximants or search space or right subspace) requiring that the residual is orthogonal to some subspace $\mathcal{W}$ (the left subspace) where $\dim \mathcal{V} = \dim \mathcal{W}$.

Projection methods

A projection method consists of approximating an eigenvector u by a vector $\tilde{u}$ belonging to some subspace $\mathcal{V}$ (the subspace of approximants or search space or right subspace) requiring that the residual is orthogonal to some subspace $\mathcal{W}$ (the left subspace) where $\dim \mathcal{V} = \dim \mathcal{W}$.

Methods of this type are called Petrov–Galerkin method, and for $\mathcal{V} = \mathcal{W}$ Galerkin method or Bubnov–Galerkin method.

Projection methods

A projection method consists of approximating an eigenvector u by a vector $\tilde{u}$ belonging to some subspace $\mathcal{V}$ (the subspace of approximants or search space or right subspace) requiring that the residual is orthogonal to some subspace $\mathcal{W}$ (the left subspace) where $\dim \mathcal{V} = \dim \mathcal{W}$.

Methods of this type are called Petrov–Galerkin method, and for $\mathcal{V} = \mathcal{W}$ Galerkin method or Bubnov–Galerkin method.

If $\mathcal{W} = \mathcal{V}$ then the method is called orthogonal projection method, if $\mathcal{W} \neq \mathcal{V}$ then the method is called oblique projection method.

Orthogonal projection method

An orthogonal projection method onto the search space $\mathcal{V}$ seeks an approximate eigenpair $(\tilde{\lambda}, \tilde{u})$ of $Ax = \lambda x$ with $\tilde{\lambda} \in \mathbb{C}$ and $\tilde{u} \in \mathcal{V}$ such that

$$v^H(A\tilde{u} - \tilde{\lambda}\tilde{u}) = 0 \quad \text{for every } v \in \mathcal{V}.$$

Orthogonal projection method

An orthogonal projection method onto the search space $\mathcal{V}$ seeks an approximate eigenpair $(\tilde\lambda, \tilde u)$ of $Ax = \lambda x$ with $\tilde\lambda \in \mathbb{C}$ and $\tilde u \in \mathcal{V}$ such that

$$v^H(A\tilde u - \tilde\lambda \tilde u) = 0 \quad \text{for every } v \in \mathcal{V}.$$

If $v^1, \dots, v^m$ denotes an orthonormal basis of $\mathcal{V}$ and $V = [v^1, \dots, v^n]$ then $\tilde u$ has a representation $\tilde u = Vy$ with $y \in \mathbb{C}^m$, and the orthogonality condition obtains the form

$$B_m y := V^H A V y = \lambda y,$$

i.e. eigenvalues $\tilde\lambda$ of the $m \times m$ matrix B_m approximate eigenvalues of A, and if $\tilde y$ is a corresponding eigenvector of B_m then $\tilde u = V\tilde y$ is an approximate eigenvector of A.

Orthogonal projection method

An orthogonal projection method onto the search space $\mathcal{V}$ seeks an approximate eigenpair $(\tilde{\lambda}, \tilde{u})$ of $Ax = \lambda x$ with $\tilde{\lambda} \in \mathbb{C}$ and $\tilde{u} \in \mathcal{V}$ such that

$$v^H(A\tilde{u} - \tilde{\lambda}\tilde{u}) = 0 \quad \text{for every } v \in \mathcal{V}.$$

If $v^1, \dots, v^m$ denotes an orthonormal basis of $\mathcal{V}$ and $V = [v^1, \dots, v^n]$ then $\tilde{u}$ has a representation $\tilde{u} = Vy$ with $y \in \mathbb{C}^m$, and the orthogonality condition obtains the form

$$B_m y := V^H A V y = \lambda y,$$

i.e. eigenvalues $\tilde{\lambda}$ of the $m \times m$ matrix B_m approximate eigenvalues of A, and if $\tilde{y}$ is a corresponding eigenvector of B_m then $\tilde{u} = V\tilde{y}$ is an approximate eigenvector of A.

