Superlinear Convergence of CG

Def. Ritz pair :
Let V be a subspace of $\mathbb{C}^n$ and $A \in \mathbb{C}^{n \times n}$. $w \in V$ and $\theta \in \mathbb{C}$ are a Ritz pair of A wrt. V if

$$Aw - \theta w \perp V$$

Note that if V is invariant under A
$A : V \to V$ (or $\forall w \in V : Aw \in V$),
then (w, θ) is an eigenpair of A

Consider m steps of Lanczos iteration

$$AV_m = V_m T_m + t_{m+1,m} v_{m+1} \eta_m^T$$

$v_1 = r_0/\|r_0\|$ (as for CG), $V_{m+1}^* V_{m+1} = I_{m+1}$
η_i is ith Cartesian basis vector
$\eta_1^T = (1\ 0 \cdots 0)$, $\eta_2^T = (0\ 1\ 0 \cdots 0)$, etc

Ritz pairs of A wrt $R(V_m)$: $w = V_m y$

$$AV_m y - V_m y \theta \perp \text{Range}(V_m) \to$$

$$V_m^* (AV_m y - V_m y \theta) = 0$$

$$V_m^*\left(V_m T_m + t_{m+1,m} v_{m+1} \eta_m^T\right) y - \underbrace{\cancel{V_m^* V_m}}_{I} y\theta = 0 \iff$$

$$T_m y - y\theta = 0$$

So, y, θ is eigenpair of T_m $\rightarrow$ eigenpairs y_i, θ_i of T_m give Ritz pairs

$(V_m y_i, \theta_i)$ of A (wrt Range(V_m))

Assume A HPD, solving $Ax = b$
Given $x_0 \rightarrow r_0 = b - Ax_0$, $v_1 = r_0 / \|r_0\|_2$,
and Lanczos iteration gives

$$AV_m = V_m T_m + t_{m+1,m} v_{m+1} \eta_m^T$$

and $r_m \neq 0$ (not yet converged)
and hence $v_{m+1} \neq 0$

(it is easy to prove that all coefs of the tridiagonal matrix T_m are nonzero $\rightarrow$ good exercise)

Theorem: The eigenvalues θ_i of T_m (Ritz values of A w.r.t. Range(V_m)) are the roots of the CG polynomial R_m ($r_m = R_m(A) r_0$)

Proof:

$$A v_1 = A V_m e_1 = \left(V_m T_m + t_{m+1,m} v_{m+1} e_m^T\right) e_1$$
$$= V_m T_m e_1 \qquad (\text{as } e_m^T e_1 = 0)$$

$$A^k v_1 = A\left(A^{k-1} V_m e_1\right)$$

Assume for $j = 1 \ldots k-1$:

$A^j V_m e_1 = V_m T_m^j e_1$ (check afterwards)

$$A^k v_1 = A(A^{k-1} V_m e_1) = A(V_m T_m^{k-1} e_1)$$
$$= \left(V_m T_m + t_{m+1,m} v_{m+1} e_m^T\right) T_m^{k-1} e_1$$

$$= V_m T_m^k e_1 + t_{m+1,m} v_{m+1} e_m^T T_m^{k-1} e_1$$

For T_m unreduced tridiagonal matrix (no zeros on upper/lower diag) and no zeros on diag as A HPD.:

$(T_m e_1)_i = 0$ for $i = 3, 4, 5, \cdots \Rightarrow$

$(T_m^2 e_1)_i = 0$ for $i = 4, 5, 6, \cdots$ etc.

$$\left(T_m^{m-2} e_1\right)_m = 0$$

$$\left(T_m^{m-1} e_1\right)_m \neq 0 \rightarrow e_m^T T_m^{m-1} e_1 \neq 0$$

So, $A^k v_1 = V_m T_m^k e_1$ for $k = 1 .. m-1$

$$A^m v_1 = A V_m T_m^{m-1} e_1 =$$

$$\left(V_m T_m + t_{m+1,m} v_{m+1} e_m^T\right) T_m^{m-1} e_1 =$$

$$V_m T_m^m e_1 + t_{m+1,m} v_{m+1} e_m^T T_m^{m-1} e_1 =$$

$$V_m T_m^m e_1 + v_{m+1} \gamma \quad (\gamma = t_{m+1,m} e_m^T T_m^{m-1} e_1)$$

