CAAM 335: Matrix Analysis

Exam 2 Solution

Problem 1 (15 points)

The standard linear least squares problem is $\min_x \|Ax-b\|^2$ where $A \in \mathbb{R}^{m\times n}$ and $m \geq n$. A vector $x \in \mathbb{R}^n$ is a solution if and only if it satisfies the normal equations $A^TAx = A^Tb$.

Now given $C \in \mathbb{R}^{m\times n}$ and $B \in \mathbb{R}^{k\times n}$ $(k < n)$, consider the matrix least squares problem

$$\min_{X\in\mathbb{R}^{m\times k}} \|C - XB\|_F^2 \tag{1}$$

where $\|\cdot\|_F$ is the Frobenius norm (square root of sum of squares of all elements).

1. Derive *a matrix equation* involving only B, C and X such that X is a solution to (1) if and only if it satisfies the derived matrix equation (i.e., a matrix form of the "normal equations" for (1)).
2. Write down an expression for the "shortest solution" (in the form $X = \cdots$) that has the minimum Frobenius-norm among all possible solutions to (1). Justify your formula.
3. Express the residue $C - XB$ at the "shortest solution" X as a product of C with an orthogonal projection matrix. Which subspace does this projection matrix project onto?

(Hint: $\|A\|_F = \|A^T\|_F$ and $(A^T)^\dagger = (A^\dagger)^T$. The sum of squares of a matrix can be done either row by row or column by column.)

Solution:

1) After transpose, the considering the sum of squares of the matrix column by column, we have the normal equations for each column of X^T (row of X):

$$\min_{X^T\in\mathbb{R}^{k\times m}} \|B^TX^T - C^T\|_F^2 \Rightarrow \quad BB^TX^Te_j = BC^Te_j,\ j = 1,\cdots,n \quad \Rightarrow \quad BB^TX^T = BC^T$$

where e_j is the j-th unit coordinate vector (i.e., column of identity). After transpose back,

$$BB^TX^T = BC^T \quad \Leftrightarrow \quad XBB^T = CB^T$$

2) The shortest solution to the j-th column-wise normal equations is: $X^Te_j = (B^T)^\dagger C^Te_j$. Hence,

$$X^T = (B^T)^\dagger C^T \quad \Rightarrow \quad X = CB^\dagger$$

Since each row of X (column of X^T) is the minimum 2-norm solution, X itself must be the minimum Frobenius-norm solution.

3) Substituting $X = CB^\dagger$ into $C - XB$ yields

$$C - XB = C - CB^\dagger B = C(I - B^\dagger B),$$

where the orthogonal projection matrix $I - B^\dagger B$ projects onto $\mathcal{N}(B)$ since $(I - B^\dagger B)x = x$ if $Bx = 0$.

Problem 2 (20 points) Recall that the Cauchy Riemann equations for $f(z) = u(x,y) + iv(x,y)$ are

$$\frac{\partial u}{\partial x}(x,y) = \frac{\partial v}{\partial y}(x,y), \qquad \frac{\partial v}{\partial x}(x,y) = -\frac{\partial u}{\partial y}(x,y),$$

or in a short form $u_x = v_y$ and $v_x = -u_y$, which are necessary for the differentiability of $f(z)$ at $z = x + iy$, but not sufficient.

Let

$$f(z) = \begin{cases} (\bar{z})^2/z, & z \neq 0, \\ 0, & z = 0. \end{cases}$$

1. Derive real functions $u(x,y)$ and $v(x,y)$ so that $f(z) = u(x,y) + iv(x,y)$.
2. Determine whether or not $f(z)$ satisfies the Cauchy-Riemann equations at $z = 0$. (Hint: $u_x(0,0) = \lim_{x\to 0} \frac{u(x,0)-u(0,0)}{x}$ and so on.)
3. Compute the limit of $f(z)/z$ as $z \to 0$ along the line $x = y$.
4. Is the function $f(z)$ differentiable at $z = 0$? Why or why not?

