Preconditioning eigenvalue problems and other approaches

- *Preconditioning eigenvalue problems: Shift-invert, polynomial*
- *Polyninial filters, Implicit restarts*
- *The Davidson approach*
- *Jacobi-Davidson*
- *Harmonic Ritz values*

Preconditioning eigenvalue problems

➤ Goal: To extract good approximations to add to a subspace in a projection process. Result: faster convergence.

➤ Best known technique: Shift-and-invert; Work with

$$B = (A - \sigma I)^{-1}$$

➤ Some success with polynomial preconditioning [Chebyshev iteration / least-squares polynomials]. Work with

$$B = p(A)$$

➤ Above preconditioners preserve eigenvectors. Other methods (Davidson) use a more general preconditioner M.

Shift-and-invert preconditioning

Main idea: to use Arnoldi, or Lanczos, or subspace iteration for the matrix $B = (A - \sigma I)^{-1}$. The matrix B need not be computed explicitly. Each time we need to apply B to a vector we solve a system with B.

➤ Factor $B = A - \sigma I = LU$. Then each solution $Bx = y$ requires solving $Lz = y$ and $Ux = z$.

How to deal with complex shifts?

➤ If A is complex need to work in complex arithmetic.

➤ If A is real, then instead of $(A - \sigma I)^{-1}$ use

$$\Re e(A - \sigma I)^{-1} = \frac{1}{2}\left[(A - \sigma I)^{-1} + (A - \bar{\sigma} I)^{-1}\right]$$

Preconditioning by polynomials

Main idea:

Iterate with $p(A)$ instead of A in Arnoldi or Lanczos,..

➤ Used very early on in subspace iteration [Rutishauser, 1959.]

➤ Usually not as reliable as Shift-and-invert techniques but less demanding in terms of storage.

Question: How to find a good polynomial (dynamically)?

Approaches:

1. Use of Chebyshev polynomials over ellipses
2. Use polynomials based on Leja points
3. Least-squares polynomials over polygons
4. Polynomials from previous Arnoldi decompositions

Polynomial filters and implicit restart

Goal: exploit the Arnoldi procedure to apply polynomial filter of the form: $p(t) = (t - \theta_1)(t - \theta_2)\ldots(t - \theta_q)$

Assume $AV_m = V_m H_m + \hat{v}_{m+1} e_m^T$
and consider first factor: $(t - \theta_1)$

$$(A - \theta_1 I)V_m = V_m(H_m - \theta_1 I) + \hat{v}_{m+1} e_m^T$$

Let $H_m - \theta_1 I = Q_1 R_1$. Then,

$$\begin{aligned} (A - \theta_1 I)V_m &= V_m Q_1 R_1 + \hat{v}_{m+1} e_m^T \quad \rightarrow \\ (A - \theta_1 I)(V_m Q_1) &= (V_m Q_1) R_1 Q_1 + \hat{v}_{m+1} e_m^T Q_1 \rightarrow \\ A(V_m Q_1) &= (V_m Q_1)(R_1 Q_1 + \theta_1 I) + \hat{v}_{m+1} e_m^T Q_1 \end{aligned}$$

Notation:

$R_1 Q_1 + \theta_1 I \equiv H_m^{(1)}$; $\quad (b_{m+1}^{(1)})^T \equiv e_m^T Q_1$; $\quad V_m Q_1 \equiv V_m^{(1)}$

- $AV_m^{(1)} = V_m^{(1)} H_m^{(1)} + v_{m+1}(b_{m+1}^{(1)})^T$
- Note that $H_m^{(1)}$ is upper Hessenberg.
- Similar to an Arnoldi decomposition.

Observe:

- $R_1 Q_1 + \theta_1 I \equiv$ matrix resulting from one step of the QR algorithm with shift θ_1 applied to H_m.
- First column of $V_m^{(1)}$ is a multiple of $(A - \theta_1 I)v_1$.
- The columns of $V_m^{(1)}$ are orthonormal.