$\gamma \neq 0$

$r_m = R_m(A) r_0 \perp V_m$ and $r_m \neq 0$

(by assumption)

$r_m \neq 0 \Rightarrow$ leading coeff. $\pi_m \neq 0$

$$R_m(A) r_0 = R_m(A) v_1 e_1 \|r_0\| =$$

$$V_m R_m(T_m) e_1 \|r_0\| + \pi_m \gamma v_{m+1} \|r_0\| \perp V_m$$

Hence $R_m(T_m) e_1 = 0 = R_m$ singular

let $T_m = U_m \Theta_m U_m^*$ (eig. decomp.)

$$e_1 = U_m \xi = \sum_{i=1}^{m} u_i \xi_i$$

$$\xi = U_m^* e_1$$

$R_m(T_m) e_1 = U_m R_m(\Theta_m) U_m^* e_1 =$

$$\sum_{i=1}^{m} u_i R_m(\theta_i) \zeta_i = 0 \Rightarrow$$

$R_m(\theta_i) = 0$ or $\zeta_i = 0$ for $i = 1..m$

(the u_i are independent)

Show $\zeta_i \neq 0$ for $i = 1..m$

$\zeta_i = u_i^* \eta_1$. Assume $\zeta_i = 0 \Rightarrow$
$u_{1,i} = 0$

Now consider $T_m u_i = u_i \theta_i \Rightarrow$
$e_1^* T_m u_i = e_1^* u_i \theta_i = 0$

Then $u_i^* T_m e_1 = 0 \Rightarrow u_i^* (t_{11} e_1 + t_{21} e_2) =$

$t_{11} \underbrace{u_i^* e_1}_{=0} + t_{21} u_i^* e_2 = 0$ (and $t_{21} \neq 0$)

But then $u_i^* e_2 = 0 \Rightarrow u_{2,i} = 0$

Next consider $e_2 T_m u_i = u_{2,i} \theta_i = 0 \Rightarrow$
$u_i^* T_m e_2 = u_i^* (t_{12} e_1 + e_2 t_{22} + e_3 t_{32}) = 0$

$\Leftrightarrow t_{12} \underbrace{u_i^* e_1}_{=0} + t_{22} \underbrace{u_i^* e_2}_{=0} + t_{32} u_i^* e_3 = 0$

$(t_{32} \neq 0)$ So, $u_{3,i} = 0$ etc

So, u_i is zero vector. Contradiction

Hence $\zeta_i \neq 0$ for $i = 1..m$

$\zeta_i \neq 0$ for $i = 1..m \Rightarrow R_m(\theta_i) = 0$ for $i = 1..m$

In addition $R_m(0) = 1$ (residual poly.)

$$R_m(t) = \frac{\theta_1 - t}{\theta_1} \, \frac{\theta_2 - t}{\theta_2} \cdots \frac{\theta_m - t}{\theta_m} \qquad \blacksquare$$

Let $A = Z \Delta Z^*$ with $Z^* Z = I$
and $0 < d_1 \leq d_2 \leq \cdots \leq d_n$

Assume at some point $\theta_1 \simeq d_1$ $(d_1 \leq \theta_1)$
$\hat{x}$ exact solution of $Ax = b$
x_m approximate solution at step m
$e_m = R_m(A) e_0$ error at step m

let $e_0 = Z\mu = \sum_{i=1}^{m} z_i \mu_i$ $\quad (\mu = Z^* e_0)$

$$e_m = \sum_{i=1}^{m} z_i R_m(d_i) \mu_i =$$

$$= \underbrace{z_1 R_m(d_1) \mu_1}_{\simeq 0} + \sum_{i=2}^{m} z_i R_m(d_i) \mu_i$$

$$= \hat{x} - x_m$$

Consider new starting vector $\tilde{x}_0$ such that
$$\hat{x} - \tilde{x}_0 = \tilde{e}_0 = \sum_{i=2}^{m} z_i R_m(d_i) \mu_i$$

