Numerical Linear Algebra
Homework 10

Alexander Harnisch
harnisc6@msu.edu

26.1. Theorem 26.1 and its successors in later lectures show that we can compute eigenvalues $\{\tilde{\lambda}_k\}$ of A numerically that are the exact eigenvalues of a matrix $A+\delta A$ with $\|\delta A\|/\|A\| = O(\epsilon_{\text{machine}})$. Does this mean they are close to the exact eigenvalues $\{\lambda_k\}$ of A? This is a question of eigenvalue perturbation theory.

One can approach such problems geometrically as follows. Given $A \in \mathbb{C}^{m\times m}$ with spectrum $\Lambda(A) \subseteq \mathbb{C}$ and $\epsilon > 0$, define the 2-norm *ϵ-pseudospectrum of A*, $\Lambda_\epsilon(A)$, to be the set of numbers $z \in \mathbb{C}$ satisfying any of the following conditions:

(i) z is an eigenvalue of $A+\delta A$ for some δA with $\|\delta A\|_2 \le \epsilon$;
(ii) There exists a vector $u \in \mathbb{C}^m$ with $\|(A-zI)u\|_2 \le \epsilon$ and $\|u\|_2 = 1$;
(iii) $\sigma_m(zI-A) \le \epsilon$;
(iv) $\|(zI-A)^{-1}\|_2 \ge \epsilon^{-1}$.

The matrix $(zI-A)^{-1}$ in (iv) is known as the *resolvent* of A at z; if z is an eigenvalue of A, we use the convention $\|(zI-A)^{-1}\|_2 = \infty$. In (iii), σ_m denotes the smallest singular value.

Prove that conditions (i)–(iv) are equivalent.

Note that if z is an ev of A, all statements are obviously true and there is nothing to proove. So assume that z is not an ev of A in the following proof. (So $A - Iz$ is non-singular!)

Proof (i) => (ii):

Let u be a normalized ev of $A+\delta A$ corresponding to ew z from (i):

$$(A+\delta A)u \overset{(i)}{=} zu, \quad \|u\|_2 = 1$$

Then: $\|(A-Iz)u\|_2 = \|Au - zu\|_2 \overset{(1)}{=} \|Au - Au - \delta A u\|_2$

$= \|\delta A u\|_2 \le \|\delta A\|_2 \|u\|_2 = \|\delta A\|_2 \overset{(i)}{\le} \epsilon$ $\square$

can always be achieved through normalization

This one actually turned out to be not required, it also follows from the proofs on the next page.

Proof (ii) => (i):

We know that $\|(A-zI)u\|_2 \leq \varepsilon$ with $\|u\|_2 = 1 = u^*u$

Let $\tilde{\varepsilon} v = (A - zI)u$ with $\|v\|_2 = 1$, $\tilde{\varepsilon} > 0$

So $\|\tilde{\varepsilon} v\|_2 \leq \tilde{\varepsilon}\|v\|_2 = \tilde{\varepsilon} \leq \varepsilon$

Then: $zu = Au - \tilde{\varepsilon} v \underbrace{u^*u}_{1} = (A - \tilde{\varepsilon} v u^*)u$

So z is an ev of $A + \delta A$ with $\delta A = -\tilde{\varepsilon} v u^*$
and $\|\delta A\|_2 = \|\tilde{\varepsilon}\|_2 \|vu^*\|_2 \leq \tilde{\varepsilon}\|v\|_2\|u^*\|_2 = \tilde{\varepsilon} \leq \varepsilon$ $\square$

Proof (i) => (iv):

Again, let u be the normalized ew of $A+\delta A$ with ev z, then:

$$1 = \|u\|_2 = \|(zI-A)^{-1}\delta A u\|_2 \leq \|(zI-A)^{-1}\|_2 \|\delta A\|_2 \overset{(i)}{\leq} \|(zI-A)^{-1}\|_2\, \varepsilon$$

So $1 \leq \|(zI-A)^{-1}\|_2\, \varepsilon \iff \|(zI-A)^{-1}\|_2 \geq \varepsilon^{-1}$ $\square$

Proof (iv) => (ii):

If $\|\delta A^{-1}\|_2 = \|(zI-A)^{-1}\|_2 \geq \varepsilon^{-1}$ (iv)
then there exists a $v \in \mathbb{C}^m$ so that (follows from (3.6))

$$\frac{\|\delta A^{-1} v\|_2}{\|v\|_2} = \|\delta A^{-1}\|_2 .$$

Define $\tilde{u} = \delta A^{-1} v$

So: $\varepsilon^{-1} \leq \|\delta A^{-1}\|_2 = \dfrac{\|\delta A^{-1} v\|_2}{\|v\|_2} = \dfrac{\|\tilde{u}\|_2}{\|\delta A\tilde{u}\|_2}$

Then $u = \dfrac{\tilde{u}}{\|\tilde{u}\|_2}$ is a unit vector satisfying

$$\|\delta A u\|_2 = \|(zI-A)u\|_2 \leq \varepsilon$$ $\square$

So on this page we have shown:

(i) <=> (ii) <=> (iv)

Proof (iii) $\Leftrightarrow$ (iv): Let $\delta A = zI - A$

(iii): $\sigma_m(\delta A) \leq \varepsilon$

Let $\delta A = U \Sigma V^*$ be the SVD of δA

V and U are unitary

Hence: $\delta A^{-1} = (U \Sigma V^*)^{-1} = V \Sigma^{-1} U^*$

So the singular values of δA^{-1} are the inverse of the singular values of δA, because Σ is diagonal.

