Arnoldi and Implicitly Restarted Arnoldi

See book.

Arnoldi recurrence : given u_1 $(\|u_1\|_2 = 1)$

for $j = 1..m$
  $\tilde{u} = A u_j$
  for $i = 1..j$
    $h_{ij} = u_i^H \tilde{u}$; $\tilde{u} = \tilde{u} - u_i h_{ij}$
  end
  $h_{j+1\,j} = \|\tilde{u}\|_2$; $u_{j+1} = \tilde{u} / h_{j+1\,j}$ *1
end

*1 if $h_{j+1\,j} < \eta \cdot \|\tilde{u}\|_2$ repeat the orthogonalization (add old + new h_{ij})

*2 we assume here $h_{j+1\,j} = 0$ (exactly) does not occur.

In matrix form : $A U_m = U_{m+1} \hat{H}_m^{(m+1)\times m}$

by construction $\hat{H} \to$ Hessenberg (upper H.) matrix

$$\begin{pmatrix} x & x & & \\ x & x & & \\ & \ddots & \ddots & \\ & & x & x \\ 0 & \cdots & 0 & x \end{pmatrix}$$

$$A U_m = U_m H_m^{m\times m} + h_{m+1,m} u_{m+1} e_m^T$$

→ Arnoldi decomposition

by construction $K_m(A, u_1) = \mathrm{span}(U_m)$
$K_{m+1}(A, u_1) = \mathrm{span}(U_{m+1})$

also $U_{m+1}^H U_{m+1} = I_{m+1}$

H_m (or $\hat{H}_m$) is reduced if $h_{j+1\,j} = 0$ for some $j = 1..m$

If H_m ($\hat{H}_m$) reduced it is in block Schur form:

$$H_m = \begin{array}{cc} j & m-j \\ \left(\begin{array}{c} H_{11} \\ \begin{matrix} 0 & \cdots & 0 & h \\ 0 & & 0 & 0 \\ \vdots & & \vdots & \vdots \\ 0 & \cdots & 0 & 0 \end{matrix} \end{array}\right. & \left.\begin{array}{c} H_{12} \\ \hline H_{22} \end{array}\right) \end{array} \quad \text{reduced if } h = 0$$

~~So, eigenpairs~~ So, eigenpairs of H_{11} are eigenpairs of H_m and

$$A(\underbrace{U_1}_{j}\ \underbrace{U_2}_{m-j}) = (U_1\ U_2)\begin{pmatrix} H_{11} & H_{12} \\ 0 & H_{22} \end{pmatrix} \Rightarrow$$

$$AU_1 = U_1 H_{11} \rightarrow$$

So, Ritz pair (H_{11}, U_1) is (exact) eigenpair.

Unfortunately ← this never ~~seems~~ to happens exactly (to me) due to rounding error.

So, "breakdown" of Arnoldi recurrence indicates we found exact invariant subspace. We can continue by picking u_{j+1} ($\|u_{j+1}\|_2 = 1$) orthogonal to U_1.

~~Which~~ We see that a reduced Hessenberg matrix corresponds to ~~an~~ an ~~an~~ eigenspace.

The ~~reverse~~ converse is true as well.

If $R(U_m)$ contains an eigenspace H_m is reduced.

If $R(U_m)$ contains eigenspace, it contains at least 1 eigenvector $U_m w$

If $AU_m w = \lambda w$, then

$$(A-\lambda I)U_m w = U_{m+1}\hat{H}_m w - U_m w \lambda$$

$$= U_{m+1}\left(\hat{H}_m - \lambda \begin{pmatrix} I \\ 0\cdots 0 \end{pmatrix}\right) w = 0 \quad \Rightarrow$$

$$\begin{pmatrix} H_m - \lambda I \\ 0\cdots 0 \; h_{m+1,m} \end{pmatrix} w = 0$$

If $\hat{H}_m$ is unreduced then $\hat{H}_m - \lambda \hat{I}$ is unreduced and hence its columns are independent. $\rightarrow$ $w=0$ which is a contradiction. So, $\hat{H}_m$ cannot be unreduced.

From $U_{m+1}^H U_{m+1} = I$ and

$$AU_m = U_{m+1}\hat{H}_m$$ we have

$$U_{m+1}^H A U_m = \hat{H}_m \;\&\; U_m^H A U_m = H_m$$

So, H_m is Rayleigh Quotient.

~~This makes Arnoldi decomposition and Arnoldi recurrence (process) particularly useful~~

This makes Arnoldi recurrence (process) and Arnoldi decomposition particularly useful for eigenvalue/pair computations.