An orthogonal projection method is called Rayleigh–Ritz method, $\tilde{\lambda}$ is called Ritz value, and $\tilde{u}$ corresponding Ritz vector. $(\tilde{\lambda}, \tilde{u})$ is called Ritz pair with respect to $\mathcal{V}$.

Oblique projection method

In an oblique projection method we are given two subspaces $\mathcal{V}$ and $\mathcal{W}$, and we seek $\tilde{\lambda} \in \mathbb{C}$ and $\tilde{u} \in \mathcal{V}$ such that

$$w^H(A - \tilde{\lambda} I)\tilde{u} = 0 \quad \text{for every } w \in \mathcal{W}.$$

Oblique projection method

In an oblique projection method we are given two subspaces $\mathcal{V}$ and $\mathcal{W}$, and we seek $\tilde{\lambda} \in \mathbb{C}$ and $\tilde{u} \in \mathcal{V}$ such that

$$w^H(A - \tilde{\lambda} I)\tilde{u} = 0 \quad \text{for every } w \in \mathcal{W}.$$

Let $W = [w^1, \dots, w^m]$ be a basis of $\mathcal{W}$, and $V = [v^1, \dots, v^m]$ be a basis of $\mathcal{V}$. We assume that these two bases are biorthogonal, i.e. $(w^i)^H v^j = \delta_{ij}$ or $W^H V = I_m$. Then writing $\tilde{u} = Vy$ as before the Petrov–Galerkin condition reads

$$B_m y := W^H A V y = \lambda y.$$

Oblique projection method

In an oblique projection method we are given two subspaces $\mathcal{V}$ and $\mathcal{W}$, and we seek $\tilde{\lambda} \in \mathbb{C}$ and $\tilde{u} \in \mathcal{V}$ such that

$$w^H(A - \tilde{\lambda} I)\tilde{u} = 0 \quad \text{for every } w \in \mathcal{W}.$$

Let $W = [w^1, \dots, w^m]$ be a basis of $\mathcal{W}$, and $V = [v^1, \dots, v^m]$ be a basis of $\mathcal{V}$. We assume that these two bases are biorthogonal, i.e. $(w^i)^H v^j = \delta_{ij}$ or $W^H V = I_m$. Then writing $\tilde{u} = Vy$ as before the Petrov–Galerkin condition reads

$$B_m y := W^H A V y = \lambda y.$$

The terms Ritz value, Ritz vector, and Ritz pair are defined in an analogous way as for the orthogonal projection method.

Oblique projection method ct.

- In order for biorthogonal bases to exist the following assumption for $\mathcal{V}$ and $\mathcal{W}$ must hold:
  For any two bases V and W of $\mathcal{V}$ and $\mathcal{W}$, respectively,
  $$\det(W^H V) \neq 0.$$
  Obviously, this condition does not depend on the particular bases selected, and it is equivalent to requiring that no vector in $\mathcal{V}$ be orthogonal to $\mathcal{W}$.

Oblique projection method ct.

- In order for biorthogonal bases to exist the following assumption for $\mathcal{V}$ and $\mathcal{W}$ must hold:
  For any two bases V and W of $\mathcal{V}$ and $\mathcal{W}$, respectively,
  $$\det(W^H V) \neq 0.$$
  Obviously, this condition does not depend on the particular bases selected, and it is equivalent to requiring that no vector in $\mathcal{V}$ be orthogonal to $\mathcal{W}$.

- The approximate problem obtained from oblique projection has the potential of being much worse conditioned than with orthogonal projection methods.

Oblique projection method ct.

- In order for biorthogonal bases to exist the following assumption for $\mathcal{V}$ and $\mathcal{W}$ must hold:
  For any two bases V and W of $\mathcal{V}$ and $\mathcal{W}$, respectively,
  $$\det(W^H V) \neq 0.$$
  Obviously, this condition does not depend on the particular bases selected, and it is equivalent to requiring that no vector in $\mathcal{V}$ be orthogonal to $\mathcal{W}$.