Can now apply second shift in same way:

$$(A - \theta_2 I)V_m^{(1)} = V_m^{(1)}(H_m^{(1)} - \theta_2 I) + v_{m+1}(b_{m+1}^{(1)})^T \quad \rightarrow$$

Similar process: $(H_m^{(1)} - \theta_2 I) = Q_2 R_2$ then $\times Q_2$ to the right:

$$(A - \theta_2 I)V_m^{(1)} Q_2 = (V_m^{(1)} Q_2)(R_2 Q_2) + v_{m+1}(b_{m+1}^{(1)})^T Q_2$$

$$AV_m^{(2)} = V_m^{(2)} H_m^{(2)} + v_{m+1}(b_{m+1}^{(2)})^T$$

Now:

1st column of $V_m^{(2)}$ = scalar $\times (A - \theta_2 I)v_1^{(1)}$
= scalar $\times (A - \theta_2 I)(A - \theta_1 I)v_1$

➤ Note that

$$(b_{m+1}^{(2)})^T = e_m^T Q_1 Q_2 = [0, 0, \cdots, 0, \eta_1, \eta_2, \eta_3]$$

➤ Let: $\hat{V}_{m-2} = [\hat{v}_1, \ldots, \hat{v}_{m-2}]$ consist of first $m-2$ columns of $V_m^{(2)}$ and $\hat{H}_{m-2} = H_m(1:m-2, 1:m-2)$. Then

$$A\hat{V}_{m-2} = \hat{V}_{m-2}\hat{H}_{m-2} + \hat{\beta}_{m-1}\hat{v}_{m-1}e_m^T \quad \text{with}$$
$$\hat{\beta}_{m-1}\hat{v}_{m-1} \equiv \eta_1 v_{m+1} + h_{m-1,m-2}^{(2)} v_{m-1}^{(2)} \quad \|\hat{v}_{m-1}\|_2 = 1$$

➤ Result: An Arnoldi process of $m-2$ steps with the initial vector $p(A)v_1$.

➤ In other words: We know how to apply polynomial 'filtering' via a form of the Arnoldi process, combined with the QR algorithm.

The Davidson approach

Goal: to use a more general preconditioner to introduce good new components to the subspace.

➤ Ideal new vector would be eigenvector itself!

➤ Next best thing: an approximation to $(A - \mu I)^{-1} r$ where $r = (A - \mu I)z$, current residual.

➤ Approximation written in the form $M^{-1}r$. Note that M can vary at every step if needed.

ALGORITHM : 1 . **Davidson's method ($A = A^T$)**

1. *Choose an initial unit vector v_1. Set $V_1 = [v_1]$.*
2. *For $j = 1, \ldots, m$ Do:*
3. $w := Av_j$.
4. *Update $H_j \equiv V_j^T A V_j$*
5. *Compute the smallest eigenpair μ, y of H_j.*
6. $z := V_j y \qquad r := Az - \mu z$
7. *Test for convergence. If satisfied Return*
8. *Compute $t := M_j^{-1} r$*
9. *Compute $V_{j+1} := ORTHN([V_j, t])$*
10. *EndDo*

➤ Note: Traditional Davidson uses diagonal preconditioning: $M_j = D - \sigma_j I$.

➤ Will work only for some matrices

Other options:

➤ Shift-and-invert using ILU [negatives: expensive + hard to parallelize.]

➤ Filtering (by averaging)

➤ Filtering by using smoothers (multigrid style)

➤ Iterative solves [e.g., Jacobi-Davidson]

Jacobi-Davidson: Introduction via Newton's metod

Assumptions: $M = A + E$ and $Az \approx \mu z$

Goal: to find an improved eigenpair $(\mu + \eta, z + v)$.

➤ Write $A(z + v) = (\mu + \eta)(z + v)$ and neglect second order terms + rearrange ➤

$$(M - \mu I)v - \eta z = -r \quad \text{with} \quad r \equiv (A - \mu I)z$$

➤ Unknowns: η and v.

➤ Underdertermined system. Need one constraint.

