Homework 4, CS 6210, Fall 2020
Instructor: Austin R. Benson
Due Friday, October 30, 2020 at 10:19am ET on CMS (before lecture)

Policies

Collaboration. You are encouraged to discuss and collaborate on the homework, but you have to write up your own solutions and write your own code.
Programming language. You can use any programming language for the coding parts of the assignment. Code snippets that we provide and demos in class will use Julia.
Typesetting. Your write-up should be typeset with LaTeX. Handwritten homeworks are not accepted.
Submission. Submit your write-up as a single PDF on CMS: `https://cmsx.cs.cornell.edu`.

Problems

1. *Good identities to know.*

   Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of a matrix A. Show that

   (a) $\det(A) = \prod_{i=1}^{n} \lambda_i$ and
   (b) $\text{trace}(A) = \sum_{i=1}^{n} \lambda_j$.

   Hint: use the definition of the characteristic polynomial $p(\lambda) = \det(A - \lambda I)$ and the identity $p(\lambda) = \prod_{i=1}^{n}(\lambda_i - \lambda)$.

2. *Approximate eigenpairs are eigenpairs of nearby matrices.*

   Let $(\hat{\lambda}, \hat{x})$ be an approximate eigenpair of A with residual $r = A\hat{x} - \hat{\lambda}\hat{x}$ and $\|\hat{x}\|_2 = 1$. Show that $(\hat{\lambda}, \hat{x})$ is an eigenpair of a matrix $A + E$ where $\|E\|_F = \|r\|_2$.

3. *Sylvester.*

   Design an $O(n^3)$ algorithm to compute the solution X to the Sylvester equation

   $$AX + XB = C,$$

   where $A, B, C \in \mathbb{R}^{n \times n}$ are known. *Hint:* use the real Schur decomposition.

4. *Power method problems and solutions.*

   Suppose that we have a symmetric matrix A with eigenvalues $|\lambda_1| = |\lambda_2| > |\lambda_3| \ldots > |\lambda_n|$.

   (a) When does the power method converge for a generic starting vector?
   (b) Provide a method to get an estimate of the set $\{\lambda_1, \lambda_2\}$, even if the power method does not converge using the matrix A.

5. *A different notion of stability.*

   Recall that an ordinary differential equation

   $$\frac{d}{dt}x(t) = \dot{x}(t) = ax(t)$$

with initial condition

$$x(0) = x_0$$

has the solution

$$x(t) = e^{ct} x_0$$

Now suppose $x(t) \in \mathbb{R}^n$ and we have

$$\dot{x}(t) = Ax(t)$$

where $A \in \mathbb{R}^{n \times n}$ and $\dot{x}(t) = [\dot{x}_1(t), \dot{x}_2(t), \ldots, \dot{x}_n(t)]^T$.

(a) Show that if A is diagonalizeable, then

$$x(t) = e^{At} x_0,$$

where t is still the scalar time parameter and $e^A = \sum_{k=0}^{\infty} \frac{1}{k!} A^k$. *Hint:* If A has diagonalization $A = V \Lambda V^{-1}$, then $e^A = V e^{\Lambda} V^{-1}$ by the spectral mapping theorem.

(b) Assume from now on that $\mathrm{Re}(\lambda_1) \geq \mathrm{Re}(\lambda_2) \geq \cdots \geq \mathrm{Re}(\lambda_n)$. Describe how the long-term behavior of $x(t)$ depends on $\mathrm{sign}(\mathrm{Re}(\lambda_1))$.

(c) Compare the behavior of $x(t)$ when $\mathrm{Re}(\lambda_1) = 0$ to the behavior of the power method in the previous question when the magnitude of the first two eigenvalues in that problem are 1.

6. *Orthogonal iteration convergence.*

Implement orthogonal iteration to compute the leading r eigenvalues of a matrix $A \in \mathbb{R}^{n \times n}$. Start with some matrix $Q^{(0)} \in \mathbb{R}^{n \times r}$ with orthonormal columns and then iterate $Q^{(k+1)} R^{(k+1)} = AQ^{(k)}$ (you should use library QR routines in your code). At each step of the algorithm, also compute the eigenvalues $\tilde{\lambda}_1, \ldots, \tilde{\lambda}_r$ of $(Q^{(k+1)})^T A Q^{(k+1)}$ as estimates to the eigenvalues of the leading r eigenvalues of A (again, use a library routine to compute the $\tilde{\lambda}_j$).

For testing the algorithm, implement a routine to create a matrix $A \in \mathbb{R}^{n \times n}$ with a Schur factorization

$$A = \begin{bmatrix} Q_1 & Q_2 \end{bmatrix} \begin{bmatrix} T_{11} & T_{12} \\ & T_{22} \end{bmatrix} \begin{bmatrix} Q_1 & Q_2 \end{bmatrix}^T = QTQ^T,$$

where Q_1 is the first r columns of Q, T_{11} is the leading $r \times r$ block of T, and the diagonal of T satisfies $|T_{i,i}| \geq |T_{i+1,i+1}| > 0$ for $i = 1, \ldots, n-1$ and $|T_{r,r}| > |T_{r+1,r+1}|$. To keep things simple, you can make Q real-valued, T real-valued and upper triangular, and just take the real-valued part of the estimates $\tilde{\lambda}_j$.

(a) Make a plot that shows how convergence depends on $|T_{r+1,r+1}|/|T_{r,r}|$.

(b) Make a plot that shows how convergence depends on $\|T_{12}\|_F$.