STAT 30900 : MATHEMATICAL COMPUTATIONS I

Fall 2019

•

Homework 5

•

Solutions by

Jinhong Du

12243476

1

Let $A \in \mathbb{C}^{n\times n}$ be nonsingular. We would like to solve $A\mathbf{x} = \mathbf{b}$ with the splitting $A = M - N$ where M is nonsingular. Let $B = M^{-1}N$ and $\mathbf{c} = M^{-1}\mathbf{b}$. Consider the iteration

$$\mathbf{x}^{(k+1)} = B\mathbf{x}^{(k)} + \mathbf{c}. \tag{1}$$

By applying Problem Set **1**, Problem **4**(d) or otherwise, show that (1) converges to the solution of $A\mathbf{x} = \mathbf{b}$ for all $\mathbf{x}^{(0)}$ and all $\mathbf{b}$ if and only if $\rho(B) < 1$.

Proof. Let $\mathbf{x}$ be the solution to $A\mathbf{x} = \mathbf{b}$. Since

$$\mathbf{x}^{(k+1)} = B\mathbf{x}^{(k)} + \mathbf{c}$$
$$\mathbf{x} = B\mathbf{x} + \mathbf{c}$$

subtracting the first equation from the second one, we get

$$\mathbf{e}^{(k+1)} = B\mathbf{e}^{(k)} = B^{k+1}\mathbf{e}^{(0)},$$

where $\mathbf{e}^{(k)} = \mathbf{x}^{(k)} - \mathbf{x}$. Then $x^{(k)}$ converges to $\mathbf{x}$ if and only if $\mathbf{e}^{(k)}$ converges to $\mathbf{0}$.

$$\begin{aligned}
\lim_{k\to\infty} B^{k+1}\mathbf{e}^{(0)} = \mathbf{0},\ \forall\ \mathbf{e}^{(0)} \quad &\iff \quad \lim_{k\to\infty} B^{k+1}\mathbf{e}_i = \mathbf{0},\ \mathbf{e}_i \text{ is a unit vector with} \\
&\qquad\quad i\text{th entry being one},\ i = 1, \ldots, n \\
&\iff \quad \lim_{k\to\infty} B^{k+1} = \mathbf{0}.
\end{aligned}$$

Let $B = XJX^{-1}$ be the Jodan form of B, with block $J_1, \ldots, J_m$. So $B^{k+1} = XJ^{k+1}X^{-1}$ with

$$J_i^{k+1} = \begin{bmatrix} \lambda_i^{k+1} & \binom{k+1}{1}\lambda_i^{k-1} & \cdots & \binom{k+1}{n_i-1}\lambda^{k-(n_i-1)} \\ & \lambda_i^{k+1} & \cdots & \vdots \\ & & \ddots & \vdots \\ & & & \lambda_i^{k+1} \end{bmatrix}$$

and

$$\begin{aligned}
\lim_{k\to\infty} B^{k+1} = \mathbf{0} \quad &\iff \quad J_i^{(k+1)} \to 0,\ i = 1, \ldots, m \\
&\iff \quad |\lambda_i| < 1,\ i = 1, \ldots, m \\
&\iff \quad \rho(B) < 1.
\end{aligned}$$

Therefore, (1.1) converges to the solution of $A\mathbf{x} = \mathbf{b}$ for all $\mathbf{x}^{(0)}$ and all $\mathbf{b}$ if and only if $\rho(B) < 1$. $\square$

In general, a *semi-iterative method* is one that comprises two steps:

$$\mathbf{x}^{(k+1)} = M\mathbf{x}^{(k)} + \mathbf{b} \tag{Iteration}$$

and

$$\mathbf{y}^{(m)} = \sum_{k=0}^{m} \alpha_k^{(m)} \mathbf{x}^{(k)}. \tag{Extrapolation}$$

As in the lectures, we will assume that $M = I - A$ with $\rho(M) < 1$ and that we are interested to solve $A\mathbf{x} = \mathbf{b}$ for some nonsingular matrix $A \in \mathbb{C}^{n\times n}$. Let

$$\mathbf{e}^{(k)} = \mathbf{x}^{(k)} - \mathbf{x} \qquad \text{and} \qquad \boldsymbol{\varepsilon}^{(m)} = \mathbf{y}^{(m)} - \mathbf{x}.$$

(a) By considering what happens when $\mathbf{x}^{(0)} = \mathbf{x}$, show that it is natural to impose

$$\sum_{k=0}^{m} \alpha_k^{(m)} = 1 \tag{2}$$

for all $m \in \mathbb{N} \cup \{0\}$. Henceforth, we will assume that (2) is satisfied for all problems in this problem set.

Proof. Since

$$\begin{aligned} M\mathbf{x} + \mathbf{b} &= (I - A)\mathbf{x} + \mathbf{b} \\ &= \mathbf{x}, \end{aligned}$$

if $\mathbf{x}^{(0)} = \mathbf{x}$, then $\mathbf{x}^{(k)} = \mathbf{x}$ for $k = 0, 1, \ldots, m$. It is natural to impose that

$$\mathbf{y}^{(m)} = \mathbf{x},$$

i.e.,

$$\sum_{k=0}^{m} \alpha_k^{(m)} \mathbf{x}^{(k)} = \sum_{k=0}^{m} \alpha_k^{(m)} \mathbf{x} = \mathbf{x}..$$

So $\sum_{k=0}^{m} \alpha_k^{(m)} = 1$. $\square$

(b) Show that for all $m \in \mathbb{N}$, we may write

$$\boldsymbol{\varepsilon}^{(m)} = P_m(M)\mathbf{e}^{(0)}$$

for some $P_m(x) = \alpha_0^{(m)} + \alpha_1^{(m)}x + \cdots + \alpha_m^{(m)}x^m \in \mathbb{C}[x]$ with $\deg(P_m) \le m$ and $P_m(1) = 1$.

