Nonstationary iterative methods and Chebyshev acceleration for $A$ SPD

Now we shall consider the nonstationary process

$$x^{(k+1)} = (I - \tau_k A) x^{(k)} + \tau_k b$$

In terms of the error $e^{(k)} = x^{(0)} - x$ we get

$$e^{(k)} = (I - \tau_k A) e^{(k-1)} = \ldots = (I - \tau_k A)(I - \tau_{k-1} A) \ldots (I - \tau_1 A) e^{(0)}$$

And therefore we have

$$\|e^{(k)}\| \leq \|Q_k(A)\| \|e^{(0)}\|$$

for any matrix norm that is subordinate to a vector norm $\|\cdot\|$.

Now we put the following question: Can we choose the iteration parameters $\tau_1, \tau_2, \ldots, \tau_k$ so that $\|Q_k(A)\|$ is as small as possible? This boils down to the construction of a matrix polynomial of such property and satisfying $Q(0) = I$.

In general, for an arbitrary nonsingular matrix $A$ this problem does not have exact mathematical solution.
However, if $A$ is an SPD matrix and $0 \neq b = N(A)$ then

$$
\|Q(A)\|_2 = \max_j |Q(j)| \leq \max_{\lambda \in \lambda = \Lambda} |Q(\lambda)|
$$

where $\lambda_j$ are eigenvalues of $A$

$$
\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n \leq \Lambda
$$

and the question translates to the following equivalent:

Find the polynomial of degree $\kappa$ that has minimal max-norm on the interval $(\lambda, \Lambda)$ and $Q(0) = 1$

In fact it is known that on $(-1, 1)$ the polynomial that has minimal max-norm is Chebyshev polynomial.

The following give equivalent definition of Chebyshev polynomials on $[-1, 1]$. 

\[ Q_n(x) = \cos(n \arccos(x)) \]

\[ T_n(x) = \frac{\sin((n+1) \arccos(x))}{\sin(\arccos(x))} \]
\[ T_m(t) = \cos(m \cos^{-1} t) \quad |t| \leq 1 \]

\[ T_m(t) = \frac{1}{2} \left( t + \sqrt{t^2 - 1} \right)^m + \frac{1}{2} \left( t - \sqrt{t^2 - 1} \right)^m \]

\[ T_0(t) = 1 \quad T_1(t) = t \quad T_m(t) = 2tT_{m-1}(t) - T_{m-2}(t) \quad m \geq 2 \]

All these satisfy \( T_m(0) = 1 \) and \( |T_m(t)| \leq 1 \) on \((-1, 1)\).

\[ Q_m(\mu) \] is obtained by transforming \(-1 \leq t \leq 1\) to \(-\mu \leq \zeta \leq \mu\)
so that \( t = \alpha + \beta \mu \) so that \(-1 = \alpha + \beta \mu \Rightarrow 1 = \nu + \beta \lambda\), i.e.
\[ \beta = \frac{2}{\lambda - \lambda} \quad \alpha = \frac{\lambda + \lambda}{\lambda - \lambda} \]
and normalizing \( \lambda = \frac{1}{\beta} = \frac{1}{\beta} \)

\[ Q_m(\mu) = \frac{T_m \left( \frac{2\mu}{\lambda - \lambda} - \frac{\lambda + \lambda}{\lambda - \lambda} \right)}{T_m \left( \frac{\lambda + \lambda}{\lambda - \lambda} \right)} \]

\[ \frac{\mu Q_m(\mu)}{2} = \left( \frac{T_m(\lambda + \lambda)}{T_m(\lambda - \lambda)} \right) \]
What are the plusses and minuses of this method?

**Positives:**
1. We have simple computations once we know $\xi$.
2. We shall show that we have much faster convergence than the standard methods: Richardson, Jacobi, G-S, SOR.

**Negatives:**
1. We need to know the bounds $\lambda$;
   - $\lambda = 1 < \ldots < 1_{n-1} = 1 \leq \overline{\lambda}$
   - $\lambda$ is easy to compute from Gershgorin circles; the difficult one is $\overline{\lambda}$.
2. Need to know in advance the number of iterations.

\[
\begin{vmatrix}
2 & -1 \\
-1 & 2 & -1 \\
& & & 0 \\
& & & -1 & 4 & -3 \\
& & & -3 & 6 & -3 \\
& & & 0 & & \\
& & & -3 & 6 & \\
\end{vmatrix}
\]

\[
\overline{\lambda} = 12 \quad \lambda = \rho = \frac{4 \sin^2 \frac{\pi}{2n}}{2n}
\]

\[
\begin{align*}
|z-2| & \leq 2 \\
|z-4| & \leq 4 \\
|z-6| & \leq 6 \\
-6 & \leq z-6 \leq 6 \\
0 & \leq z \leq 12
\end{align*}
\]
Thus the solution of our minimization problem is

\[ Q_m(M) = \frac{T_m \left( \frac{2^m}{N-\lambda} - \frac{N+1}{N-\lambda} \right)}{T_m \left( \frac{N+1}{N-\lambda} \right)} = \frac{N}{N+1} \]

