Final Examination

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

1. The trace of $\boldsymbol{A}\in\mathbb{C}^{m\times m}$, $\boldsymbol{A}=(a_{ij})_{1\leqslant i,j\leqslant m}$, $i,j\in\mathbb{N}$, is defined as $\operatorname{tr}(\boldsymbol{A})=\sum_{i=1}^m a_{ii}$.
   a) Is $\operatorname{tr}(\boldsymbol{A})$ invariant under similarity transformations?
   Solution. The characteristic polynomial of $\boldsymbol{A}$ is

$$p_{\boldsymbol{A}}(\lambda)=\det(\boldsymbol{A}-\lambda\boldsymbol{I})=\lambda^m-\operatorname{tr}(\boldsymbol{A})\lambda^{m-1}+\cdots,$$

   and that of $\boldsymbol{B}=\boldsymbol{Q}\boldsymbol{A}\boldsymbol{Q}^{-1}$ is

$$p_{\boldsymbol{B}}(\lambda)=\det(\boldsymbol{B}-\lambda\boldsymbol{I})=\det(\boldsymbol{Q}\boldsymbol{A}\boldsymbol{Q}^{-1}-\lambda\boldsymbol{Q}\boldsymbol{Q}^{-1})=\det(\boldsymbol{Q}(\boldsymbol{A}-\lambda\boldsymbol{I})\boldsymbol{Q}^{-1})=\det(\boldsymbol{Q})\det(\boldsymbol{A}-\lambda\boldsymbol{I})\det(\boldsymbol{Q}^{-1})$$

   hence $p_{\boldsymbol{B}}(\lambda)=p_{\boldsymbol{A}}(\lambda)$, and the trace is invariant.
   b) Consider $\boldsymbol{B}\in\mathbb{C}^{m\times n}$, $\boldsymbol{C}\in\mathbb{C}^{n\times m}$. Prove that $\operatorname{tr}(\boldsymbol{B}\boldsymbol{C})=\operatorname{tr}(\boldsymbol{C}\boldsymbol{B})$.
   Solution. With $\boldsymbol{B}=(b_{ij})$, $\boldsymbol{C}=(c_{ij})$, compute

$$(\boldsymbol{B}\boldsymbol{C})_{ii}=\sum_{k=1}^n b_{ik}c_{ki},\ (\boldsymbol{C}\boldsymbol{B})_{ii}=\sum_{k=1}^m c_{ik}b_{ki}\Rightarrow$$

$$\operatorname{tr}(\boldsymbol{B}\boldsymbol{C})=\sum_{i=1}^m\sum_{k=1}^n b_{ik}c_{ki}=\sum_{k=1}^m\sum_{i=1}^n b_{ki}c_{ik}=\sum_{k=1}^m\sum_{i=1}^n c_{ik}b_{ki}=\sum_{i=1}^n\sum_{k=1}^m c_{ik}b_{ki}=\operatorname{tr}(\boldsymbol{C}\boldsymbol{B}).$$

   c) Does the matrix equation $\boldsymbol{A}\boldsymbol{X}-\boldsymbol{X}\boldsymbol{A}=\boldsymbol{I}$ have a solution $\boldsymbol{X}\in\mathbb{C}^{m\times m}$?.
   Solution. Take traces

$$\operatorname{tr}(\boldsymbol{A}\boldsymbol{X})=\operatorname{tr}(\boldsymbol{I}+\boldsymbol{X}\boldsymbol{A})=m+\operatorname{tr}(\boldsymbol{X}\boldsymbol{A})=m+\operatorname{tr}(\boldsymbol{A}\boldsymbol{X})\Rightarrow m=0,$$

   hence no solution is possible.
   d) For $\boldsymbol{A}$ hermitian, express $\operatorname{tr}(\boldsymbol{A})$ and $\operatorname{tr}(\boldsymbol{A}^{-1})$ in terms of eigenvalues of $\boldsymbol{A}$.
   Solution. $\boldsymbol{A}$ is unitarily diagonalizable, $\boldsymbol{A}=\boldsymbol{Q}\boldsymbol{\Lambda}\boldsymbol{Q}^*$, $\boldsymbol{A}^{-1}=\boldsymbol{Q}\boldsymbol{\Lambda}^{-1}\boldsymbol{Q}^*$ hence by (a)