Proof.

$$\begin{aligned}
\boldsymbol{\epsilon}^{(m)} &= \mathbf{y}^{(m)} - \mathbf{x} \\
&= \sum_{k=0}^{m} \alpha_k^{(m)} \mathbf{x}^{(k)} - \mathbf{x} \\
&= \sum_{k=0}^{m} \alpha_k^{(m)} \mathbf{e}^{(k)} \\
\mathbf{e}^{(k)} &= M\mathbf{x}^{(k-1)} + \mathbf{b} - \mathbf{x} \\
&= M(\mathbf{x}^{(k-1)} - \mathbf{x}) + \mathbf{b} + (M - I)\mathbf{x} \\
&= M\mathbf{e}^{(k-1)} \\
&= M^k \mathbf{e}^{(0)}
\end{aligned}$$

so

$$\begin{aligned}
\boldsymbol{\epsilon}^{(m)} &= \left(\sum_{k=0}^{m} \alpha_k^{(m)} M^k \right) \mathbf{e}^{(0)} \\
&= P_m(M)\mathbf{e}^{(0)}
\end{aligned}$$

where $P_m(x) = \alpha_0^{(m)} + \cdots + \alpha_m^{(m)} x^m \in \mathbb{C}[x]$ with $\deg(P_m) \leq m$ and $P_m(1) = \sum_{k=0}^{m} \alpha_k^{(m)} = 1$. □

(c) Hence deduce that a necessary condition for $\boldsymbol{\varepsilon}^{(m)} \to \mathbf{0}$ is that

$$\lim_{m\to\infty} \|P_m(M)\|_2 < 1$$

where $\|\cdot\|_2$ is the spectral norm. Is this condition also sufficient?

Proof. Since $\boldsymbol{\epsilon}^{(m)} \to \mathbf{0}$, we have $P_m(M)\mathbf{e}^{(0)} \to \mathbf{0}$ for all $\mathbf{e}^{(0)}$. So $\lim_{m\to\infty} P_m(M) = \mathbf{0}$. So $\lim_{m\to\infty} \|P_m(M)\|_2 < 1$.

This condition is not sufficient. For example, let $\mathbf{e}^{(0)} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $\alpha_0^{(m)} = 0.9$, $\alpha_1^{(m)} = 0.1$, $M = \begin{bmatrix} -1 & \\ & -1 \end{bmatrix}$, then $P_m(M) = \begin{bmatrix} 0.8 & \\ & 0.8 \end{bmatrix}$ and $\boldsymbol{\epsilon}^{(m)} = P_m(M)\mathbf{e}^{(0)} = \begin{bmatrix} 0.8 \\ 0.8 \end{bmatrix} \nrightarrow \mathbf{0}$. □

(d) Consider the case when

$$\alpha_0^{(m)} = \alpha_1^{(m)} = \cdots = \alpha_m^{(m)} = \frac{1}{m+1}$$

for all $m \in \mathbb{N} \cup \{0\}$. Show that if a sequence (any sequence, not necessarily one generated as in (Iteration)) is convergent and

$$\lim_{k\to\infty} \mathbf{x}^{(k)} = \mathbf{x}$$

then

$$\lim_{m\to\infty} \mathbf{y}^{(m)} = \mathbf{x}.$$

Is the converse also true?

Proof. Since $\lim_{k\to\infty} x^{(k)} = \mathbf{x}$, we have that $\lim_{k\to\infty} \|x^{(k)} - \mathbf{x}\|_2 = 0$. So $\forall\ \epsilon > 0$, $\exists\ N \in \mathbb{N}$, s.t. $n > N$, $\|\mathbf{x}^{(k)} - \mathbf{x}\|_2 < \frac{\epsilon}{2}$. Then for large m such that $\frac{1}{m}\sum_{k=1}^{N} \|\mathbf{x}^{(m)} - \mathbf{x}\|_2 < \frac{\epsilon}{2}$, we have

$$
\begin{aligned}
\left\|\mathbf{y}^{(m)} - \mathbf{x}\right\|_2 &= \left\|\frac{1}{m}\sum_{k=1}^{m}(\mathbf{x}^{(m)} - \mathbf{x})\right\|_2 \\
&\leq \frac{1}{m}\sum_{k=1}^{m} \|\mathbf{x}^{(m)} - \mathbf{x}\|_2 \\
&= \frac{1}{m}\sum_{k=1}^{N} \|\mathbf{x}^{(m)} - \mathbf{x}\|_2 + \frac{1}{m}\sum_{k=N+1}^{m} \|\mathbf{x}^{(m)} - \mathbf{x}\|_2 \\
&< \frac{\epsilon}{2} + \frac{m-N+1}{m}\cdot\frac{\epsilon}{2} \\
&\leq \epsilon.
\end{aligned}
$$

Therefore, $\lim_{m\to\infty} \|y^{(m)} - \mathbf{x}\|_2 = 0$, i.e., $\lim_{m\to\infty} y^{(m)} = \mathbf{x}$.

The converse is not true in general. For example, if $x = 0$, $x^{(k)} = (-1)^k$, then $\lim_{k\to\infty} x^{(k)} = \infty$. Since $y^{(m)} = \begin{cases} 0 & \text{, if } m \text{ is odd} \\ \frac{1}{m+1} & \text{, if } m \text{ is even} \end{cases}$, $\lim_{m\to\infty} y^{(m)} = 0 = x$. □

3

It is clear that in any semi-iterative method defined by some $M \in \mathbb{C}^{n\times n}$ with $\rho(M) < 1$, we would like to solve the problem

$$\min_{P\in\mathbb{C}[x],\ \deg(P)=m,\ P(1)=1} \|P(M)\|_2. \tag{3}$$

Note that the condition $P(1) = 1$ is motivated by Problem **2**(a).

(a) Show that if $m \geq n$, then a solution to (3) is given by

$$P_m(x) = \frac{x^{m-n}\det(xI - M)}{\det(I - M)}.$$

You may assume the Cayley–Hamilton Theorem. How do we know that the denominator is non-zero?

Proof. From the Cayley-Hamilton Theorem, the characteristic polynomial of M is given by $p_M(x) = \det(xI - M)$ and $p_M(M) = \mathbf{0}$. Since $p_M(x)$ has degree n and sum of coefficients $p_M(1) = \det(I - M)$, $P_m(x) = \frac{x^{m-n}\det(xI-M)}{\det(I-M)}$ is a degree-m polynomial with $P_m(1) = 1$. Also, $P_m(M) = \frac{x^{m-n}p_M(M)}{\det(I-M)} = \mathbf{0}$, i.e., $\|P_m(M)\|_2 = 0$ which minimizes problem (3). Since $\rho(M) < 1$, all eigenvalues of M is strictly less than 1. Therefore, the eigenvalues of $I - M$ are all larger than 0, and so $\det(I - M) \neq 0$. $\square$

(b) From now on assume that M is Hermitian with minimum and maximum eigenvalues $\lambda_{\min} := a$ and $\lambda_{\max} := b \in \mathbb{R}$. Define

$$\|f\|_\infty = \sup_{x\in[a,b]} |f(x)|.$$

Emulating our discussions in the lectures, show that for $m = 0, 1, \ldots, n-1$, the solution to the relaxed problem

$$\min_{P\in\mathbb{C}[x],\ \deg(P)=m,\ P(1)=1} \|P\|_\infty \tag{4}$$

would yield an upper bound to (3).