- The approximate problem obtained from oblique projection has the potential of being much worse conditioned than with orthogonal projection methods.

- Problems obtained from oblique projection may be able to compute good approximations to both, left and right eigenvectors, simultaneously.

Oblique projection method ct.

- In order for biorthogonal bases to exist the following assumption for $\mathcal{V}$ and $\mathcal{W}$ must hold:
  For any two bases V and W of $\mathcal{V}$ and $\mathcal{W}$, respectively,

  $$\det(W^H V) \neq 0.$$

  Obviously, this condition does not depend on the particular bases selected, and it is equivalent to requiring that no vector in $\mathcal{V}$ be orthogonal to $\mathcal{W}$.

- The approximate problem obtained from oblique projection has the potential of being much worse conditioned than with orthogonal projection methods.

- Problems obtained from oblique projection may be able to compute good approximations to both, left and right eigenvectors, simultaneously.

- There are methods based on oblique projection which require much less storage than similar orthogonal projection methods.

Error bound

Let $(\tilde{\lambda}, \tilde{u})$ be an approximation to an eigenpair of A. If A is normal, i.e. $AA^H = A^H A$, then the following error estimate holds.

Error bound

Let $(\tilde{\lambda}, \tilde{u})$ be an approximation to an eigenpair of A. If A is normal, i.e. $AA^H = A^H A$, then the following error estimate holds.

THEOREM

Let $\lambda_1, \dots, \lambda_n$ be the eigenvalues of the normal matrix A. then it holds

$$\min_{j=1,\dots,n} |\lambda_j - \tilde{\lambda}| \le \frac{\|r\|_2}{\|\tilde{u}\|_2} \quad \text{where } r := A\tilde{u} - \tilde{\lambda}\tilde{u}.$$

Proof

Let $u^1, \dots, u^n$ be a unitary basis of eigenvectors of A. Then it holds

$$\tilde{u} = \sum_{i=1}^{n} (u^i)^H \tilde{u} \cdot u^i, \; A\tilde{u} = \sum_{i=1}^{n} \lambda_i (u^i)^H \tilde{u} \cdot u^i, \; \|\tilde{u}\|_2^2 = \sum_{i=1}^{n} |\tilde{u}^H u^i|^2.$$

Proof

Let $u^1, \dots, u^n$ be a unitary basis of eigenvectors of A. Then it holds

$$\tilde{u} = \sum_{i=1}^{n} (u^i)^H \tilde{u} \cdot u^i, \; A\tilde{u} = \sum_{i=1}^{n} \lambda_i (u^i)^H \tilde{u} \cdot u^i, \; \|\tilde{u}\|_2^2 = \sum_{i=1}^{n} |\tilde{u}^H u^i|^2.$$

Hence,

$$\begin{aligned} \|A\tilde{u} - \tilde{\lambda}\tilde{u}\|_2^2 &= \|\sum_{i=1}^{n} (\lambda_i - \tilde{\lambda})(u^i)^H \tilde{u} \cdot u^i\|_2^2 = \sum_{i=1}^{n} |\lambda_i - \tilde{\lambda}|^2 |\tilde{u}^H u^i|^2 \\ &\geq \min_{i=1,\dots,n} |\lambda_i - \tilde{\lambda}|^2 \sum_{i=1}^{n} |\tilde{u}^H u^i|^2 = \min_{i=1,\dots,n} |\lambda_i - \tilde{\lambda}|^2 \|\tilde{u}\|_2^2, \end{aligned}$$

from which we obtain the error bound.

Backward error

For general matrices

$$\frac{\|r\|_2}{\|\tilde{u}\|_2} \quad \text{with } r := A\tilde{u} - \tilde{\lambda}\tilde{u}.$$

is the backward error

$$\min_E \{\|E\|_2 \,:\, (A+E)\tilde{u} = \tilde{\lambda}\tilde{u}\}$$

of $(\tilde{\lambda}, \tilde{u})$.