➤ Add the condition: $w^H v = 0$ for some vector w.

In matrix form:

$$\begin{bmatrix} M - \mu I & -z \\ w^H & 0 \end{bmatrix} \begin{bmatrix} v \\ \eta \end{bmatrix} = \begin{bmatrix} -r \\ 0 \end{bmatrix}$$

➤ Eliminate v from second equation:

$$\begin{aligned} (M - \mu I)v - \eta z &= -r \\ w^H (M - \mu I)^{-1} z . \eta &= w^H (M - \mu I)^{-1} r \end{aligned}$$

➤ Solution: [Olsen's method]

$$\eta = \frac{w^H (M - \mu I)^{-1} r}{w^H (M - \mu I)^{-1} z} \qquad v = -(M - \mu I)^{-1}(r - \eta z)$$

When $M = A$, corresponds to Newton's method for solving

$$\begin{cases} (A - \lambda I)u = 0 \\ w^T u = Constant \end{cases}$$

Another characterization of the solution:

$$v = -(M - \mu I)^{-1} r + \eta (M - \mu I)^{-1} z,$$
$$\eta \text{ such that } \quad w^H v = 0$$

Alternative expression using projectors.

➤ Let P_z= projector in direction of z, s.t. $P_z r = r$:

$$P_z = I - \frac{z s^H}{s^H z} \quad \text{with} \quad s \perp r$$

➤ Similarly let P_w any projector that leaves v inchanged. Then Olsen's solution can be rwritten in mathematically equivalent form:

$$[P_z (M - \mu I) P_w] v = -r \qquad w^H v = 0$$

The Jacobi-Davidson approach

➤ In orthogonal projection methods (e.g. Arnoldi) we have $r \perp z$

➤ Also it is natural to take $w \equiv z$. Assume $\|z\|_2 = 1$

With the above assumptions, Olsen's correction equation is mathematically equivalent to finding v such that :

$$(I - zz^H)(M - \mu I)(I - zz^H)v = -r \qquad v \perp z$$

➤ Main attraction: can use iterative method for the solution of the correction equation. (M -solves not explicitly required).

Harmonic Ritz values

Main idea: take $L = AK$ in projection process

➤ In context of Arnoldi's method. Write $\tilde{u} = V_m y$ then:

$$(A - \tilde{\lambda} I)V_m y \perp \{AV_m\}$$

Using $AV_m = V_{m+1}\underline{H}_m$ ➤

$$\underline{H}_m^H V_{m+1}^H \left[V_{m+1}\underline{H}_m y - \tilde{\lambda} V_m y\right] = 0$$

Notation: $H_m = \underline{H}_m -$ last row. Then

$$\underline{H}_m^H \underline{H}_m y - \tilde{\lambda} H_m^H y = 0$$

or

$$\left(H_m^H H_m + h_{m+1,m}^2 e_m e_m^H\right) y = \tilde{\lambda} H_m^H y$$

Remark:

Assume H_m is nonsingular and multiply both sides by H_m^{-H}. Then, the problem is equivalent to

$$\left(H_m + z_m e_m^H\right) y = \tilde{\lambda} y$$

with $z_m = h_{m+1,m}^2 H_m^{-H} e_m$.

➤ Modified from H_m only in the last column.

Implementation within Davidson framework

➤ Slight varation to standard Davidson: Introduce $z_i = M_i^{-1} r_i$ to subspace. Proceed as in FGMRES: $v_{j+1} = Orthn(Az_j, V_j)$.

➤ From Gram-Schmidt process: $Az_j = \sum_{i=1}^{j+1} h_{ij} v_i$

➤ Hence the relation

$$AZ_m = V_{m+1}\bar{H}_m$$

Approximation: λ, $\tilde{u} = Z_m y$

Galerkin Condition: $r \perp AZ_m$ gives the generalized problem

$$\bar{H}_m^H \bar{H}_m \, y = \lambda \, \bar{H}_m^H V_{m+1}^H Z_m \, y$$