Norwegian University of Science and Technology
Department of Mathematical Sciences

TMA4205 Numerical Linear Algebra Fall 2014

Solutions to exercise set 6

1

a) A is a lower triangular matrix with non-zero diagonal $\implies$ non-singular. A direct computation shows that $x_i^* = (-1)^{i+1}$.

b) $r_0 = e_1$. An inductive argument utilizing the equality $Ae_i = e_i + e_{i+1}$, $i < n$, shows that $K_m(A, r_0) = \operatorname{span}\langle e_1, \ldots, e_m \rangle$, $1 \le m \le n$.

c) By construction $x_m \in 0 + K_m(A, r_0)$, in particular $(x_m)_{m+1} = \cdots = (x_m)_n = 0$. This and the formula for x^* imply immediately that $\|x_m - x^*\|_2^2 \ge \sum_{i=m+1}^{n} (-1)^{2(i+1)} = n - m$. As $\|-A^{-1} r_m\|_2 = \|x_m - x^*\|_2$ and therefore $\|r_m\|_2 \ge \|A^{-1}\|_2^{-1} \|x_m - x^*\|_2$, from which the required bound follows.

d) Perhaps the easiest way in the present situation when the matrix is small is to find the basis vectors $r_0, Ar_0, \ldots, A^{m-1} r_0$ directly and then solve a few small least squares problems producing the polynomial coefficients of p_{m-1}.
For example, when $m = 1$ we solve $\|r_0 - y_1 A r_0\|_2 \to \min$, or in Matlab notation

```
r_0 = [1 0 0 0 0]'
A = spdiags(ones(5,2),-1:0,5,5)
y_1 = (A*r_0)\r_0
```

from which $y_1 = 0.5$, $\tilde{p}_1(t) = 1 - 0.5t$.
For $m = 2$ we have $\|r_0 - y_1 A r_0 - y_2 A^2 r_0\|_2 \to \min$, or in Matlab notation

```
y_2 = [A*r_0, A^2*r_0]\r_0
```

which yields $y = [1, -1/3]'$ and $\tilde{p}_2(t) = 1 - t + 1/3t^2$.
Proceeding in this way we find that $\tilde{p}_3(t) = 1 - 3/2t + t^2 - 1/4t^3$, $\tilde{p}_4(t) = 1 - 2t + 2t^2 - 1t^3 + 1/5t^4$, $\tilde{p}_5(t) = 1 - 5t + 10t^2 - 10t^3 + 5t^4 - t^5 = (1-t)^5 = \det(I - tA)$, see Fig. 1 (left).
Finally

i	$\|\tilde{p}_i(A)\|_2$	$\|\tilde{p}_i(A) r_0\|_2$	$\|\tilde{p}_i((A + A^{\mathrm{T}})/2)\|_2$	$\|\tilde{p}_i((A + A^{\mathrm{T}})/2) r_0\|_2$
1	0.9595	0.7071	0.8928	0.4472
2	0.8921	0.5774	0.7474	0.2673
3	0.8036	0.5000	0.5797	0.1826
4	0.7027	0.4472	0.4071	0.1348
5	0	0	0	0

e) The eigenvalues of a lower triangular matrix appear on its diagonal, thus $\sigma(A) = \{1\}$. The eigenvalue-based error bounds rely critically on the diagonalizability of A, say $A = X^{-1} \Lambda X$, when the estimate