Proof. Since M is Hermitian, it can be unitarily decomposed as $M = Q\Lambda Q^\top$ where $\Lambda \in \mathbb{C}^{n\times n}$ is a diagonal matrix with entries $\lambda_1 \geq \ldots \geq \lambda_n$ being the eigenvalues of M and $Q \in \mathbb{C}^{n\times n}$ is unitary.

$$\begin{aligned}
\|P(M)\|_2 &= \|QP(\Lambda)Q^\top\|_2 = \|P(\Lambda)\|_2 \\
&= \left\| \begin{bmatrix} P(\lambda_1) & & \\ & \ddots & \\ & & P(\lambda_n) \end{bmatrix} \right\|_2 \\
&= \max_{i=1,\ldots,n} |P(\lambda_i)| \\
&\leq \max_{\lambda\in[a,b]} |P(\lambda)| \\
&= \|P\|_\infty.
\end{aligned}$$

So (4) yields an upper bound to (3).

$\square$

(c) Consider the Chebyshev polynomials defined by

$$C_m(x) = \begin{cases} \cos(m\cos^{-1}(x)) & -1 \le x \le 1, \\ \cosh(m\cosh^{-1}(x)) & x > 1, \\ (-1)^m \cosh(m\cosh^{-1}(-x)) & x < -1. \end{cases}$$

Suppose $-1 < a < b < +1$. Show that the polynomials defined by

$$P_m(x) = \frac{C_m\left(\dfrac{2x-(b+a)}{b-a}\right)}{C_m\left(\dfrac{2-(b+a)}{b-a}\right)} \tag{5}$$

satisfy $\deg(P_m) = m$, $P_m(1) = 1$, and

$$\|P_m\|_\infty = \frac{1}{C_m\left(\dfrac{2-(b+a)}{b-a}\right)}.$$

Proof. Since

$$\begin{aligned} C_{m+1}(x) &= \cos[(m+1)\cos^{-1}(x)] \\ &= \cos(m\cos^{-1}(x) + \cos^{-1}(x)) \\ &= \cos(m\cos^{-1}(x))x - \sin(m\cos^{-1}(x))\sin(\cos^{-1}(x)) \\ C_{m-1}(x) &= \cos[(m-1)\cos^{-1}(x)] \\ &= \cos(m\cos^{-1}(x) - \cos^{-1}(x)) \\ &= \cos(m\cos^{-1}(x))x + \sin(m\cos^{-1}(x))\sin(\cos^{-1}(x)) \\ C_{m+1}(x) + C_{m-1}(x) &= 2C_m(x)x \end{aligned}$$

and so

$$C_{m+1}(x) = 2C_m(x)x - C_{m-1}(x),$$

for $m \ge 1$ with $C_0(x) = 1$ and $C_1(x) = x$. Suppose that $C_m(x)$ is degree-m for $m \le k$, then $C_{k+1}(x) = 2C_k(x)x - C_{k-1}(x)$ has degree $k+1$. By induction, $C_m(x)$ has degree m for $m \ge 0$. So $\deg(P_m) = m$. Also, $P_m(1) = \frac{C_m(\frac{2-(b+a)}{b-a})}{C_m(\frac{2-(b+a)}{b-a})} = 1$.
Suppose that $C_m(1) = 1$ for $m \le k$, then $C_{k+1}(1) = 2C_k(1) - C_{k-1}(1) = 2 - 1 = 1$. By induction, $C_m(1) = 1$ for all $m \ge 0$. Also, by definite $|C_m(x)| = |\cos(m\cos^{-1}(x))| \le 1$. So $\|C_m\|_\infty = 1$ and $\|P_m\|_\infty = \frac{1}{C_m(\frac{2-(b+a)}{b-a})}$. □

(d) By emulating our discussions in the lectures, show that the solution to (4) is given by P_m. Note that this solves (4) for all $m \in \mathbb{N}$ and not just $m \le n-1$.

Proof. Notice that $y_k = \cos\left(\frac{\pi k}{m}\right) \in [-1, 1]$ for $k = 0, \dots, m$ are $m+1$ different points such that C_m achieves extreme values $C_m(y_k) = \pm 1$ and $\|P_m(x)\|_\infty = \frac{1}{C_m(\frac{2-(b+a)}{b-a})}$ as well. And for degree-m polynomial $P_m(x)$, it has at most $m-1$ different extreme values in $\mathbb{C}$, and at most $m+1$ different

Solution (cont.)

extreme values in the interval $[a, b]$. So y_k, $k = 0, \dots, m$ are all extreme values of $P_m(x)$ in $[a, b]$. Suppose that $P_m(x)$ is not a solution to (4), then there exists another polynomial $f_m(x)$ such that $\deg f_m = m$, $f_m(1) = 1$ and $\|f_m\|_\infty < \|P_m\|_\infty$. Then

$$\begin{cases} f_m(y_k) < P_m(y_k) & , \text{ if } k \text{ is even} \\ f_m(y_k) > P_m(y_k) & , \text{ if } k \text{ is odd} \end{cases}$$

for $k = 0, \dots, m$. So the polynomial $f_m(x) - P_m(x)$ changes sign m times at (y_k, y_{k+1}) for $k = 0, \dots, m-1$, i.e., $f_m(x) - P_m(x)$ has at least m roots in $[a, b]$. Furthermore, $f_m(1) - P_m(1) = 0$, i.e. $1 > b$ is another root of $f_m(x) - P_m(x)$. So $f_m(x) - P_m(x)$ has at least $m + 1$ roots on $\mathbb{C}$. However, $\deg[f_m(x) - P_m(x)] \le m$, which means $f_m = P_m$. But f_m and P_m are different at y_k, which is a contradiction. So $P_m(x)$ achieves minimum ∞-norm in $[a, b]$ among all polynomials in $\mathbb{C}[x]$ satisfy $\deg(P) = m$ and $P(1) = 1$, which means that $P_m(x)$ is the solution to (4). $\square$

(e) Show that the solution in (d) is unique.

