Newton-based Methods

let $Az - \mu z = r \quad \|z\| = 1, \ \|r\|$ small, $\mu = z^H A z$

Assume z not ill-conditioned, μ not ill-conditioned
So, nearby eigenpair:

$$A(z+v) = (z+v)(\mu+\eta) \qquad \text{where } \|v\| = \varepsilon$$

$$\Leftrightarrow \ Az - \mu z + Av - \mu v - \eta z - \cancel{\eta v} = 0 \quad \Leftrightarrow$$

$$(A-\mu I) v \simeq -r + \eta z$$ ($\simeq$ because quadr. small terms ignored)

(or $(A - \mu I)\tilde{v} = -r + \eta z$)

If we assume correction v orthogonal to z $\rightarrow$

$$(I - zz^*)(A-\mu I) v = -(I - zz^*) r + \cancel{\eta (I - zz^*) z}$$

$$\left(\mu = z^H A z \Rightarrow z^H r = z^H A z - \mu = 0\right)$$

$$(I - zz^H)(A - \mu I)\overbrace{(I - zz^H) v}^{= v} = -r$$

Correction equation $\rightarrow$ consistent since $r \perp z$

$$(A - \mu I) v = -r + \eta z \Leftrightarrow v = -(A-\mu I)^{-1} r + \eta (A-\mu I)^{-1} z$$

$$v \perp z \Rightarrow z^H v = -z^H (A-\mu I)^{-1} r + \eta z^H (A-\mu I)^{-1} z = 0$$

$$\Rightarrow \eta = \frac{z^H (A-\mu I)^{-1} r}{z^H (A-\mu I)^{-1} z} = \frac{z^H z}{z^H (A-\mu I)^{-1} z} = \frac{1}{z^H (\dots)^{-1} z}$$

$$\left((A - \mu I) z = r \Leftrightarrow z = (A-\mu I)^{-1} r\right)$$

$$v = -z + \frac{1}{z^H(A-\mu I)^{-1}z}(A-\mu I)^{-1}z$$

$$\hat{z} = z+v = \eta(A-\mu I)^{-1}z$$

(Note $\hat{z}$ not exact because of neglected quadratic term)

New residual: $A\hat{z} - (\mu+\eta)\hat{z} = \hat{r} \iff$

$$(A-\mu I)\hat{z} - \eta\hat{z} = \hat{r} \iff$$

$$\eta \cancel{(A-\mu I)(A-\mu I)^{-1}} z - \eta^2(A-\mu I)^{-1}z = \hat{r} \iff$$

$$\eta(z - \eta(A-\mu I)^{-1}z) = -\eta v = \hat{r}$$

$|\eta|$? Take exact $v \to A(z+v) = (z+v)(\mu+\eta)$

Then $z^H A v = z^H A(z+v) - z^H A z =$

$$\underbrace{\cancel{z^H(z+v)}}_{1}(\mu+\eta) - \mu = \eta$$

$$|\eta| = |z^H A v| \le \|z\|_2 \cdot \|Av\|_2 \le \alpha\varepsilon \quad (\alpha = \|A\|_2)$$

$$\left.\begin{array}{l} \text{exact } v \to (A-\mu I)v = -r + \eta z + \eta v \\ \text{approx } \tilde{v} \to (A-\mu I)\tilde{v} = -r + \eta z \end{array}\right\} \Rightarrow$$

$$(A-\mu I)(v-\tilde{v}) = \eta v \quad (O(\varepsilon^2))$$

$$\|v-\tilde{v}\| = \|\Delta v\| = |\eta| \, \|(A-\mu I)^{-1}v\|_2$$

Since $\{\eta+\mu, z+v\}$ not ill-cond $\to$

$\|\Delta v\| = O(\varepsilon^2)$ typically.

If "other" eig. value too close to μ $\|\Delta v\| = O(\varepsilon)$

$\|(z+v)-(z+\tilde{v})\|_2 = \|Av\|_2$

We also have $\|\hat{r}\| = |\eta|\,\|v\|_2 = O(\varepsilon^2)$

If A Hermitian we get cubic (!) convergence.

In practice (and in principle) we should not use $\mu+\eta$ but new Rayleigh quotient:

$$\hat{z}^H A \hat{z} = (z+\tilde{v})^H A (z+\tilde{v})$$

(also for residual). More detailed analysis pp. 401–408

In Jacobi-Davidson method we combine "Newton correction" with Rayleigh-Ritz approx. over subspace. In practice correction equation solved approximately.

