Announcements

1) HW 6 due Thursday

2) Final Friday 4/25

11:30 – 2:30

CB 2062

Cumulative!

Stage 1 of the Algorithm

Householder reflections, but not applied to all of A! We choose our first unitary in the form

$$\begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & \tilde{Q}_1 & \\ 0 & & & \end{bmatrix}$$

where $\tilde{Q}_1$ sends e_1^{m-2} to an eigenvector of the bottom right submatrix of A.

Theorem: (backwards stability)

Given $A \in \mathbb{C}^{m \times m}$ and let $\tilde{Q}$ unitary and $\tilde{H}$ Hessenberg be computed from A via the algorithm detailed in Algorithm 26.1. Then $\exists\ \delta A \in \mathbb{C}^{m \times m}$ with

$$\boxed{\tilde{Q}\tilde{H}(\tilde{Q})^* = A + \delta A = \tilde{A}}$$

and

$$\frac{\|\delta A\|}{\|A\|} = O(\varepsilon_{machine})$$

In other words, the algorithm produces a $\tilde{Q}$ and $\tilde{H}$ such that $\exists\ \tilde{A}$ whose Hessenberg form is $(\tilde{Q})^* A \tilde{Q} = \tilde{H}$;

so the algorithm is backwards stable.

Stage 2 of the Algorithm

Reduction: (symmetric matrices)

We'll assume A is real and symmetric: the Hessenberg form of A will then be tridiagonal.

Recall: (unitary diagonalizability)

$\exists$ Q real and unitary and diagonal matrix D with $A = QDQ^*$. D has real entries. In particular, the columns of Q, $\{q_1, \dots, q_m\}$, for an orthonormal basis of eigenvectors for A over $\mathbb{C}^m$.

Definition: (Rayleigh quotient)

Given $A \in \mathbb{C}^{m \times m}$, define the Rayleigh Quotient $r : \mathbb{C}^m \setminus \{0\} \to \mathbb{C}$ by

$$\boxed{r(x) = \frac{x^t A x}{x^t x}}.$$

Example 1: (eigenvectors)

Suppose that $x \in \mathbb{C}^m$ is an eigenvector for eigenvalue λ of A. Then

$$r(x) = \frac{x^t A x}{x^t x}$$

$$= \frac{x^t (\lambda x)}{x^t x}$$

$$= \lambda \frac{(x^t x)}{x^t x} = \lambda$$

Property of Rayleigh Quotient

If x is an eigenvector of A corresponding to λ,

$$\boxed{\begin{gathered}\lambda - q(x_n) = O\left(\|x - x_n\|_2^2\right) \\ \text{as} \quad \|x - x_n\|_2^2 \to 0\end{gathered}}$$

We see that we get convergence to eigenvalues out of convergence to eigenvectors.

Step 1: power iteration

We want λ_1 = the largest e-value of A (in absolute value).

If the eigenvalues of A are $\{\lambda_1, \lambda_2, \ldots, \lambda_n\}$,

$$|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \cdots \geq |\lambda_n|$$

Algorithm: Initialize at some $v^{(0)}$, $\|v^{(0)}\|_2 = 1$.

for $k = 1, 2, 3, \ldots$

$$w = A v^{(k-1)}$$

$$v^{(k)} = w / \|w\|$$

$$\lambda^{(k)} = \left(v^{(k)}\right)^t A v^{(k)}$$

Example 2:

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}$$

eigenvalue of larger absolute value $= 2 + \sqrt{5}$

$$\approx 4.23607$$

In three steps of the algorithm, we recover this approximation.

Theorem: (rate of convergence)

The power iteration algorithm 27.1 converges as follows:

Let $|\lambda_1| > |\lambda_2| \geq \cdots \geq |\lambda_m|$ be the eigenvalues for A, symmetric. Let q_1 be an eigenvector for λ_1 of unit length.

Then

$$\boxed{\| v^{(k)} - \pm e_1 \|_2 = O\left(\left| \frac{\lambda_2}{\lambda_1} \right|^k \right)}$$

and

$$\boxed{| \lambda^{(k)} - \lambda_1 | = O\left(\left| \frac{\lambda_2}{\lambda_1} \right|^{2k} \right)}$$

as $k \to \infty$.

The "$\pm$" refers to the fact that $\dim(E_{\lambda_1}) = 1$ and that sometimes the algorithm chooses the wrong vector of unit length.

proof: Observe that

since $v^{(k)} = w / \|w\|_2$

and $w = A v^{(k-1)}$,

1) $v^{(k)}$ is a unit vector for all k.

2) $$v^{(k)} = \frac{A^k v^{(0)}}{\|A^k v^{(0)}\|_2}$$

Let $\{q_1, q_2, \ldots, q_n\}$ be an orthonormal basis of eigenvectors of A corresponding to $\{\lambda_1, \lambda_2, \ldots \lambda_m\}$, respectively.