Proof. For fix m, suppose that there exists another solution $g(x) \in \mathbb{C}[x]$ to (4) such that $\deg(g) = m$ and $g(1) = 1$, then $\|g\|_\infty = \|P_m\|_\infty$ and $\forall\ \alpha \in (0, 1)$, the polynomial $g_\alpha(x) = \alpha P_m(x) + (1 - \alpha) g(x)$ is also a solution to (4) since $\deg(g_\alpha) = m$, $g_\alpha(1) = \alpha P_m(1) + (1 - \alpha) g(1) = 1$ and $\|g_\alpha\|_\infty \le \alpha \|P_m\|_\infty + (1 - \alpha)\|g(x)\|_\infty = \|P_m\|_\infty$. As we know that the minimum of (4) is $\|P_m\|_\infty$, $\|g_\alpha\|_\infty = \|P_m\|_\infty$.
Then for y_k $(k = 0, \dots, m)$ defined in (f), we have that $g(y_k) = P_m(y_k) \in \{\pm \|P_m\|_\infty\}$. Otherwise, there exists k_0 such that $g(y_{k_0}) \neq P_m(y_{k_0})$, if $P_m(y_{k_0}) = \|P_m\|_\infty > 0$, then

$$\begin{aligned} g_\alpha(y_{k_0}) &= \alpha P_m(y_{k_0}) + (1 - \alpha) g(y_{k_0}) \\ &< \alpha P_m(y_{k_0}) + (1 - \alpha) P_m(y_{k_0}) \\ &= P_m(y_{k_0}) \\ &= \|P_m\|_\infty, \end{aligned}$$

and

$$\begin{aligned} g_\alpha(y_{k_0}) &= \alpha P_m(y_{k_0}) + (1 - \alpha) g(y_{k_0}) \\ &\ge \alpha P_m(y_{k_0}) - (1 - \alpha)\|g\|_\infty \\ &\ge -\alpha P_m(y_{k_0}) - (1 - \alpha)\|P_m\|_\infty \\ &> -\|P_m\|_\infty, \end{aligned}$$

i.e. $|g_\alpha(y_{k_0})| < \|P_m\|_\infty$ and $\|g_\alpha\|_\infty < \|P_m\|_\infty$, which is a contradiction. Analogously, if $P_m(y_{k_0}) = -\|P_m\|_\infty$, $\|g_\alpha\|_\infty < \|P_m\|_\infty$, which is also a contradiction.
Then we have the polynomial $P_m(x) - g(x)$ has at least $m + 1$ roots y_k $(k = 0, \dots, m)$ but at most degree m, which means that $P_m(x) - g(x) \equiv 0$, i.e., $g(x) = P_m(x)$. Therefore, the solution to (4) is unique. $\square$

Let $M \in \mathbb{C}^{n\times n}$ be Hermitian with $\rho(M) = \rho < 1$. Moreover, suppose that

$$\lambda_{\min} = -\rho, \quad \lambda_{\max} = \rho.$$

(a) Show that the P_m's in (5) satisfy a three-term recurrence relation

$$C_{m+1}\left(\frac{1}{\rho}\right) P_{m+1}(x) = \frac{2x}{\rho} C_m\left(\frac{1}{\rho}\right) P_m(x) - C_{m-1}\left(\frac{1}{\rho}\right) P_{m-1}(x)$$

for all $m \in \mathbb{N}$.

Proof. $a = -\rho$ and $b = \rho$. Since $C_{m+1}(y) = 2yC_m(y) - C_{m-1}(y)$, we have

$$C_{m+1}(y) = 2yC_m(y) - C_{m-1}(y)$$

Let $y = \frac{2x-(b+a)}{b-a} = \frac{x}{\rho}$. Since $P_m(x) = \frac{C_m\left(\frac{1}{\rho}y\right)}{C_m\left(\frac{1}{\rho}\right)}$, we have

$$C_{m+1}\left(\frac{1}{\rho}\right) P_{m+1}(x) = \frac{2x}{\rho} C_m\left(\frac{1}{\rho}\right) P_m(x) - C_{m-1}\left(\frac{1}{\rho}\right) P_{m-1}(x)$$

□

(b) Show that the semi-iterative method with $\alpha_k^{(m)}$ given by the coefficient of P_m in (5) may be written as

$$\mathbf{y}^{(m+1)} = \omega_{m+1}(M\mathbf{y}^{(m)} - \mathbf{y}^{(m-1)} + \mathbf{b}) + \mathbf{y}^{(m-1)}$$

where $\mathbf{y}^{(-1)} := \mathbf{0}$, $\omega_1 := 1$, and

$$\omega_{m+1} = \frac{2C_m(1/\rho)}{\rho C_{m+1}(1/\rho)}$$

for $m = 0, 1, 2, \ldots$. This is a slightly different Chebyshev method where we choose the normalization (2) instead of $\alpha_m^{(m)} = 1$ in the lecture.

Proof. Since

$$P_{m+1}(x) = \frac{2xC_m(\frac{1}{\rho})}{\rho C_{m+1}(\frac{1}{\rho})} P_m(x) - \frac{C_{m-1}(\frac{1}{\rho})}{C_{m+1}(\frac{1}{\rho})} P_{m-1}(x)$$

and $\mathbf{y}^{(m+1)} = \sum_{k=0}^{m+1} \alpha_k^{(m+1)} \mathbf{x}^{(k)}$ with $\alpha_k^{(m+1)}$ given by the coefficient of P_{m+1}, we have

$$\begin{aligned}
\mathbf{y}^{(m+1)} &= \sum_{k=0}^{m+1} \alpha_k^{(m+1)} \mathbf{x}^{(k)} \\
&= \sum_{k=1}^{m+1} \frac{2C_m(\frac{1}{\rho})}{\rho C_{m+1}(\frac{1}{\rho})} \alpha_{k-1}^{(m)} \mathbf{x}^{(k)} - \sum_{k=0}^{m-1} \frac{C_{m-1}(\frac{1}{\rho})}{C_{m+1}(\frac{1}{\rho})} \alpha_k^{(m-1)} \mathbf{x}^{(k)} \\
&= \frac{2C_m(\frac{1}{\rho})}{\rho C_{m+1}(\frac{1}{\rho})} \sum_{k=1}^{m+1} \alpha_{k-1}^{(m)} (M\mathbf{x}^{(k-1)} + \mathbf{b}) - \frac{C_{m-1}(\frac{1}{\rho})}{C_{m+1}(\frac{1}{\rho})} \mathbf{y}^{(m-1)} \\
&= w_{m+1}(\mathbf{y}^{(m)} + \mathbf{b}) + \frac{C_{m+1}(\frac{1}{\rho}) - \frac{2}{\rho} C_m(\frac{1}{\rho})}{C_{m+1}(\frac{1}{\rho})} \mathbf{y}^{(m-1)} \\
&= \omega_{m+1}(M\mathbf{y}^{(m)} - \mathbf{y}^{(m-1)} + \mathbf{b}) + \mathbf{y}^{(m-1)}
\end{aligned}$$

□

(c) Show that

$$\|P_m(M)\|_2 = \frac{1}{C_m(1/\rho)} = \frac{1}{\cosh(m\sigma)}$$

where $\sigma = \cosh^{-1}(1/\rho)$. Deduce that $\|P_m(M)\|_2$ is a strictly decreasing sequence for all $m = 0, 1, 2 \ldots$.