JD step $V = [v_1 \cdots v_m] \rightarrow H_m = V_m^H A V_m$

(focal point / target τ)

~~eigen~~ Ritz pair $(\mu, z) \rightarrow r = Az - \mu z$

Choose shift κ near τ or μ

Solve $(I - zz^H)(A - \kappa I)(I - zz^H)v = -r \quad (v \perp z)$

↳ typically (prec) Krylov method

$$v_{m+1} = (v - V_m V_m^H v)/\|v - V_m V_m^H v\|_2$$

$$V_{m+1} = (V_m \;\; v_{m+1}) \rightarrow H_{m+1} = \begin{pmatrix} H_m & V_m^H A v_{m+1} \\ v_{m+1}^H A V_m & v_{m+1}^H A v_{m+1} \end{pmatrix}$$

How to solve for multiple eig. pairs $\rightarrow$

Use (partial) Schur decomposition.

$AU = UT \rightarrow$ already found

$U^H U = I$, $\quad U^H A U = T$ upper triangular

Find next Schur vector:

$(U\ U_\perp)$ unitary $\rightarrow (U\ U_\perp)^H A (U\ U_\perp) = \begin{pmatrix} T & H \\ 0 & B \end{pmatrix}$

$Bz = \mu z \quad (\|z\|_2 = 1) \quad (z\ Z_\perp)$ unitary

$$\begin{pmatrix} z^H \\ Z_\perp^H \end{pmatrix} B \,(z\ Z_\perp) = \begin{pmatrix} \mu & g^H \\ 0 & C \end{pmatrix} \Rightarrow$$

$$(U\ U_\perp z\ U_\perp Z_\perp)^H A\, (U\ U_\perp z\ U_\perp Z_\perp) =$$

$$\begin{pmatrix} T & HU_\perp z & HU_\perp Z_\perp \\ 0 & \mu & g^h \\ 0 & 0 & C \end{pmatrix} \rightarrow$$

$$A(U\ U_\perp z) = (U\ U_\perp z)\begin{pmatrix} T & HU_\perp z \\ 0 & \mu \end{pmatrix}$$

part. Schur decomp. of A of one order higher.

Implementation?

$P_\perp = U_\perp U_\perp^H = (I - UU^H) \equiv I - P$

$A_\perp = P_\perp A P_\perp = U_\perp B U_\perp^H$

If $Bz = \mu z \rightarrow y = U_\perp z \rightarrow$

$A_\perp y = U_\perp B U_\perp^H U_\perp z = \mu U_\perp z = \mu y$

(obviously $y \perp U$)

Conversely, let $y \perp U$ and $A_\perp y = \mu y$

Then $U_\perp U_\perp^H y = U_\perp z = y$ and

$$A_\perp y = U_\perp B z = \mu U_\perp z \Rightarrow \underset{(\text{mult } U_\perp^H)}{Bz = \mu z}$$

So, we can compute (μ, y) using JD for $A_\perp$, and makes $A_\perp = (I - UU^H) A (I - UU^H)$ does not require anything beyond what we have.

Once we have $A_\perp y = \mu y$ we set

$$A(U \; y) = (U \; y)\begin{pmatrix} T & t \\ 0 & \mu \end{pmatrix} \quad \text{where } t = U^H A y$$

(Since search space $V \perp U$ we can compute $V^H A V$ rather than $V^H (I - UU^H) A (I - UU^H) V$

For Ritz pair: $r = A_\perp z - \mu z$

Want $s = Az - Ut - \mu z$ sufficiently small

$$U^H s = 0 \qquad U_\perp U_\perp^H s = U_\perp U_\perp^H A z - \mu U_\perp U_\perp^H z$$

$$\cancel{U_\perp U_\perp^H A U_\perp} \qquad\qquad z \perp U$$

$$= P_\perp A P_\perp z - \mu z = r$$

$\Rightarrow \|s\|_2 = \|r\|_2 \rightarrow$ use r in conv. test.

Solve correction equation:

$$(I - zz^H) A_\perp (I - zz^H) v = r$$

Truncation typically done by computing (ordered) Schur decomp and selecting ~~???~~ eigenvalues closest to target.