Then $\exists$ real numbers $a_1, a_2, \ldots, a_m$ with

$$\boxed{r^{(0)} = \sum_{i=1}^{m} a_i q_i}$$

$$v^{(k)} = \frac{A^k v^{(0)}}{\|A^k v^{(0)}\|_2}$$

$$= \frac{1}{\|A^k v^{(0)}\|_2} A^k \left(\sum_{i=1}^{m} a_i q_i \right)$$

$$= \frac{1}{\|A^k v^{(0)}\|_2} \sum_{i=1}^{m} a_i A^k q_i$$

$$= \frac{1}{\|A^k v^{(0)}\|_2} \sum_{i=1}^{m} a_i \lambda_i^k q_i$$

Since each q_i is an eigenvector for λ_i.

Then substituting,

$$\|v^{(k)} - q_1\|_2$$

$$= \left\| \frac{1}{\|A^k v^{(0)}\|_2} \sum_{i=1}^{m} a_i \lambda_i^k q_i - q_1 \right\|_2$$

$$= \left\| \frac{a_1 \lambda_1^k}{\|A^k v^{(0)}\|_2} q_1 - q_1 + \frac{1}{\|A^k v^{(0)}\|_2} \sum_{i=2}^{m} a_i \lambda_i^k q_i \right\|_2$$

$$\leq \left\| \frac{a_1 \lambda_1^k}{\|A^k v^{(0)}\|_2} q_1 - q_1 \right\|_2 +$$

$$\left\| \frac{1}{\|A^k v^{(0)}\|_2} \sum_{i=2}^{m} a_i \lambda_i^k q_i \right\|_2$$

by the triangle inequality

Now since $\{q_i\}_{i=1}^m$ is orthonormal,

$$\left\| \sum_{i=2}^{m} \lambda_i^k a_i q_i \right\|_2^2$$

$$= \sum_{i=2}^{m} |\lambda_i|^{2k} |a_i|^2$$

$$\leq \sum_{i=2}^{m} |\lambda_2|^{2k} |a_i|^2$$

Since $|\lambda_2| \geq |\lambda_3| \geq \cdots \geq |\lambda_m|$.

Then

$$\sum_{i=2}^{m} |\lambda_2|^{2k} |a_i|^2$$

$$= \frac{|\lambda_2|^{2k}}{|\lambda_1|^{2k}} \cdot |\lambda_1|^{2k} \sum_{i=2}^{m} |a_i|^2$$

$$= |\lambda_1|^{2k} \frac{|\lambda_2|^{2k}}{|\lambda_1|^{2k}} \sum_{i=2}^{m} |a_i|^2$$

Hence

$$\left\| \frac{1}{\|A^k r^{(0)}\|_2} \sum_{i=2}^{m} \lambda_i^k a_i q_i \right\|_2$$

$$= \frac{1}{\|A^k r^{(0)}\|_2} \left\| \sum_{i=2}^{m} \lambda_i^k a_i q_i \right\|_2$$

$$= \frac{1}{\|A^k r^{(0)}\|_2} |\lambda_1|^k \left(\frac{|\lambda_2|^k}{|\lambda_1|^k} \sqrt{\sum_{i=2}^{n} |a_i|^2} \right)$$

This last quantity is

$$O\left(\frac{|\lambda_2|^k}{|\lambda_1|^k}\right)$$ since

$$|\lambda_1| > |\lambda_2|$$

$$\Rightarrow \frac{|\lambda_2|^k}{|\lambda_1|^k} = \left(\frac{|\lambda_2|}{|\lambda_1|}\right)^k$$

$$\to 0$$

as $k \to \infty$.

Then as $k \to \infty$,

the remaining term

is $\left\| \frac{a_1 \lambda_1}{\|A^k v^{(0)}\|_2} q_1 - q_1 \right\|_2$.

But as $k \to \infty$, the first

term approaches $v^{(k)}$ which

is of norm one, and so

$$\frac{|a_1 \lambda_1|}{\|A^k v^{(0)}\|_2} \to 1 .$$

Then with the appropriate signs,

$$\left\| \frac{a_1\lambda_1}{\|A^k v^{(0)}\|_2} q_1 \pm q_1 \right\|_2$$

$$= \left| \frac{a_1\lambda_1}{\|A^k v^{(0)}\|_2} \pm 1 \right| \to 0$$

as $k \to \infty$.

This shows

$$\|v^{(k)} - q_1\|_2 = O\left(\left|\frac{\lambda_2}{\lambda_1}\right|^2\right).$$

Finally,

$$|\lambda^{k} - \lambda_1| = |q(v^{(k)}) - q(q_1)|$$

$$= O\left(\|v^{(k)} - q_1\|_2^2\right)$$

$$= O\left(\left|\frac{\lambda_2}{\lambda_1}\right|^{2k}\right)$$

by the property of Rayleigh quotients. $\square$