Proof.

$$\|P_m(M)\|_2 = \frac{1}{C_m(\frac{1}{\rho})}\|C_m(\tfrac{1}{\rho}M)\|_2$$

Sicne M is Hermitian, M has a unitarily eigen-decomposition $M = Q\Lambda Q^\top$ with $\Lambda \in \mathbb{C}^{n\times n}$ is a diagonal matrix with diagonal entries $\lambda_1 \geq \ldots \geq \lambda_n$. Since $\rho(M) = \rho$, we have $\rho(M) = \max_i |\lambda_i| = \rho$ and $\rho(\frac{1}{\rho}M) = \frac{1}{\rho}\max_i |\lambda_i| = 1$.

$$\begin{aligned}
\|C_m(\tfrac{1}{\rho}M)\|_2 &= \|QC_m(\tfrac{1}{\rho}\Lambda)Q^\top\|_2 \\
&= \|C_m(\tfrac{1}{\rho}\Lambda)\|_2 \\
&= \left\| \begin{bmatrix} C_m(\frac{1}{\rho}\lambda_1) & & \\ & \ddots & \\ & & C_m(\frac{1}{\rho}\lambda_n) \end{bmatrix} \right\|_2 \\
&= 1
\end{aligned}$$

where the last equality holds since $|C_m(\frac{1}{\rho}\lambda_i)| \leq 1$ and there exists a λ_{i_0} such that $\frac{1}{\rho}\lambda_{i_0} = \pm 1$ and $|C_m(\pm 1)| = 1$. To see this, if $\frac{1}{\rho}\lambda_{i_0} = 1$, then $C_m(1) = 1$. If $\frac{1}{\rho}\lambda_{i_0} = -1$, then $C_m(-1) = (-1)^m$ (By induction, $C_{m+1}(-1) = -2C_m(-1) - C_{m-1}(-1) = -2 \times (-1)^m - (-1)^{m-1} = (-1)^{m+1}$).
So

$$\|P_m(M)\|_2 = \frac{1}{C_m(1/\rho)} = \frac{1}{\cosh(m\sigma)},$$

sicne $\frac{1}{\rho} > 1$. Since $\cosh(x) = \frac{e^x + e^{-x}}{2}$, $\cosh'(x) = \frac{e^x - e^{-x}}{2} > 0$ when $x > 0$, $\cosh(x)$ is strictly increasing with respect to x when $x > 1$. So $\cosh(m\sigma)$ is strictly increasing with respect to m and $\|P_m(M)\|_2 = \frac{1}{\cosh(m\sigma)}$ is strictly decreasing with respect to m for $m = 0, 1, \ldots$. □

(d) Show that

$$e^{-\sigma} = (\omega - 1)^{1/2}$$

where

$$\omega = \frac{2}{1 + \sqrt{1 - \rho^2}} \tag{6}$$

and deduce that

$$\|P_m(M)\|_2 = \frac{2(\omega - 1)^{m/2}}{1 + (\omega - 1)^m}.$$

Proof. Since $y = \cosh(x) = \frac{e^x + e^{-x}}{2}$, we have $(e^x)^2 - 2ye^x + 1 = 0$. Solving for $e^x > 1$, we have $e^x = y + \sqrt{y^2 - 1}$, i.e., $\cosh^{-1}(y) = \ln(y + \sqrt{y^2 - 1})$. So

Solution (cont.)

$$
\begin{aligned}
e^{-\sigma} = e^{-\cosh^{-1}(\frac{1}{\rho})} &= e^{-\ln\left(\frac{1}{\rho}+\sqrt{\frac{1}{\rho^2}-1}\right)} \\
&= \frac{1}{\frac{1}{\rho}+\sqrt{\frac{1}{\rho^2}-1}} \\
&= \frac{\rho}{1+\sqrt{1-\rho^2}} \\
&= \frac{\sqrt{(1-\sqrt{1-\rho^2})(1+\sqrt{1-\rho^2})}}{1+\sqrt{1-\rho^2}} \\
&= \left(\frac{1-\sqrt{1-\rho^2}}{1+\sqrt{1-\rho^2}}\right)^{\frac{1}{2}} \\
&= (\omega-1)^{\frac{1}{2}}
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
\|P_m(M)\|_2 &= \frac{1}{\cosh(m\sigma)} \\
&= \frac{2}{e^{m\sigma}+e^{-m\sigma}} \\
&= \frac{2}{(\omega-1)^{-\frac{m}{2}}+(\omega-1)^{\frac{m}{2}}} \\
&= \frac{2(\omega-1)^{\frac{m}{2}}}{1+(\omega-1)^m}.
\end{aligned}
$$

□

(e) Hence show that $(\omega_m)_{m=0}^{\infty}$ is strictly decreasing for $m \geq 2$ and that

$$\lim_{m\to\infty} \omega_m = \omega.$$

Proof. For $m \geq 2$,

$$
\begin{aligned}
\omega_m &= \frac{2C_{m-1}(\frac{1}{\rho})}{C_m(\frac{1}{\rho})} \\
&= \frac{2\|P_m(M)\|_2}{\rho\|P_{m-1}(M)\|_2} \\
&= \frac{2\frac{2(\omega-1)^{\frac{m}{2}}}{1+(\omega-1)^m}}{\rho\frac{2(\omega-1)^{\frac{m-1}{2}}}{1+(\omega-1)^{m-1}}} \\
&= \frac{2(w-1)^{\frac{1}{2}}[1+(w-1)^{m-1}]}{\rho[1+(w-1)^m]} \\
\frac{\mathrm{d}\omega_m}{\mathrm{d}m} &= \frac{2(w-1)^{\frac{2m-1}{2}}\ln(w-1)(2-w)}{\rho[1+(w-1)^m]^2} < 0
\end{aligned}
$$

since $w-1 = \frac{1-\sqrt{1-\rho^2}}{1+\sqrt{1-\rho^2}} \in (0,1)$, and $\ln(w-1) < 0$. So ω_m is strictly decreasing for $m \geq 2$, and $\omega_m \to \frac{2(w-1)^{\frac{1}{2}}}{\rho} = \frac{2}{\rho}\cdot\frac{\rho}{1+\sqrt{1-\rho^2}} = \frac{2}{1+\sqrt{1-\rho^2}} = \omega.$ □