$$\|p(A)\|_2 \le \kappa_2(X) \max_i |p(\lambda_i)|$$

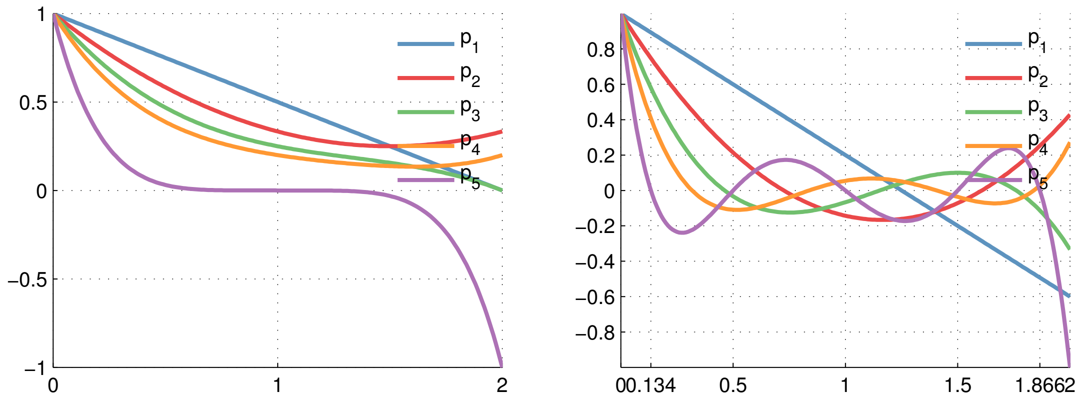

Figure 1: Optimal polynomials $\tilde{p}_i$ for A (left) and $(A+A^{\mathrm{T}})/2$ (right). Eigenvalues are marked with vertical dotted lines.

is used. Without diagonalizability this argument cannot be applied; one may only claim that

$$\|p(A)\|_2 \leq \kappa_2(X)\|p(J)\|_2,$$

where $A = X^{-1}JX$ is the Jordan canonical form of A. The behaviour of powers of Jordan blocks is relatively complicated: see Section 4.2.1 in [Saad].

$\boxed{2}$ Proceeding as in Proposition 6.32 in [Saad] we obtain the estimate

$$\begin{aligned}\frac{\|r_m\|_2}{\|r_0\|_2} &\leq \kappa_2(X) \min_{\tilde{p}_m \in \mathbb{P}_m : p(0)=1} \max_i |\tilde{p}_m(\lambda_i)| \\ &\leq \kappa_2(X) \min_{\tilde{p}_m \in \mathbb{P}_m : p(0)=1} \max\{|\tilde{p}_m(\bar{\lambda})|, \max_{\lambda_{\min} \leq \lambda \leq \lambda_{\max}} |\tilde{p}_m(\lambda)|\}\end{aligned}$$

We now replace the minimum polynomial $\tilde{p}_m$ with

$$\bar{p}_m(\lambda) = \frac{C_{m-1}(t(\lambda))}{C_{m-1}(t(0))} \frac{\bar{\lambda}-\lambda}{\bar{\lambda}},$$

which is mth degree polynomial renormalized so that $\bar{p}_m(0) = 1$. By estimating the first factor as in Theorem 6.29 in [Saad] we obtain the folowing:

$$\frac{\|r_m\|_2}{\|r_0\|_2} \leq 2\kappa_2(X) \left[\frac{\sqrt{\lambda_{\max}}-\sqrt{\lambda_{\min}}}{\sqrt{\lambda_{\max}}+\sqrt{\lambda_{\min}}}\right]^{m-1} \frac{\max\{|\bar{\lambda}-\lambda_{\min}|, |\bar{\lambda}-\lambda_{\max}|\}}{|\bar{\lambda}|}.$$

If A is normal and $|\bar{\lambda}| >> \max\{\lambda_{\min}, \lambda_{\max}\}$ then $\kappa_2(X) = 1$ and $\max\{|\bar{\lambda}-\lambda_{\min}|, |\bar{\lambda}-\lambda_{\max}|\}/|\bar{\lambda}| \approx 1$.

$\boxed{3}$ The short story is: all residuals need to be scaled by δ, but the search space and the constraint space remain the same. Indeed, $\operatorname{span}\langle r_0, Ar_0, \ldots, A^{m-1}r_0\rangle = \operatorname{span}\langle \delta r_0, \delta^2 Ar_0, \ldots, \delta^m A^{m-1} r_0\rangle$. Furthermore, $r_m \perp \mathscr{L}$ if and only if $\delta r_m \perp \mathscr{L}$, since $\mathscr{L}$ is a linear space.