(1)

\[ t = \frac{2^m}{N-\lambda} - \frac{N+1}{N-\lambda} \]

Recall \[ Q_m(A) = \prod_{j=1}^{m} (I - t_j A) \quad t_j \text{ - parameters} \]

\[ Q_m(M) = \prod_{j=1}^{m} (1 - t_j M) \quad t_j = \frac{1}{t_j} \text{ - roots of } Q_m(M) \]

Thus the roots of \( Q_m(M) \) and \( \frac{N}{N+1} \) are related through the transformation (1)

\[ 0 = T_m(t_j) = \cos m \cos^{-1} t_j \quad \Rightarrow m \omega_0^{-1} t_j = \frac{2j-1}{2} \]

\[ t_j = \cos \frac{2j-1}{2m} \pi \]

\[ j = 0, \ldots, m \]

\[ t_j = \frac{2m}{N-\lambda} - \frac{N+1}{N-\lambda} \quad \iff \quad \mu_j = \frac{1}{2} (N-\lambda) t_j + \frac{1}{2} (N+1) \frac{1}{t_j} \]

\[ j = 1, \ldots, m \]

\[ t_j = \frac{1}{\mu_j} = \frac{2}{N+1+(N-1)t_j} \quad t_j = \omega \frac{2j-1}{2m} \pi \]

\[ j = 1, \ldots, m \]
Is this a better method than the simplest ones?

\[-\frac{\lambda + \lambda}{\lambda - \lambda} = \frac{1}{50} \quad \rho_0 = \frac{1 - \xi}{1 + \xi} \quad \xi = \frac{\lambda}{\Lambda} < 1\]

\[T_n(H) = \frac{1}{2} \left(1 + \sqrt{1 + 4(\lambda + \lambda)} \right) \frac{1}{\lambda} \left(1 - \sqrt{1 + 4(\lambda + \lambda)} \right)\]

\[\|Q_n(A)\| = \frac{1}{\|T_n(1/50)\|} = \frac{2}{\left(\frac{1}{\rho_0} + 1 + \sqrt{\left(\frac{1}{\rho^2} - 1\right)^m + \left(\frac{1}{\rho_0} - \sqrt{\frac{4}{\rho^2} - 1}\right)^m}\right)}\]

\[\frac{1}{\rho_0} + \sqrt{\frac{4}{\rho^2} - 1} = \frac{1 + \xi}{1 - \xi} + \sqrt{\frac{(1 + \xi)^2}{(1 - \xi)^2} - 1} = \frac{1 + \xi + \sqrt{1 + 4\xi^2}}{1 - \xi} =\]

\[\frac{1 + \xi + \sqrt{1 + 4\xi^2}}{1 - \xi} = \frac{(1 + \sqrt{1 + 4\xi^2})^2}{(1 - \xi)(1 + \sqrt{1 + 4\xi^2})} =\]

\[= \frac{1 + \sqrt{1 + 4\xi^2}}{1 - \xi} = \frac{1}{\rho_0}\]

\[\frac{1}{\rho_0} - \sqrt{\frac{4}{\rho^2} - 1} = \frac{1 - \xi}{1 + \sqrt{1 + 4\xi^2}} = \rho_1\]

\[\frac{1}{\|T_n(1/50)\|} = \frac{2}{\rho_0^m + \rho_1^m} = \frac{2\rho_1^m}{1 + \rho_1^m} = 9^m \approx 2\rho_1^m\]

\[1 \sim \rho_1 < 1\]

\[\rho_1 \approx \frac{1 - \xi}{1 + \sqrt{1 + 4\xi^2}} \approx (1 - \sqrt{1 + 4\xi^2} = \approx 1 - 2\sqrt{\xi}\]

\[\|Q_n(A)\|_2 \approx 2 \left(1 - 2\sqrt{\xi}\right)^m\]

Richardson optimal is just \(m = 1\)

\[\|G_{\text{Rich}}(A)\| \approx \left(\frac{1 - \xi}{1 + \xi}\right)^m \approx \left(1 - 2\xi\right)^m\]
How to compute by Chebyshev acceleration method?

Possibility 1: As a nonstationary Richardson method.

Given \( \lambda \) and \( \Lambda \) (or more accurate \( \lambda = t_1, \Lambda = t_n \)),

1. tolerance \( \epsilon \) you want to reduce the initial error
2. compute \( m \) or fix \( m \) (say \( m = 64 \)).
3. compute the iteration parameters

\[ t_j = \frac{2}{\lambda + \lambda + (\Lambda - \lambda) t_j}, \quad t_j = \omega \cdot \frac{\pi (2j - 1)}{2m}, \quad j = 1, \ldots, m \]

and perform \( m \) steps of Richardson with \( t_j \).