Let $M \in \mathbb{C}^{n\times n}$ be nonsingular with $\rho(M) < 1$ and suppose we are interested in solving

$$M\mathbf{x} = \mathbf{b}. \tag{7}$$

(a) Show that SOR applied to the system

$$\begin{bmatrix} I & -M \\ -M & I \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{z} \end{bmatrix} = \begin{bmatrix} \mathbf{b} \\ \mathbf{b} \end{bmatrix} \tag{8}$$

yields the following iterations

$$\begin{aligned} \mathbf{x}^{(m+1)} &= \omega(M\mathbf{z}^{(m)} - \mathbf{x}^{(m)} + \mathbf{b}) + \mathbf{x}^{(m)}, \\ \mathbf{z}^{(m+1)} &= \omega(M\mathbf{x}^{(m+1)} - \mathbf{z}^{(m)} + \mathbf{b}) + \mathbf{z}^{(m)}, \end{aligned}$$

for $m = 0, 1, 2, \ldots$.

Proof. The matrix form of SOR iteration is given by

$$(D + \omega L)\mathbf{w}^{(m+1)} = \omega\mathbf{c} + [(1-\omega)D - \omega U]\mathbf{w}^{(m)}$$

where

$$D = I_{2n} \qquad L = \begin{bmatrix} 0 & 0 \\ -M & 0 \end{bmatrix} \qquad U = \begin{bmatrix} 0 & -M \\ 0 & 0 \end{bmatrix} \qquad \mathbf{w}^{(m)} = \begin{bmatrix} \mathbf{x}^{(m)} \\ \mathbf{z}^{(m)} \end{bmatrix} \qquad \mathbf{c} = \begin{bmatrix} \mathbf{b} \\ \mathbf{b} \end{bmatrix}.$$

So

$$\begin{bmatrix} I & 0 \\ -\omega M & I \end{bmatrix} \begin{bmatrix} \mathbf{x}^{(m+1)} \\ \mathbf{z}^{(m+1)} \end{bmatrix} = \begin{bmatrix} \omega\mathbf{b} \\ \omega\mathbf{b} \end{bmatrix} + \begin{bmatrix} (1-\omega)I & \omega M \\ 0 & (1-\omega)I \end{bmatrix} \begin{bmatrix} \mathbf{x}^{(m)} \\ \mathbf{z}^{(m)} \end{bmatrix},$$

i.e.

$$\begin{aligned} \mathbf{x}^{(m+1)} &= \omega\mathbf{b} + (1-\omega)\mathbf{x}^{(m)} + \omega M\mathbf{z}^{(m)} \\ &= \omega(M\mathbf{z}^{(m)} - \mathbf{x}^{(m)} + \mathbf{b}) + \mathbf{x}^{(m)}, \\ -\omega M\mathbf{x}^{(m+1)} + \mathbf{z}^{(m+1)} &= \omega\mathbf{b} + (1-\omega)\mathbf{z}^{(m)} \\ \mathbf{z}^{(m+1)} &= \omega\mathbf{b} + (1-\omega)\mathbf{z}^{(m)} + \omega M\mathbf{x}^{(m+1)} \\ &= \omega(M\mathbf{x}^{(m+1)} - \mathbf{z}^{(m)} + \mathbf{b}) + \mathbf{z}^{(m)}. \end{aligned}$$

$\square$

(b) Define the sequence of iterates $\mathbf{y}^{(m)}$ by

$$\mathbf{y}^{(m)} = \begin{cases} \mathbf{x}^{(k)} & \text{if } m = 2k, \\ \mathbf{z}^{(k)} & \text{if } m = 2k+1. \end{cases}$$

Show that the iterations obtained in (a) are exactly the iterations in Problem **4**(b). This shows that SOR applied to (8) is equivalent to Chebyshev applied to (7) but with $\omega_m = \omega$ for all $m \in \mathbb{N}$. Note that if ω is chosen to be the value in (6), then this is in fact the optimal SOR parameter.

Proof. For $m \geq 0$, if $m + 1 = 2k$ for some $k \in \mathbb{Z}$, then

$$\begin{aligned} \mathbf{y}^{(m+1)} &= \mathbf{x}^{(k)} \\ &= \omega(M\mathbf{z}^{(k-1)} - \mathbf{x}^{(k-1)} + \mathbf{b}) + \mathbf{x}^{(k-1)} \\ &= \omega(M\mathbf{y}^{(m)} - \mathbf{y}^{(m-1)} + \mathbf{b}) + \mathbf{y}^{(m-1)}. \end{aligned}$$

if $m + 1 = 2k + 1$ for some $k \in \mathbb{Z}$, then

$$\begin{aligned} \mathbf{y}^{(m+1)} &= \mathbf{z}^{(k)} \\ &= \omega(M\mathbf{x}^{(k)} - \mathbf{z}^{(k-1)} + \mathbf{b}) + \mathbf{z}^{(k-1)} \\ &= \omega(M\mathbf{y}^{(m)} - \mathbf{y}^{(m-1)} + \mathbf{b}) + \mathbf{y}^{(m-1)}. \end{aligned}$$

So $\mathbf{y}^{(m+1)} = \omega(M\mathbf{y}^{(m)} - \mathbf{y}^{(m-1)} + \mathbf{b}) + \mathbf{y}^{(m-1)}$ for all $m = 0, 1, \ldots$. Therefore, it is equivalent to the iterations in Problem 4(b) when $\omega_m = \omega$ for all $m \in \mathbb{N}$. $\square$

6

Let $A \in \mathbb{R}^{n \times n}$ be symmetric positive definite and $\mathbf{b} \in \mathbb{R}^n$. As usual, we write

$$\mathbf{r}_k = \mathbf{b} - A\mathbf{x}_k. \tag{9}$$

We assume that $\mathbf{x}_0$ is initialized in some manner. In the lectures we assumed $\mathbf{x}_0 = \mathbf{0}$ and so $\mathbf{r}_0 = \mathbf{b}$ but we will do it a little more generally here. Consider the quadratic functional

$$\varphi(\mathbf{x}) = \mathbf{x}^\top A\mathbf{x} - 2\mathbf{b}^\top \mathbf{x}.$$

(a) Show that

$$\nabla\varphi(\mathbf{x}_k) = -2\mathbf{r}_k$$

and hence if $\mathbf{x}_* \in \mathbb{R}^n$ is a stationary point of φ, then

$$A\mathbf{x}_* = \mathbf{b}.$$

Show also that $\mathbf{x}_*$ must be a minimizer of φ.

