April 14, 2021

\[Ax = b\]

\[A = M - N\]
\[Mx_{k+1} = Nx_k + b\]
\[x_{k+1} = M^{-1}N x_k + M^{-1}b\]

\[e_{k+1} = R e_k = R^k e_0\]
\[e_{k+1} \approx p(R) e_k\]
\[p(R) = \max_j |\lambda_j(R)| \quad (< 1 \text{ for convergence})\]

\[x_k - x_{k+1} = e_k - e_{k+1} \approx (1 - p(R)) e_k\]
\[x_* = x_k - e_k \approx x_k = \frac{x_k - x_{k+1}}{1 - p(R)}\]

Can we do better by combining iterates? \[\frac{R + M^{-1}b}{x_k}\]

If \[e_k \to 0\], can't really do worse \[x_{k+1} \to x^*\]

Recall heavy ball: \[x_{k+1} = x_k - \alpha_h \nabla f_h + \beta_h (x_k - x_{k-1})\]
\( K_h = \text{span} \{ x_0, \ldots, x_k \} \) (subspace)

\[ x_{k+1} = \arg\min_x f(x) = \frac{1}{2} x^T A x - x^T b \]

s.t. \( x \in K_h \)

1. How good is this?  
2. How to solve?  
3. Efficiency? (growing \( k \))

\[ x_{k+1} = R x_k + M^{-1} b \]

\[ x_0 = M^{-1} b \Rightarrow x_1 = R x_0 + M^{-1} b = R M^{-1} b + M^{-1} b = (R + I) M^{-1} b \]

\[ x_2 = R x_1 + M^{-1} b = R (R + I) M^{-1} b + M^{-1} b = (R^2 + R + I) M^{-1} b \]

\[ x_{k+1} = \sum_{j=0}^{k+1} R^j M^{-1} b \quad (R^0 = I) \]

\[ x_{k+1} = \sum_{j=0}^{k+1} c_j R^j M^{-1} b = q(R) M^{-1} b \]

\[ x_{k+1} \in \text{span} \{ M^{-1} b, R M^{-1} b, \ldots, R^k M^{-1} b \} \]

Krylov subspace

\[ K_{k+2}(A, b) = \text{span} \{ b, Ab, \ldots, A^{k+1} b \} \]

\[ K_{k+2}(R, M^{-1} b) \]
How good is this?

\[ \min_{x} \frac{1}{2} x^T A x - x^T b \quad (A, b) \quad R = V \Lambda V^{-1} \]

s.t. \( x \in \mathbb{R}^n \in \text{span}\{M^{-1}b, RM^{-1}b, \ldots, R^{k-1}M^{-1}b\} \)

\[ e_k = x_k - x^* = \rho(R) M^{-1}b - x^* \]

\[ \rho(R) = c_0 I + c_1 R + c_2 R^2 + \cdots \]

\[ (\rho(R) = \sum_{j=0}^{k-1} c_j R^j) \quad M^{-1}b - x^* = e_0 \]

\[ \|e_k\|_2 = \|V \rho(\Lambda) V^{-1} e_0\| \leq k_2(V) \|e_0\|_2 \|\rho(\Lambda)\|_2 \]

\[ \min_{\rho \in \mathbb{R}^n} \|\rho(\Lambda)\|_2 \quad \max \{\rho(\Lambda)\} \]

\[ \min_{\rho \in \mathbb{R}^n} \max_{j} \{\rho(\Lambda)\} \]

Chebyshev \( T_0(z) = 1 \)
\[ T_1(z) = z \]
\[ T_{n+1}(z) = 2zT_n(z) - T_{n-1}(z) \]
How to solve?

\[ M = I, \quad N = I - A \Rightarrow x_0 = b, \quad R = I - A \]
\[ \text{span}\{b, (I-A)b, \ldots, (I-A)^{k-1}b\} = \text{span}\{b, Ab, \ldots, A^{k-1}b\} \]

\[ F_k = [b, Ab, \ldots, A^{k-1}b] \quad \text{n x h basis for } \ker(A, b) \]

\[
\begin{align*}
\min_{x \in \mathbb{R}^n} & \quad \frac{1}{2} x^T A x - x^T b \\
\text{s.t. } & \quad x = F_k c
\end{align*}
\]
\[
\Rightarrow \min_{c \in \mathbb{R}^n} \quad \frac{1}{2} c^T F_k^T A F_k c - c^T F_k b
\]
\[ F_k^T A F_k c = F_k^T b \]

Is \( F_k \) a good basis?

\[ A^{k-1}b \Rightarrow \text{first order} \]
\[ \Rightarrow \text{ill-conditioned} \]
$F_k = Q_k R_k \Rightarrow Q_k$ as ortho basis

$\text{Range}(AQ_k) = \text{Range}(A [b \ldots A^{k-1} b]) \leq \text{Range}( [b, Ab, \ldots, A^{k-1} b]) = \text{Range}(Q_{k+1})$

$Q_{k+1} = [q_1 \ldots q_{k+1}] \Rightarrow h_{n+1}, h_{q_{k+1}} = \frac{1}{2} \sum_{j=1}^{k} h_{j, k} q_j$

$$Q_h^T AQ_h = [Q_h^T Q_h] [H_k] = [Q_h^T H_{k+1} e_h^T] = \sum_{i=1}^{k+1} Q_h H_{h+1, k} + \sum_{i=1}^{k+1} h_{k+1, q_{k+1}} e_k$$

$Q_h^T A Q_h = H_k = (\text{symm}) \Rightarrow H_k = (0\ 0\ 0)$

$h_{n+1, q_{k+1}} + h_{k, k} q_k + h_{k-1, k} q_{k-1} = A_{kk}$