Final Examination Solution (Practice)

Solve the following problems (4 course points each). Present a brief motivation of your method of solution.

* Prove that $\boldsymbol{A} \in \mathbb{C}^{m \times m}$ can be uniquely factored as $\boldsymbol{A} = \boldsymbol{R}\boldsymbol{U}$, with $\boldsymbol{R}$ hermitian positive definite and $\boldsymbol{U}$ unitary. (This is known as a polar factorization and generalizes the relation $z = r e^{i\theta}$)

  Solution. Inspired by $r^2 = z\,\bar{z}$, set $\boldsymbol{R}^2 = \boldsymbol{A}\boldsymbol{A}^*$. From SVD $\boldsymbol{A} = \boldsymbol{V}\boldsymbol{\Sigma}\boldsymbol{W}^*$ obtain

$$\boldsymbol{A} = \boldsymbol{V}\boldsymbol{\Sigma}\boldsymbol{V}^*\boldsymbol{V}\boldsymbol{W}^*$$

  identify $\boldsymbol{R} = \boldsymbol{V}\boldsymbol{\Sigma}\boldsymbol{V}^*$, $\boldsymbol{U} = \boldsymbol{V}\boldsymbol{W}^*$ satisfying all conditions.

* Consider $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{C}^{m \times m}$, both diagonalizable. Prove that $\boldsymbol{A}\boldsymbol{B} = \boldsymbol{B}\boldsymbol{A}$ if and only if $\boldsymbol{A}, \boldsymbol{B}$ are simultaneously diagonalizable, i.e., the same invertible matrix $\boldsymbol{P} \in \mathbb{C}^{m \times m}$ leads to $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \boldsymbol{D}_1$, $\boldsymbol{P}^{-1}\boldsymbol{B}\boldsymbol{P} = \boldsymbol{D}_2$, with $\boldsymbol{D}_1$, $\boldsymbol{D}_2$ diagonal.

  Solution. $\boldsymbol{A}\boldsymbol{B} = \boldsymbol{B}\boldsymbol{A} \Leftrightarrow \boldsymbol{A}, \boldsymbol{B}$ simultaneously diagonalizable.

  $''\Leftarrow''$: Assume $\boldsymbol{A} = \boldsymbol{P}\boldsymbol{\Lambda}_1\boldsymbol{P}^{-1}$, $\boldsymbol{B} = \boldsymbol{P}\boldsymbol{\Lambda}_2\boldsymbol{P}^{-1}$, and compute

$$\boldsymbol{A}\boldsymbol{B} = \boldsymbol{P}\boldsymbol{\Lambda}_1\boldsymbol{P}^{-1}\boldsymbol{P}\boldsymbol{\Lambda}_2\boldsymbol{P}^{-1} = \boldsymbol{P}\boldsymbol{\Lambda}_1\boldsymbol{\Lambda}_2\boldsymbol{P}^{-1}, \boldsymbol{B}\boldsymbol{A} = \boldsymbol{P}\boldsymbol{\Lambda}_2\boldsymbol{P}^{-1}\boldsymbol{P}\boldsymbol{\Lambda}_1\boldsymbol{P}^{-1} = \boldsymbol{P}\boldsymbol{\Lambda}_2\boldsymbol{\Lambda}_1\boldsymbol{P}^{-1}.$$

  Since $\boldsymbol{\Lambda}_1\boldsymbol{\Lambda}_2 = \boldsymbol{\Lambda}_2\boldsymbol{\Lambda}_1$, it results that $\boldsymbol{A}\boldsymbol{B} = \boldsymbol{B}\boldsymbol{A}$.

  $''\Rightarrow''$: Assume $\boldsymbol{A}\boldsymbol{B} = \boldsymbol{B}\boldsymbol{A}$, and consider $\boldsymbol{A}\boldsymbol{x} = \lambda\boldsymbol{x}$, $\boldsymbol{B}\boldsymbol{y} = \mu\boldsymbol{y}$. To gain insight,first assume $\boldsymbol{A}, \boldsymbol{B}$, have distinct eigenvalues (algebraic multiplicity equals 1 for all eigenvalues). From $(\boldsymbol{B}\boldsymbol{A} - \boldsymbol{A}\boldsymbol{B})\boldsymbol{x} = \boldsymbol{0}$ obtain

$$\boldsymbol{B}\boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}\boldsymbol{B}\boldsymbol{x} \Rightarrow \lambda\boldsymbol{B}\boldsymbol{x} = \boldsymbol{A}\boldsymbol{B}\boldsymbol{x}$$

  so $\boldsymbol{B}\boldsymbol{x}$ must be the eigenvector of $\boldsymbol{A}$ associated with eigenvalue λ. Since eigenvectors are determined up to a scaling, write this as $\boldsymbol{B}\boldsymbol{x} = \nu\boldsymbol{x}$, and deduce that $\boldsymbol{x}$ is also an eigenvector of $\boldsymbol{B}$. Similar considerations imply $\boldsymbol{y}$ eigenvector of $\boldsymbol{B}$ is also an eigenvector of $\boldsymbol{A}$. For eigenvalues λ with algebraic multiplicity $n > 1$, define $\boldsymbol{X} \in \mathbb{C}^{m \times n}$ as an orthonormal basis for the eigenspace of $\boldsymbol{A}$, $\boldsymbol{A}\boldsymbol{X} = \lambda\boldsymbol{X}$. By the same argument,

$$\boldsymbol{B}\boldsymbol{A}\boldsymbol{X} = \boldsymbol{A}\boldsymbol{B}\boldsymbol{X} \Rightarrow \lambda\boldsymbol{B}\boldsymbol{X} = \boldsymbol{A}\boldsymbol{B}\boldsymbol{X},$$

  hence, $\boldsymbol{B}\boldsymbol{X} \in C(\boldsymbol{X})$, hence $\boldsymbol{X}$ is also a basis for the eigenspace of $\boldsymbol{B}$, and with the same multiplicity.