Proof.

$$\begin{aligned}\nabla_{\mathbf{x}_k}\varphi(\mathbf{x}_k) &= 2A\mathbf{x}_k - 2\mathbf{b}\\ &= 2(A\mathbf{x}_k - \mathbf{b})\\ &= -2\mathbf{r}_k.\end{aligned}$$

If $\mathbf{x}_*$ is a stationary point of φ, then $\nabla_{\mathbf{x}_*}\varphi(\mathbf{x}_*) = 2(A\mathbf{x}_* - \mathbf{b}) = \mathbf{0}$, i.e., $A\mathbf{x}_* = \mathbf{b}$. Also, $\nabla^2_{\mathbf{x}}\varphi(\mathbf{x}) = 2A \succ \mathbf{0}$ since A is positive definite, so the stationary point is also a minimizer of φ. $\square$

(b) Consider an iterative method

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{p}_k \tag{10}$$

where $\mathbf{p}_0, \mathbf{p}_1, \mathbf{p}_2, \ldots$ are search directions to be chosen later. Show that if we want α_k so that the function $f : \mathbb{R} \to \mathbb{R}$,

$$f(\alpha) = \varphi(\mathbf{x}_k + \alpha\mathbf{p}_k)$$

is minimized, then we must have

$$\alpha_k = \frac{\mathbf{r}_k^\top \mathbf{p}_k}{\mathbf{p}_k^\top A\mathbf{p}_k}. \tag{11}$$

Proof. Let

$$\nabla_\alpha f(\alpha) = 2\mathbf{p}_k^\top (A(\mathbf{x}_k + \alpha\mathbf{p}_k) - \mathbf{b}) = 0,$$

we have $\alpha\mathbf{p}_k^\top A\mathbf{p}_k = \mathbf{p}_k^\top(\mathbf{b} - A\mathbf{x}_k) = \mathbf{p}_k^\top \mathbf{r}_k$. Since $A \succ \mathbf{0}$ and $\mathbf{p}_k \neq \mathbf{0}$, we have $\mathbf{p}_k^\top A\mathbf{p}_k > 0$ and $\alpha_k = \frac{\mathbf{r}_k^\top \mathbf{p}_k}{\mathbf{p}_k^\top A\mathbf{p}_k}$ is a stationary point of $f(\alpha)$. Also,

$$\nabla^2_\alpha f(\alpha) = 2\mathbf{p}_k^\top A\mathbf{p}_k > 0,$$

so α_k minimizes $f(\alpha)$. $\square$

(c) Deduce that

$$\varphi(\mathbf{x}_{k+1}) - \varphi(\mathbf{x}_k) = -\frac{(\mathbf{r}_k^\top \mathbf{p}_k)^2}{\mathbf{p}_k^\top A \mathbf{p}_k}$$

and therefore $\varphi(\mathbf{x}_{k+1}) < \varphi(\mathbf{x}_k)$ as long as $\mathbf{r}_k^\top \mathbf{p}_k \neq 0$.

Proof.

$$\begin{aligned}
\varphi(\mathbf{x}_{k+1}) - \varphi(\mathbf{x}_k) &= \mathbf{x}_{k+1}^\top A \mathbf{x}_{k+1} - 2\mathbf{b}^\top \mathbf{x}_{k+1} - \mathbf{x}_k^\top A \mathbf{x}_k + 2\mathbf{b}^\top \mathbf{x}_k \\
&= (\mathbf{x}_k + \alpha_k \mathbf{p}_k)^\top A(\mathbf{x}_k + \alpha_k \mathbf{p}_k) - 2\mathbf{b}^\top(\mathbf{x}_k + \alpha_k \mathbf{p}_k) - \mathbf{x}_k^\top A \mathbf{x}_k + 2\mathbf{b}^\top \mathbf{x}_k \\
&= 2\alpha_k \mathbf{p}_k^\top A \mathbf{x}_k - 2\alpha_k \mathbf{b}^\top \mathbf{p}_k + \alpha_k^2 \mathbf{p}_k^\top A \mathbf{p}_k \\
&= 2\alpha_k \mathbf{p}_k^\top (A\mathbf{x}_k - \mathbf{b}) + \alpha_k^2 \mathbf{p}_k^\top A \mathbf{p}_k \\
&= \frac{-2(\mathbf{r}_k^\top \mathbf{p}_k)^2 + (\mathbf{r}_k^\top \mathbf{p}_k)^2}{\mathbf{p}_k^\top A \mathbf{p}_k} \\
&= -\frac{(\mathbf{r}_k^\top \mathbf{p}_k)^2}{\mathbf{p}_k^\top A \mathbf{p}_k} < 0,
\end{aligned}$$

so $\varphi(\mathbf{x}_{k+1}) < \varphi(\mathbf{x}_k)$ if $\mathbf{r}_k^\top \mathbf{p}_k \neq 0$. □

(d) Show that if we choose

$$\mathbf{p}_k = \mathbf{r}_k, \tag{12}$$

we obtain the steepest descent method discussed in the lectures.

Proof. If $\mathbf{p}_k = \mathbf{r}_k$, then $\alpha_k = \frac{\mathbf{r}_k^\top \mathbf{r}_k}{\mathbf{r}_k^\top A \mathbf{r}_k}$. Then $\mathbf{x}_{k+1} = \mathbf{x}_k + \frac{\mathbf{r}_k^\top \mathbf{r}_k}{\mathbf{r}_k^\top A \mathbf{r}_k} \mathbf{r}_k$ is the steepest descent. □

(e) Let the eigenvalues of A be $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n > 0$ and $P \in \mathbb{R}[t]$. Show that

$$\|P(A)\mathbf{x}\|_A \leq \max_{1 \leq i \leq n} |P(\lambda_i)| \|\mathbf{x}\|_A$$

for every $\mathbf{x} \in \mathbb{R}^n$. [*Hint:* $A \succ 0$ and so has an eigenbasis].