So $\sigma_m(\delta A) = \sigma_1(\delta A^{-1})^{-1} = \|\delta A^{-1}\|_2^{-1}$

So $\underbrace{\sigma_m(\delta A) \leq \varepsilon}_{(iii)} \Leftrightarrow \sigma_m(\delta A^{-1})^{-1} = \|\delta A^{-1}\|_2^{-1} \leq \varepsilon$

$\Leftrightarrow \underbrace{\|\delta A^{-1}\|_2 \geq \varepsilon^{-1}}_{(iv)}$ $\square$

We already showed that (i) $\Leftrightarrow$ (ii) $\Leftrightarrow$ (iv).
Since (iii) $\Leftrightarrow$ (iv) we conclude:

(i) $\Leftrightarrow$ (ii) $\Leftrightarrow$ (iii) $\Leftrightarrow$ (iv)

q.e.d.

27.1. Let $A \in \mathbb{C}^{m\times m}$ be given, not necessarily hermitian. Show that a number $z \in \mathbb{C}$ is a Rayleigh quotient of A if and only if it is a diagonal entry of Q^*AQ for some unitary matrix Q. Thus Rayleigh quotients are just diagonal entries of matrices, once you transform orthogonally to the right coordinate system.

Forward: Let $z \in \mathbb{C}$ be a Rayleigh coefficient of A corresponding to the vector x:

$$z = \frac{x^* A x}{x^* x}$$

Then let $\tilde{x} = \frac{x}{\sqrt{x^* x}}$ (vector x normalized)

And construct an orthonormal basis $\{\tilde{x}, q_2, \dots, q_m\}$ of $\mathbb{C}^m$. Then the matrix Q, which colums are the vectors $\{\tilde{x}, q_1, \dots, q_m\}$, is unitary. The order does not matter. Assume that x is the i-th column of Q.

Then the i-th column of AQ is Ax and the (i,i) entry of Q^*AQ is:

$$(Q^*AQ)_{ii} = \tilde{x}^* A \tilde{x} = \left(\frac{x}{\sqrt{x^* x}}\right)^* A \left(\frac{x}{\sqrt{x^* x}}\right) = \frac{x^* A x}{x^* x} = z$$

So z is the i-th diagonal entry of Q^*AQ.

Conversely, let z be the i-th diagonal entry of Q^*AQ, with Q unitary. That means:

$$z = (Q^*AQ)_{ii} = q_i^* A q_i = \frac{q_i^* A q_i}{\underbrace{q_i^* q_i}_{=1,\text{ because } Q^*Q = I}} = R(q_i)$$

where q_i is the i-th column of Q.

Thus, z is a Rayleigh coefficient of A. □

27.3. Show that for a nonhermitian matrix $A \in \mathbb{C}^{m\times m}$, the Rayleigh quotient $r(x)$ gives an eigenvalue estimate whose accuracy is generally linear, not quadratic. Explain what convergence rate this suggests for the Rayleigh quotient iteration applied to nonhermitian matrices.

Non hermitian: $r(x) = \dfrac{x^* A x}{x^* x}$

Let's derive the gradient:

$$\frac{\partial}{\partial x_j} r(x) = \frac{\frac{\partial}{\partial x_j}(x^* A x)}{x^* x} - \frac{(x^* A x)\frac{\partial}{\partial x_j}(x^* x)}{(x^* x)^2}$$

$$= \frac{\left(\frac{\partial}{\partial x_j} x^*\right) A x + x^*\left(\frac{\partial}{\partial x_j} A x\right)}{x^* x} - \frac{(x^* A x)\, 2 x_j}{(x^* x)^2} = \frac{(Ax + x^* A)_j}{x^* x} - \frac{2 r(x) x_j}{x^* x}$$

$$= \frac{1}{x^* x}\left(Ax + x^* A - 2 r(x) x\right)_j$$

$$\Rightarrow \quad \text{grad}(r(x)) = \frac{1}{x^* x}\big(\underbrace{Ax + x^* A}_{\neq 2Ax \text{ for } A^* \neq A} - 2 r(x) x\big)$$

So for q being an ev with ew λ of A:

$$\text{grad}(r(q)) = \frac{1}{\|q\|_2^2}\left(\lambda q + q^* A - 2\lambda q\right) = \frac{1}{\|q\|_2^2}\left(q^* A - \lambda q\right) \neq 0$$

So the Taylor Expansion of r around a ev looks like:

$$r(x) \overset{x \to q}{=} \underbrace{r(q)}_{\lambda} + \underbrace{\text{grad}(r(q))}_{\neq 0\,!}(x-q) + O((x-q)^2)$$

So: $r(q) - \lambda = O(\|x - q\|)$ as $x \to q$ $\square$

This suggests: $\|v^{(k+1)} - q_J\| = O(\underbrace{|\lambda^{(k)} - \lambda_J|}_{\varepsilon}\, \underbrace{\|v^{(k)} - q_J\|}_{\varepsilon}) = O(\epsilon^2)$.

quadratic instead of cubic convergence for the Rayleigh quotient iteration!