* Do similar matrices have the same characteristic polynomial?

  Solution. $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{C}^{m \times m}$ are similar if there exists $\boldsymbol{T}$, s.t. $\boldsymbol{A} = \boldsymbol{T}\boldsymbol{B}\boldsymbol{T}^{-1}$. Let $p_{\boldsymbol{A}}$ denote the characteristic polynomial of $\boldsymbol{A}$. By Cayley-Hamilton theorem

$$p_{\boldsymbol{A}}(\boldsymbol{A}) = \boldsymbol{A}^m + a_1\boldsymbol{A}^{m-1} + \cdots + a_m\boldsymbol{A}^0 = \boldsymbol{0}.$$

  Replace $\boldsymbol{A} = \boldsymbol{T}\boldsymbol{B}\boldsymbol{T}^{-1}$ and obtain

$$\boldsymbol{T}p_{\boldsymbol{A}}(\boldsymbol{B})\boldsymbol{T}^{-1} = \boldsymbol{0} \Rightarrow p_{\boldsymbol{A}}(\boldsymbol{B}) = \boldsymbol{0},$$

  hence $p_{\boldsymbol{A}}$ is also the characteristic polynomial of $\boldsymbol{B}$.

* Consider $\boldsymbol{A} \in \mathbb{R}^{m \times m}$ symmetric.

  a) Construct an orthogonal matrix $\boldsymbol{Q}$ to carry out the similarity transformation

$$\boldsymbol{Q}\boldsymbol{A}\boldsymbol{Q}^T = \begin{pmatrix} \lambda & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{B} \end{pmatrix}.$$

  b) Write pseudo-code that would use the above relation to carry out eigenvalue deflation, e.g., during a QR or Lanczos algorithm.

Solution. From the above λ is an eigenvalue of $\boldsymbol{A}$, and the eigenvector $\boldsymbol{A}\boldsymbol{x} = \lambda\boldsymbol{x}$, $\|\boldsymbol{x}\| = 1$, is determined by finding a basis vector for the null space $N(\boldsymbol{A} - \lambda\boldsymbol{I})$. Note that $\lambda\boldsymbol{x}^T = (\boldsymbol{A}\boldsymbol{x})^T = \boldsymbol{x}^T\boldsymbol{A}^T = \boldsymbol{x}^T\boldsymbol{A}$. Define $\boldsymbol{C}$ to complete an orthonormal basis set $\boldsymbol{Q}^T = (\ \boldsymbol{x}\ \ \boldsymbol{C}\ )$, and compute

$$\boldsymbol{Q}\boldsymbol{A}\boldsymbol{Q}^T = \left(\begin{array}{c} \boldsymbol{x}^T \\ \boldsymbol{C}^T \end{array}\right)\boldsymbol{A}(\ \boldsymbol{x}\ \ \boldsymbol{C}\ ) = \left(\begin{array}{c} \boldsymbol{x}^T \\ \boldsymbol{C}^T \end{array}\right)(\ \boldsymbol{A}\boldsymbol{x}\ \ \boldsymbol{A}\boldsymbol{C}\ ) = \left(\begin{array}{cc} \boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x} & \boldsymbol{x}^T\boldsymbol{A}\boldsymbol{C} \\ \boldsymbol{C}^T\boldsymbol{A}\boldsymbol{x} & \boldsymbol{C}^T\boldsymbol{A}\boldsymbol{C} \end{array}\right) = \left(\begin{array}{cc} \lambda & \lambda\boldsymbol{x}^T\boldsymbol{C} \\ \lambda\ \ \boldsymbol{C}^T\boldsymbol{x} & \boldsymbol{C}^T\boldsymbol{A}\boldsymbol{C} \end{array}\right) \Rightarrow$$

$$\boldsymbol{Q}\boldsymbol{A}\boldsymbol{Q}^T = \left(\begin{array}{cc} \lambda & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{B} \end{array}\right).$$

Let $(\lambda, \boldsymbol{X}) = \mathrm{eig1}(\boldsymbol{A})$ be some eigenvalue algorithm, returning an eigenvalue λ and associated orthonormal basis $\boldsymbol{X}$ for $N(\boldsymbol{A} - \lambda\boldsymbol{I})$. Let $\mathrm{comp}(\boldsymbol{X})$ denote a procedure to return the orthogonal completion of basis set $\boldsymbol{X}$. A recursive algorithm $\mathrm{eig}(\boldsymbol{A})$ to find all eigenvalues using deflation would be carried out as

```
Input: A ∈ ℝ^{m×m}
Output: λ ∈ ℝ^m, Q ∈ ℝ^{m×m}

eig(A)
  if dim(A) ⩾ 1
    (λ, X) = eig(A); C = comp(X)
    return ( (λ, X), eig(C^T A C) )
```