$$\operatorname{tr}(\boldsymbol{A})=\operatorname{tr}(\boldsymbol{\Lambda})=\sum_{i=1}^m\lambda_i,\ \operatorname{tr}(\boldsymbol{A}^{-1})=\sum_{i=1}^m\lambda_i^{-1},\text{if }\boldsymbol{A}^{-1}\text{ exists, }i.e.\ \lambda_i\neq 0.$$

2. Consider $\boldsymbol{A}\in\mathbb{C}^{m\times m}$ nonsingular, and $\boldsymbol{c},\boldsymbol{d}\in\mathbb{C}^m$ such that $1+\boldsymbol{d}^*\boldsymbol{A}^{-1}\boldsymbol{c}\neq 0$.
   a) Prove the Sherman-Morrison formula

$$(\boldsymbol{A}+\boldsymbol{c}\boldsymbol{d}^*)^{-1}=\boldsymbol{A}^{-1}-\frac{\boldsymbol{A}^{-1}\boldsymbol{c}\boldsymbol{d}^*\boldsymbol{A}^{-1}}{1+\boldsymbol{d}^*\boldsymbol{A}^{-1}\boldsymbol{c}}.$$

   Solution. Multiply both sides by $\boldsymbol{A}+\boldsymbol{c}\boldsymbol{d}^*$ and obtain

$$\boldsymbol{0}=\boldsymbol{c}\boldsymbol{d}^*\boldsymbol{A}^{-1}-\frac{(\boldsymbol{A}+\boldsymbol{c}\boldsymbol{d}^*)\boldsymbol{A}^{-1}\boldsymbol{c}\boldsymbol{d}^*\boldsymbol{A}^{-1}}{1+\boldsymbol{d}^*\boldsymbol{A}^{-1}\boldsymbol{c}}\Rightarrow$$

$$\boldsymbol{0}=\boldsymbol{c}\boldsymbol{d}^*\boldsymbol{A}^{-1}+\boldsymbol{c}\boldsymbol{d}^*\boldsymbol{A}^{-1}\boldsymbol{d}^*\boldsymbol{A}^{-1}\boldsymbol{c}-\boldsymbol{c}\boldsymbol{d}^*\boldsymbol{A}^{-1}-\boldsymbol{c}\boldsymbol{d}^*\boldsymbol{A}^{-1}\boldsymbol{c}\boldsymbol{d}^*\boldsymbol{A}^{-1}=\boldsymbol{c}\boldsymbol{d}^*\boldsymbol{A}^{-1}\boldsymbol{d}^*\boldsymbol{A}^{-1}\boldsymbol{c}-\boldsymbol{c}\boldsymbol{d}^*\boldsymbol{A}^{-1}\boldsymbol{c}\boldsymbol{d}^*\boldsymbol{A}^{-1}$$

   Since $\boldsymbol{d}^*\boldsymbol{A}^{-1}\boldsymbol{c}$ is a scalar, $\boldsymbol{c}\boldsymbol{d}^*\boldsymbol{A}^{-1}\boldsymbol{c}\boldsymbol{d}^*\boldsymbol{A}^{-1}=\boldsymbol{c}\boldsymbol{d}^*\boldsymbol{A}^{-1}\boldsymbol{d}^*\boldsymbol{A}^{-1}\boldsymbol{c}$, and the identity is verified.
   b) Comment on the computational utility of the Sherman-Morrison formula.
   Solution. The Sherman-Morrison formula allows use of a computed $\boldsymbol{A}^{-1}$, (in practice a factorization such as $\boldsymbol{L}\boldsymbol{U}=\boldsymbol{A}$) to also compute the inverse of rank-1 perturbation of $\boldsymbol{A}$.
3. For $\boldsymbol{A},\boldsymbol{B}\in\mathbb{R}^{m\times m}$ prove:
   a) $\|\boldsymbol{A}\boldsymbol{B}\|\leqslant\|\boldsymbol{A}\|\,\|\boldsymbol{B}\|$