Proof. Since $A \succ \mathbf{0}$, it has a unitary eigen-decomposition as $A = Q\Lambda Q^\top$ where $\Lambda \in \mathbb{R}^{n \times n}$ with diagonal entries $\lambda_1 \geq \ldots \geq \lambda_n > 0$.

$$\begin{aligned}
\|P(A)\mathbf{x}\|_A &= \|QP(\Lambda)Q^\top \mathbf{x}\|_A \\
&= \sqrt{\mathbf{x}^\top QP(\Lambda)Q^\top AQP(\Lambda)Q^\top \mathbf{x}} \\
&= \sqrt{\mathbf{x}^\top QP(\Lambda)\Lambda P(\Lambda)Q^\top \mathbf{x}}
\end{aligned}$$

Notice that $P(\Lambda)\Lambda P(\Lambda)$ is a diag matrix and is positive definite, we have $P(\Lambda)\Lambda P(\Lambda) \preceq \max\limits_{1 \leq i \leq n} P(\lambda_i)^2 \Lambda$. So

$$\begin{aligned}
\|P(A)\mathbf{x}\|_A &\leq \sqrt{\max_{1 \leq i \leq n} P(\lambda_i)^2 \mathbf{x}^\top Q\Lambda Q^\top \mathbf{x}} \\
&= \max_{1 \leq i \leq n} |P(\lambda_i)| \|\mathbf{x}\|_A
\end{aligned}$$

□

(f) Using (e) and $P_\alpha(t) = 1 - \alpha t$, show that if we have (12), then

$$\|\mathbf{x}_k - \mathbf{x}_*\|_A \leq \max_{1\leq i\leq n} |P_\alpha(\lambda_i)| \|\mathbf{x}_{k-1} - \mathbf{x}_*\|_A$$

for all $\alpha \in \mathbb{R}$.

Proof. If $\mathbf{p}_k = \mathbf{r}_k$, then $\mathbf{x}_k = \mathbf{x}_{k-1} + \alpha_{k-1}\mathbf{r}_{k-1}$. Since $\mathbf{b} = A\mathbf{x}_*$, we have

$$\begin{aligned}\mathbf{r}_{k-1} &= \mathbf{b} - A\mathbf{x}_{k-1}\\ &= A(\mathbf{x}_* - \mathbf{x}_{k-1}).\end{aligned}$$

So

$$\begin{aligned}\mathbf{x}_k - \mathbf{x}_* &= \mathbf{x}_{k-1} - \mathbf{x}_* + \alpha_{k-1}\mathbf{r}_{k-1}\\ &= \mathbf{x}_{k-1} - \mathbf{x}_* - \alpha_{k-1}A(\mathbf{x}_{k-1} - \mathbf{x}_*)\\ &= (I - \alpha_{k-1}A)(\mathbf{x}_{k-1} - \mathbf{x}_*)\\ \|\mathbf{x}_k - \mathbf{x}_*\|_A &= \|P_{\alpha_{k-1}}(A)(\mathbf{x}_{k-1} - \mathbf{x}_*)\|_A\\ &\leq \max_{1\leq i\leq n} |P_{\alpha_{k-1}}(\lambda_i)| \cdot \|\mathbf{x}_{k-1} - \mathbf{x}_*\|_A.\end{aligned}$$

for all $\alpha_{k-1} \in \mathbb{R}$. □

(g) Using properties of Chebyshev polynomials, show that

$$\min_{\alpha\in\mathbb{R}} \max_{\lambda_n\leq t\leq\lambda_1} |1 - \alpha t| = \frac{\lambda_1 - \lambda_n}{\lambda_1 + \lambda_n}$$

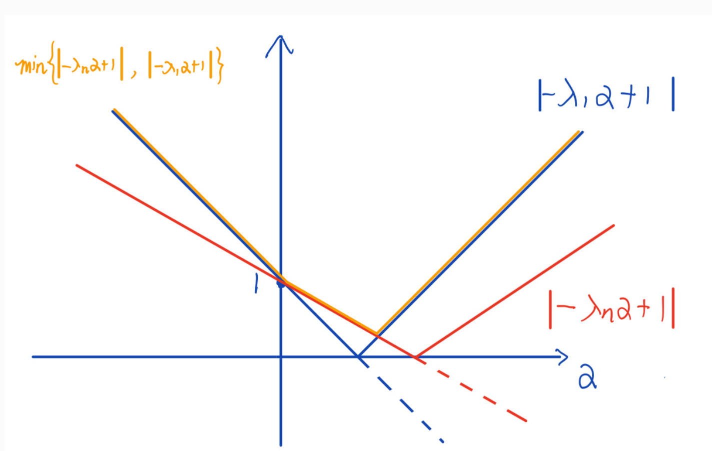
and hence deduce that

$$\|\mathbf{x}_k - \mathbf{x}_*\|_A \leq \frac{\lambda_1 - \lambda_n}{\lambda_1 + \lambda_n} \|\mathbf{x}_{k-1} - \mathbf{x}_*\|_A.$$

Proof. Since $1 - \alpha t$ is a linear function with respect to t, the maximum value of $|1 - \alpha t|$ can be achieved only at $t = \lambda_1$ or $t = \lambda_n$. So

$$\min_{\alpha\in\mathbb{R}} \max_{\lambda_n\leq t\leq\lambda_1} |1 - \alpha t| = \min_{\alpha\in\mathbb{R}} \{|1 - \lambda_1\alpha|, |1 - \lambda_n\alpha|\}.$$

If $\lambda_1 = \lambda_n$, then $\min_{\alpha\in\mathbb{R}} \max_{\lambda_n\leq t\leq\lambda_1} |1 - \alpha t| = 0 = \frac{\lambda_1 - \lambda_n}{\lambda_1 + \lambda_n}$ and the minimum is achieved for all $\alpha \in \mathbb{R}$. If $\lambda_1 > \lambda_n$,



Solution (cont.)

As we can see, the minimum is achieved when $|1-\alpha\lambda_1| = |1-\alpha\lambda_n|$ and $\alpha > 0$. Also, $\lambda_1 > \lambda_n$, so $\alpha = \frac{2}{\lambda_1+\lambda_n}$ and the minimum value is $\left|1 - \frac{2\lambda_1}{\lambda_1+\lambda_n}\right| = \frac{\lambda_1-\lambda_n}{\lambda_1+\lambda_n}$.
Therefore, from (f) we have

$$\begin{aligned}
\|\mathbf{x}_k - \mathbf{x}_*\|_A &\le \max_{1\le i\le n} |P_{\alpha_{k-1}}(\lambda_i)| \cdot \|\mathbf{x}_{k-1} - \mathbf{x}_*\|_A \\
&\le \min_{\alpha\in\mathbb{R}} \max_{\lambda_n \le t \le \lambda_1} |1-\alpha t| \|\mathbf{x}_{k-1} - \mathbf{x}_*\|_A \\
&\le \frac{\lambda_1-\lambda_n}{\lambda_1+\lambda_n} \|\mathbf{x}_{k-1} - \mathbf{x}_*\|_A.
\end{aligned}$$

□