~~[illegible] be Ritz~~

Computing Ar. decomp. is main cost – see notes book. ←

Ritz pairs can be computed cheaply from Arnoldi decomposition.

Refined Ritz vectors and harmonic Ritz pairs can also be computed cheaply.

Let $\mu_1 \cdots \mu_m$ be Ritz values.

Refined Ritz vectors are computed as follows ($AU_m = U_{m+1}\hat{H}_m$) :

$$\hat{x}_\mu = \arg\min_z \|(A-\mu I)U_m z\|_2$$

$$\hat{I}_m = \begin{pmatrix} I_m \\ 0 \cdots 0 \end{pmatrix}$$

$$(A-\mu I)U_m z = (U_{m+1}\hat{H}_m - \mu U_m) z$$

$$= U_{m+1}(\hat{H}_m - \mu \hat{I})$$

Since $U_{m+1}^H U_{m+1} = I_{m+1}$

$$\|U_{m+1}(\hat{H}_m - \mu\hat{I}) z\|_2 = \|(\hat{H}_m - \mu\hat{I}) z\|_2$$

Hence, we only need to compute the SVD for a small $(m+1)\times m$ matrix.

Harmonic Ritz pairs are computed as follows:

$(\varkappa+\delta, U w)$ is harmonic Ritz pair of

$$\underbrace{U_m^H (A-\varkappa I)^H}\underbrace{(A-\varkappa I) U_m} w = \delta U_m^H (A-\varkappa I)^H U_m w$$

$$(A-\varkappa I) U_m = U_{m+1}(\hat{H}_m - \varkappa \hat{I}) \longrightarrow$$

$$(\hat{H}_m - \varkappa \hat{I})^H (\hat{H}_m - \varkappa \hat{I}) w = \delta (\hat{H}_m - \varkappa \hat{I})^H \hat{I}_m w$$

$$= \delta H_m^H w$$

This amounts to solving a small generalized eigenvalue problem ~~…~~ defined by H_m (already computed).

Neither Ref. Ritz vectors or harm. Ritz pairs involve $O(n)$ computations.

We cannot continue Ar. recurrence until wanted eig. pairs are found suff. accurately. Too expensive.

Build Arnoldi decomposition and "prune" (truncate) subspace. ~~…~~ (remove components that do not contribute)

Keep space of fixed dim. and reduce angle w. eig. space of interest.

→ Do in such a way that we maintain Arnoldi decomp. that we can extend again.

Key concept is that Krylov space is space of polynomials.

$K_m(A, u_1) = \text{span}\{u_1, Au_1, \dots, A^{m-1}u_1\}$

$= \{p_{m-1}(A)u_1 : p_{m-1} \text{ poly of degree at most } m-1\}$

$$p_{m-1}(A)u_1 = \sum_{i=0}^{m-1} \alpha_i A^i u_1$$

$$= \sum_{i=0}^{m-1} \alpha_i A^i U_m e_1$$

$$A U_m e_1 = U_{m+1} \hat{H}_m e_1 = U_m H_m e_1$$

(induction) $A^k U_m e_1 = U_m H_m^k e_1 \quad (k < m)$

$\hookrightarrow A \cdot (A^{k-1} U_m e_1) = A U_m \underbrace{H_m^{k-1} e_1}_{\begin{pmatrix} x \\ \vdots \\ x \\ 0 \\ \vdots \\ 0 \end{pmatrix}]k}$

$= U_{m+1} \hat{H}_m H_m^{k-1} e_1$

$= U_m (H_m^k e_1)$

$$p_{m-1}(A) u_1 = U_m \underbrace{p_{m-1}(H_m) e_1}$$

* So, we can work with poly in H_m, Kr. space in H_m, etc.

We want to keep Ar. decomp. of order k.

$\to$ Krylov space of dim k in new starting vector; subspace of current space.

Take $\tilde{u}_1 = p^{(1)}(A) u_1 = U_m p^{(1)}(H_m) e_1$

let $p^{(1)}$ be of degree $m-k \to p^{(1)}_{m-k}$.

Then $p_j(A) p^{(1)}_{m-k}(A) u_1 \in R(U_m)$ if $j \le k-1$

Good $\tilde{u}_1$ $\rightarrow$ large components (relative) in wanted eig. space; small components in complement.

Use $p^{(1)}_{m-k}(A)$ to dampen comp.s in unwanted eig. pairs.

$u = \sum_i \gamma_i v_i \qquad (A v_i = d_i v_i)$

$p(A)\, u = \sum_i \gamma_i\, p(d_i)\, v_i \rightarrow p(d_i)$ very small removes component in direction of eig. vector v_i.

