STAT 309: MATHEMATICAL COMPUTATIONS I
FALL 2017
PROBLEM SET 5

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

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

and

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

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

$$\sum_{k=0}^{m} \alpha_k^{(m)} = 1 \qquad (1.1)$$

for all $m \in \mathbb{N} \cup \{0\}$. Henceforth, we will assume that (1.1) is satisfied for all problems in this problem set.

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

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

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

(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?

(d) Consider the case when

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

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?

Date: November 20, 2017 (Version 1.0); due: November 29, 2017.

2. 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

$$\min_{P\in\mathbb{C}[x],\ \deg(P)=m,\ P(1)=1} \|P(M)\|_2. \tag{2.2}$$

Note that in the lectures, we required the polynomial P to satisfy $P(0) = 1$. Here we use a different condition, $P(1) = 1$, motivated by Problem **1**(a).

(a) Show that if $m \geq n$, then a solution to (2.2) 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?

(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

$$\min_{P\in\mathbb{C}[x],\ \deg(P)=m,\ P(1)=1} \|P\|_\infty \tag{2.3}$$

would yield an upper bound to (2.2).

(c) Consider the Chebyshev polynomials defined by

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

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

$$P_m(x) = \frac{C_m\left(\dfrac{2x-(b+a)}{b-a}\right)}{C_m\left(\dfrac{2-(b+a)}{b-a}\right)} \tag{2.4}$$

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

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

(d) By emulating our discussions in the lectures, show that the solution to (2.3) is given by P_m. Note that this solves (2.3) for all $m \in \mathbb{N}$ and not just $m \leq n-1$.

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

3. 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 (2.4) 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 $m \in \mathbb{N}$.

(b) Show that the semi-iterative method with $\alpha_k^{(m)}$ given by the coefficient of P_m in (2.4) 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 (1.1) instead of $\alpha_m^{(m)} = 1$ in the lecture.

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

(d) Show that

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

where

$$\omega = \frac{2}{1 + \sqrt{1 - \rho^2}} \tag{3.5}$$

and deduce that

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

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

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

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

$$M\mathbf{x} = \mathbf{b}. \tag{4.6}$$

(a) Show that SOR applied to the system

$$\begin{bmatrix} I & -M \\ -M & I \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{z} \end{bmatrix} = \begin{bmatrix} \mathbf{b} \\ \mathbf{b} \end{bmatrix} \tag{4.7}$$

yields the following iterations

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

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

(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 **3**(b). This shows that SOR applied to (4.7) is equivalent to Chebyshev applied to (4.6) but with $\omega_m = \omega$ for all $m \in \mathbb{N}$. Note that if ω is chosen to be the value in (3.5), then this is in fact the optimal SOR parameter.

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

$$\mathbf{r}_k = \mathbf{b} - A\mathbf{x}_k. \tag{5.8}$$

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

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

(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.

(b) Consider an iterative method

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{p}_k \tag{5.9}$$

where $\mathbf{p}_0, \mathbf{p}_1, \mathbf{p}_2, \dots$ 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

$$\alpha_k = \frac{\mathbf{r}_k^{\mathsf{T}}\mathbf{p}_k}{\mathbf{p}_k^{\mathsf{T}} A\mathbf{p}_k}. \tag{5.10}$$

(c) Deduce that

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

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

(d) Show that if we choose

$$\mathbf{p}_k = \mathbf{r}_k, \tag{5.11}$$

we obtain the steepest decent method discussed in the lectures.

(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].

(f) Using (e) and $P_\alpha(t) = 1 - \alpha t$, show that if we have (5.11), 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 $\alpha \in \mathbb{R}$.

(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

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