Solution. Assume matrix norms induced by vector norms,

$$\|\boldsymbol{A}\| = \sup_{\boldsymbol{x}} \frac{\|\boldsymbol{A}\boldsymbol{x}\|}{\|\boldsymbol{x}\|} \Rightarrow \|\boldsymbol{A}\boldsymbol{x}\| \leqslant \|\boldsymbol{A}\| \, \|\boldsymbol{x}\|.$$

Now compute

$$\|\boldsymbol{A}\boldsymbol{B}\boldsymbol{x}\| \leqslant \|\boldsymbol{A}\| \, \|\boldsymbol{B}\boldsymbol{x}\| \leqslant \|\boldsymbol{A}\| \, \|\boldsymbol{B}\| \, \|\boldsymbol{x}\| \Rightarrow \|\boldsymbol{A}\boldsymbol{B}\| \leqslant \|\boldsymbol{A}\| \, \|\boldsymbol{B}\|.$$

b)

$$\left\| \begin{pmatrix} \boldsymbol{A} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{B} \end{pmatrix} \right\|_2 = \max(\|\boldsymbol{A}\|_2, \|\boldsymbol{B}\|_2).$$

Solution. Introduce SVDs of $\boldsymbol{A}, \boldsymbol{B}$

$$\begin{pmatrix} \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{V}^T & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{P}\boldsymbol{\Sigma}\boldsymbol{Q}^T \end{pmatrix} = \begin{pmatrix} \boldsymbol{U} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{P} \end{pmatrix} \begin{pmatrix} \boldsymbol{\Lambda} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{\Sigma} \end{pmatrix} \begin{pmatrix} \boldsymbol{P} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{Q} \end{pmatrix}^T = \boldsymbol{W}\boldsymbol{D}\boldsymbol{Z}^T.$$

Note that $\boldsymbol{W}, \boldsymbol{Z}$ are orthogonal, and $\boldsymbol{D}$ is diagonal semi-positive definite, hence $\boldsymbol{W}\boldsymbol{D}\boldsymbol{Z}^T = \boldsymbol{C}$ is an SVD up to a permutation, and $\|\boldsymbol{C}\|_2$ is the largest singular value, $\|\boldsymbol{C}\|_2 = \max(\|\boldsymbol{A}\|_2, \|\boldsymbol{B}\|_2)$.

4. Consider the following analogy:

Scalars	**Matrices**
Real numbers, $z = \bar{z}$	Hermitian matrices, $\boldsymbol{A} = \boldsymbol{A}^*$
Imaginary numbers, $z = -\bar{z}$	Skew-Hermitian matrices, $\boldsymbol{A} = -\boldsymbol{A}^*$
Unit circle numbers, $z = e^{i\theta}$, $z\,\bar{z} = \bar{z}\,z = 1$	Unitary matrices, $\boldsymbol{A}\boldsymbol{A}^* = \boldsymbol{A}^*\boldsymbol{A} = \boldsymbol{I}$

Table 1. Analogy between scalars $z \in \mathbb{C}$ and matrices $\boldsymbol{A} \in \mathbb{C}^{m \times m}$

a) What is the image of the imaginary axis through the function $f(z) = (1-z)/(1+z)$?
Solution. Evaluate

$$\overline{f(z)}\, f(z) = \frac{1-\bar{z}}{1+\bar{z}} \cdot \frac{1-z}{1+z} = \frac{1-(z+\bar{z})+z\bar{z}}{1+(z+\bar{z})+z\bar{z}} = 1, \text{for } z = -\bar{z},$$

hence $\mathrm{Im}(f)$ is the unit circle.