Backward error

For general matrices

$$\frac{\|r\|_2}{\|\tilde{u}\|_2} \quad \text{with } r := A\tilde{u} - \tilde{\lambda}\tilde{u}.$$

is the backward error

$$\min_E \{\|E\|_2 \;:\; (A+E)\tilde{u} = \tilde{\lambda}\tilde{u}\}$$

of $(\tilde{\lambda}, \tilde{u})$.

This follows from

$$(A+E)\tilde{u} = \tilde{\lambda}\tilde{u} \quad \Rightarrow \quad E\tilde{u} = r \quad \Rightarrow \quad \|E\|_2 \geq \|\frac{\|E\tilde{u}\|_2}{\|\tilde{u}\|_2} = \frac{\|r\|_2}{\|\tilde{u}\|_2},$$

and on the other hand we have for $E := -r\tilde{u}^H/\|\tilde{u}\|_2^2$

$$\|E\|_2^2 = \rho(E^H E) = \frac{1}{\|\tilde{u}\|_2^4}\rho(\tilde{u}r^H r\tilde{u}^H) = \frac{\|r\|_2}{\|\tilde{u}\|_2}.$$

Iterative projection methods

The dimension of the eigenproblem is reduced by projecting it upon a subspace of small dimension. The reduced problem is handled by a fast technique for dense problems.

Iterative projection methods

The dimension of the eigenproblem is reduced by projecting it upon a subspace of small dimension. The reduced problem is handled by a fast technique for dense problems.

The errors of approximating Ritz pairs to wanted eigenvalues are 'estimated'.

Iterative projection methods

The dimension of the eigenproblem is reduced by projecting it upon a subspace of small dimension. The reduced problem is handled by a fast technique for dense problems.

The errors of approximating Ritz pairs to wanted eigenvalues are 'estimated'.

If an error tolerance is not met the search space is expanded in the course of the algorithm in an iterative way with the aim that some of the eigenvalues of the reduced matrix become good approximations of some of the wanted eigenvalues of the given large matrix.

General iterative projection method

Choose initial vector u^1 with $\|u^1\| = 1$, $U_1 = [u^1]$
for $j = 1, 2, \dots$ until convergence **do**
    $w^j = Au^j$
    for $k = 1, \dots, j-1$ **do**
        $b_{kj} = (u^k)^H w^j$
        $b_{jk} = (u^j)^H w^k$
    end for
    $b_{jj} = (u^j)^H w^j$
    Determine wanted eigenvalue θ of B
    and corresponding eigenvector s such that $\|s\| = 1$
    $y = U_j s$
    $r = Ay - \theta y$
    Determine expansion direction q
    $q = q - U_j U_j^H q$
    $u^{j+1} = q / \|q\|$
    $U_{j+1} = [U_j, u^{j+1}]$
end for

Two types of iterative projection methods

Krylov subspace methods: like the Lanczos, Arnoldi, and rational Krylov method, where the expansion by

$$q = A * \text{last column of } V$$

is independent of the eigensolution of the reduced problem. Problem is projected to Krylov space

$$\mathcal{K}_k(v^1, A) = \text{span}\{v^1, Av^1, A^2v^1, \dots, A^{k-1}v^1\}.$$

Two types of iterative projection methods

Krylov subspace methods: like the Lanczos, Arnoldi, and rational Krylov method, where the expansion by

$$q = A * \text{last column of } V$$

is independent of the eigensolution of the reduced problem. Problem is projected to Krylov space

$$\mathcal{K}_k(v^1, A) = \text{span}\{v^1, Av^1, A^2v^1, \dots, A^{k-1}v^1\}.$$

General iterative projection methods: like the Davidson, or the Jacobi–Davidson method where the expansion direction q is chosen such that the resulting search space has a high approximation potential for the eigenvector wanted next.