Take $p^{(1)}_{m-k}$ to be filter polynomial.

Implicitly (cheaply) generate Arnoldi decomp. for new starting vector.

$$A U_m = U_{m+1} \hat{H}_m \longrightarrow A \tilde{U}_k = \tilde{U}_{k+1} \hat{\tilde{H}}_k$$

dampen unwanted part of spectrum using zeros of polynomial.

$$p(t) = (t-\kappa_1)(t-\kappa_2)\cdots(t-\kappa_{m-k})$$

Apply using shifted QR step:

$$(A-\kappa_1 I)\, U_m = U_m H_m - \kappa_1 U_m + \beta_m u_{m+1} e_m^T$$

Compute $Q_1 R_1 = H_m - \kappa_1 I \rightarrow$

$$(A-\kappa_1 I)\underbrace{U_m Q_1}_{U_m^{(1)}} = \underbrace{U_m Q_1}_{U_m^{(1)}} \cdot R_1 Q_1 + \beta_m u_{m+1} e_m^T Q_1$$

$$A U_m^{(1)} = U_m^{(1)} \underbrace{\overbrace{(R_1 Q_1 + \kappa_1 I)}^{\to\, H_m^{(1)}}}_{\text{implicit QR step (w. shift) on Hessenberg matrix}} + \beta_m u_{m+1} b_{m+1}^{(1)H}$$

$\rightarrow U_m^{(1)}$ still orthonormal

$\rightarrow H_m^{(1)}$ is upper Hessenberg $\rightarrow$ QR preserves H. form in Q_i

($\cancel{p}Q_i$ is product of Givens rotations) $\downarrow$

obvious from structure

$$\rightarrow b_{m+1}^{(1)H} = e_m^T Q_1 = (0 \cdots 0 * *)$$

$\rightarrow u_1^{(1)} = U_m^{(1)} e_1$ given by

$$(A - \varkappa_1 I)\, u_1 = U_m (H_m - \varkappa_1 I) e_1 + \beta_m u_{m+1} \underbrace{\cancel{e_m^T e_1}}_{=0}$$

$$= \underbrace{U_m Q_1} R_1 e_1$$

$$= \cancel{U_m} U_m^{(1)} e_1 r_{11} \quad \Longleftrightarrow$$

$$r_{11}^{-1}(A - \varkappa_1 I)\, u_1 = u_1^{(1)} \Rightarrow u_1^{(1)} = p^{(1)}(A)\, u_1$$

(analogous) $u_1^{(2)} = c_2 (A - \varkappa_2 I)\, u_1^{(1)} = p^{(2)}(A)\, u_1$

After m-k shifts (implicit QR steps on H_m):

$$A U_m^{(m-k)} = U_m^{(m-k)} H_m^{(m-k)} + \beta_m u_{m+1} b_{m+1}^{(m-k)H}$$

$\uparrow$ orthonorm. $\qquad \uparrow$ upperhess

$u_1^{(m-k)} = p(A)\, u_1$

$$b_{m+1}^{(m-k)} = e_m^T Q_1 Q_2 \cdots Q_{m-k}$$

$$= (\underbrace{0 \cdots 0}_{k-1} \underbrace{* \cdots *}_{m-k+1})$$

Take first k columns of $U_m^{(m-k)}$, leading $(k+1)\times k$ part of $H_m^{(m-k)}$ and rewrite

$$A\tilde{U}_k = \tilde{U}_k \tilde{H}_k + h^{(m-k)}_{k+1,k}\, u^{(m-k)}_{k+1}\, e_k^T + \beta_m q_{mk}\, u_{m+1}\, e_k^T$$

where $q_{mk} = (b^{(m-k)}_{m+1})_k$

$$A\tilde{U}_k = \tilde{U}_k \tilde{H}_k + \tilde{\beta}_k\, \tilde{u}_{k+1}\, e_k^T$$

$$\tilde{\beta}_k = \| h^{(m-k)}_{k+1,k}\, u^{(m-k)}_{k+1}\, e_k^T + \beta_m q_{mk}\, u_{m+1}\, e_k^T \|_2$$

$$\tilde{u}_{k+1} = \tilde{\beta}_k^{-1}\left(h^{(m-k)}_{k+1,k}\, u_{k+1}\, e_k^T + \beta_m q_{mk}\, u_{m+1}\, e_k^T \right)$$

One obvious (usual) strategy: use unwanted Ritz values as shifts

(For H_m these are exact shifts $\Leftrightarrow$ eigenvalues)