4. If the tolerance is not achieved, repeat the process until you reach desired tolerance in your residual.

The negative of this is that you need to know in advance the number of iterations in order to determine the iteration parameters.
Our polynomial $Q_m(t)$ was expressed through Chebyshev polynomials — shifted and normalized:

$$p(t) = \frac{2t}{\Lambda - \lambda} - \frac{\Lambda + \lambda}{\Lambda - \lambda}$$

$$Q_m(t) = \frac{T_m(p(t))}{T_m(p(10))}$$

Chebyshev polynomials have some very nice properties. One of them as many other orthogonal polynomials are generated by proper three-term recurrence relations:

$$T_0(t) = 1 \quad T_1(t) = t \quad T_n(t) = 2tT_{n-1}(t) - T_{n-2}(t) \quad m \geq 2$$

This allows to reformulate the method using three-term recurrence relation:

$$\delta = \frac{2}{\Lambda + \lambda} \quad p_1 = 2 \quad p_k = \frac{1}{\lambda - \lambda_k} \quad \lambda = \frac{1}{4}(\frac{\Lambda - \lambda}{\Lambda + \lambda})^2$$

$$G = I - A$$

$x(0)$ given $x(1) = y(Gx(0) + b) + (1 - \delta)x(0)$

The same as one step Richardson

$$x^{k+1} = p_k \left\{ y(Gx^k + b) + (1 - \delta)x^k \right\} + (1 - p_k)x^{k-1}$$

1. No number of iterations needed
2. But you need to keep 3 iterates $x^{(k-1)}$, $x^{(k)}$, $x^{(k+1)}$
\[ T_2(t) = 2t^2 - 1 \quad t_{1,2} = \pm \frac{1}{\sqrt{2}} = \pm \frac{\sqrt{2}}{2} \]

\[ Q_2(M) = \left[ 2 \left( \frac{2M}{\Lambda - \lambda} - \frac{\Lambda + \lambda}{\Lambda - \lambda} \right)^2 - 1 \right] \sqrt{\frac{2}{\left( \frac{2M}{\Lambda - \lambda} - \frac{\Lambda + \lambda}{\Lambda - \lambda} \right)^2 - 1}} \]

\[ \left[ \sqrt{2} \left( \frac{2M}{\Lambda - \lambda} - \frac{\Lambda + \lambda}{\Lambda - \lambda} \right) - 1 \right] \left[ \sqrt{2} \left( \frac{2M}{\Lambda - \lambda} - \frac{\Lambda + \lambda}{\Lambda - \lambda} \right) + 1 \right] \]

\[ \left( \frac{2\sqrt{2} M}{\Lambda - \lambda} - \sqrt{2} \frac{\Lambda + \lambda}{\Lambda - \lambda} - 1 \right) \left( \frac{2\sqrt{2} M}{\Lambda - \lambda} - \sqrt{2} \frac{\Lambda + \lambda}{\Lambda - \lambda} + 1 \right) \]

\[ \frac{2\sqrt{2}}{\Lambda - \lambda} \left[ M - \frac{\Lambda + \lambda}{2} - \frac{\Lambda - \lambda}{2\sqrt{2}} \right] \quad t_{1,2} = \pm \frac{1}{\sqrt{2}} \]

\[ \times \frac{2\sqrt{2}}{\Lambda - \lambda} \left[ M - \frac{\Lambda + \lambda}{2} + \frac{\Lambda - \lambda}{2\sqrt{2}} \right] \]

\[ (1 - t_1 M)(1 - t_2 M) \]

\[ T_1 = \frac{1}{\frac{\Lambda + \lambda}{2} + \frac{\Lambda - \lambda}{2\sqrt{2}}} \]

\[ T_2 = \frac{1}{\frac{\Lambda + \lambda}{2} - \frac{\Lambda - \lambda}{2\sqrt{2}}} \]

\[ \frac{2\sqrt{2}}{\Lambda - \lambda} \left( M - \frac{1}{t_1} \right) \left( M - \frac{1}{t_2} \right) \]

\[ = \frac{2\sqrt{2}}{\Lambda - \lambda} \left( \frac{1}{t_1 t_2} \right) (1 - t_1 M)(1 - t_2 M) \]

\[ \tau_1 = \frac{2}{\Lambda + \lambda + (A - \lambda)\frac{1}{\sqrt{2}}} \]

\[ \tau_2 = \frac{2}{\Lambda + \lambda - (A - \lambda)\frac{1}{\sqrt{2}}} \]