b) For $\boldsymbol{A}$ skew-hermitian, prove that the result of the Cayley transformation

$$f(\boldsymbol{A}) = (\boldsymbol{I} - \boldsymbol{A})(\boldsymbol{I} + \boldsymbol{A})^{-1},$$

is unitary.
Solution. By analogy to above, with $\boldsymbol{Q} = \boldsymbol{f}(\boldsymbol{A})$, establish that $\boldsymbol{Q}^*\boldsymbol{Q} = \boldsymbol{Q}\boldsymbol{Q}^* = \boldsymbol{I}$. The difference from the scalar case is the adjoint involves both conjugation and transposition, $(\boldsymbol{A}\boldsymbol{B})^* = \bar{\boldsymbol{B}}^T\bar{\boldsymbol{A}}^T$, and commutativity must be considered. In the scalar case $f(z) = (1-z)(1+z)^{-1} = (1+z)^{-1}(1-z)$. Could this also hold for matrices? Yes:

$$(\boldsymbol{I} - \boldsymbol{A})(\boldsymbol{I} + \boldsymbol{A})^{-1} = -(-\boldsymbol{I} + \boldsymbol{A})(\boldsymbol{I} + \boldsymbol{A})^{-1} = -(-2\boldsymbol{I} + \boldsymbol{I} + \boldsymbol{A})(\boldsymbol{I} + \boldsymbol{A})^{-1} = 2(\boldsymbol{I} + \boldsymbol{A})^{-1} - \boldsymbol{I} =$$

$$(\boldsymbol{I} + \boldsymbol{A})^{-1}[2\boldsymbol{I} - (\boldsymbol{I} + \boldsymbol{A})] = (\boldsymbol{I} + \boldsymbol{A})^{-1}(\boldsymbol{I} - \boldsymbol{A}).$$

Replace $\boldsymbol{A}$ by $\boldsymbol{A}^*$ to obtain $(\boldsymbol{I} - \boldsymbol{A}^*)(\boldsymbol{I} + \boldsymbol{A}^*)^{-1} = (\boldsymbol{I} + \boldsymbol{A}^*)^{-1}(\boldsymbol{I} - \boldsymbol{A}^*)$. Also note that $(\boldsymbol{A}^*)^{-1} = (\boldsymbol{A}^{-1})^*$ (e.g., use SVD). With this established, by analogy to the scalar calculation, evaluate

$$f(\boldsymbol{A})^* f(\boldsymbol{A}) = [(\boldsymbol{I} - \boldsymbol{A})(\boldsymbol{I} + \boldsymbol{A})^{-1}]^*(\boldsymbol{I} - \boldsymbol{A})(\boldsymbol{I} + \boldsymbol{A})^{-1} = (\boldsymbol{I} + \boldsymbol{A}^*)^{-1}(\boldsymbol{I} - \boldsymbol{A}^*)(\boldsymbol{I} + \boldsymbol{A})^{-1}(\boldsymbol{I} - \boldsymbol{A}) =$$

$$(\boldsymbol{I} - \boldsymbol{A}^*)(\boldsymbol{I} + \boldsymbol{A}^*)^{-1}(\boldsymbol{I} + \boldsymbol{A})^{-1}(\boldsymbol{I} - \boldsymbol{A}) = (\boldsymbol{I} - \boldsymbol{A}^*)[(\boldsymbol{I} + \boldsymbol{A})(\boldsymbol{I} + \boldsymbol{A}^*)]^{-1}(\boldsymbol{I} - \boldsymbol{A}) =$$

$$(\boldsymbol{I} - \boldsymbol{A}^*)(\boldsymbol{I} + \boldsymbol{A}\boldsymbol{A}^*)^{-1}(\boldsymbol{I} - \boldsymbol{A}) = (\boldsymbol{I} - \boldsymbol{A}^*)(\boldsymbol{I} + \boldsymbol{A}\boldsymbol{A}^*)^{-1}(\boldsymbol{I} + \boldsymbol{A}\boldsymbol{A}^*)(\boldsymbol{I} - \boldsymbol{A}^*)^{-1} = \boldsymbol{I}.$$