Note $\|\tilde{e}_0\|_A^2 \leq \|e_m\|_A^2$

Now, we start a separate (new) CG iteration for the purpose of comparison. Next, we compare $\|e_{m+k}\|_A$ (from our original CG iteration after another k iterations) and $\|\tilde{e}_k\|_A$ (from our comparison CG iteration after k iterations)

$e_{m+k} = R_{m+k}(A)\, e_0$ optimal in A-norm

Now consider the following polynomial

$\tilde{R}_k(A)\, Q_m(A)$ where

$\tilde{R}_k(A)$ is the residual poly for k steps of CG starting with $\tilde{x}_0 = \hat{x} - \tilde{e}_0$

$$Q_m(t) = \frac{\lambda_1 - t}{\lambda_1} \cdot \frac{\theta_2 - t}{\theta_2} \cdot \ldots \cdot \frac{\theta_m - t}{\theta_m}$$

Note $\tilde{R}_k(0) = 1$ (residual poly) ~~an~~ and $Q_m(0) = 1$ (by construction) $\Rightarrow$ $\tilde{R}_k(0)\, Q_m(0) = 1$

$$Q_m(t) = \frac{\lambda_1 - t}{\lambda_1} \cdot \frac{\theta_1}{\theta_1 - t} R_m(t) = \frac{\theta_1}{\lambda_1} \cdot \frac{\lambda_1 - t}{\theta_1 - t} \cdot R_m(t)$$

for $i = 2, \ldots, m$: $|Q_m(\lambda_i)| = \frac{\theta_1}{\lambda_1} \left| \frac{\lambda_1 - \lambda_i}{\theta_1 - \lambda_i} \right| |R_m(\lambda_i)|$

Let $F_m = \max\limits_{i=2..m} \frac{\Theta_1}{d_1} \left| \frac{d_1 - d_i}{\Theta_1 - d_i} \right|$

$|Q_m(d_i)| \leq F_m \, |R_m(d_i)|$

(note $Q_m(d_1) = 0$)

$$\|e_{m+k}\|_A^2 \leq \|\tilde{R}_k(A) Q_m(A) e_0\|_A^2 =$$

$$\sum_{i=2}^{n} d_i \tilde{R}_k^2(d_i) Q_m^2(d_i) |\mu_i|^2 \leq$$

$$\sum_{i=2}^{n} d_i \tilde{R}_k^2(d_i) F_m^2 R_m^2(d_i) |\mu_i|^2 =$$

$$F_m^2 \|\tilde{R}_k(A) \tilde{e}_0\|_A = F_m^2 \|\tilde{e}_k\|_A^2 \leq$$

$$F_m^2 \frac{\|e_m\|_A^2}{\|\tilde{e}_0\|_A^2} \|\tilde{e}_k\|_A^2 \qquad \left(\frac{\|e_m\|_A^2}{\|\tilde{e}_0\|_A^2} \geq 1 \right)$$

$\Longleftrightarrow$

$$\|e_{m+k}\|_A \leq F_m \frac{\|\tilde{e}_k\|_A}{\|\tilde{e}_0\|_A} \|e_m\|_A$$

$$\leq F_m \cdot 2 \left(\frac{\sqrt{d_n/d_2} - 1}{\sqrt{d_n/d_2} + 1} \right)^k \|e_m\|_A$$

faster rate
(spectrum w. d_1 removed)

Note that for $\Theta_1 < \frac{d_1 + d_2}{2}$, $F_m < 2$.