Solution:
1)

$$\frac{(\bar{z})^2}{z} = \frac{(\bar{z})^3}{z\bar{z}} = \frac{(x-iy)^3}{x^2+y^2} = \frac{x^3 - 3x^2(iy) + 3x(iy)^2 - (iy)^3}{x^2+y^2} = \frac{x^3 - 3xy^2}{x^2+y^2} + i\frac{y^3 - 3x^2y}{x^2+y^2}$$

Therefore,

$$u(x,y) = \begin{cases} \frac{x^3-3xy^2}{x^2+y^2}, & x^2+y^2 \neq 0, \\ 0, & x^2+y^2 = 0. \end{cases} \qquad v(x,y) = \begin{cases} \frac{y^3-3x^2y}{x^2+y^2}, & x^2+y^2 \neq 0, \\ 0, & x^2+y^2 = 0. \end{cases}$$

2) Since $u(0,0) = v(0,0) = 0$,

$$u_x(0,0) = \lim_{x\to 0} \frac{u(x,0)}{x} = \lim_{x\to 0} \frac{x}{x} = 1, \quad v_y(0,0) = \lim_{y\to 0} \frac{v(0,y)}{y} = \lim_{y\to 0} \frac{y}{y} = 1,$$

$$v_x(0,0) = \lim_{x\to 0}\frac{v(x,0)}{x} = \lim_{x\to 0}\frac{0}{x} = 0, \quad u_y(0,0) = \lim_{y\to 0}\frac{u(0,y)}{y} = \lim_{y\to 0}\frac{0}{y} = 0,$$

Therefore, the Cauchy-Riemann equations are satisfied at $z = 0$.

3) Let $z(t) = (1+i)t$ and $t \to 0$,

$$\lim_{t\to 0}\frac{f(z(t))}{z(t)} = \lim_{t\to 0}\frac{((1-i)t)^2}{((1+i)t)^2} = \frac{(1-i)^2}{(1+i)^2} = \frac{-2i}{2i} = -1$$

4) From the calculations in 2) and 3),

$$1 = u_x(0,0) + iv_x(0,0) = \lim_{z=x\to 0}\frac{f(z)}{z} \neq \lim_{z=(1+i)t\to 0}\frac{f(z)}{z} = -1.$$

Therefore, $\lim_{z\to 0}\frac{f(z)-f(0)}{z-0}$ does not exist. Hence, f is not differentiable at $z = 0$.

Problem 3 (25 points) Let $\lambda_1, \lambda_2 \in \mathbb{C}$ and $\lambda_1 \neq \lambda_2$,

$$B = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 1 \\ 0 & 0 & \lambda_2 \end{pmatrix}, \quad x_0 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

We want to solve the dynamical system $x'(t) = Bx(t)$, $t > 0$, $x(0) = x_0$ using the Laplace transform and inverse Laplace transform. Let us denote the Laplace transform of $x(t)$ by

$$X(s) \equiv \mathcal{L}(x)(s) = \int_0^\infty e^{-st}x(t)dt.$$

For a polynomial g and a differentiable function f the inverse Laplace transform of f/g is

$$\mathcal{L}^{-1}\left(\frac{f}{g}\right)(t) = \frac{1}{2\pi i}\int_C \frac{f(z)}{g(z)}e^{zt}dz,$$

where C is a closed curve that encloses each of the roots of g.

Show all details of your work!

1. Show that the Laplace transform $X(s)$ satisfies

$$X(s) = R(s)x_0.$$

   where $R(s) = (sI - B)^{-1}$ is the resolvent of B.

2. Find the resolvent and its partial fraction expansion,

$$R(s) = \sum_{j=1}^{h} \sum_{k=1}^{m_j} \frac{R_{jk}}{(s-\lambda_j)^k}.$$

   Then compute $X(s) = R(s)x_0$.

3. Use the residue theorem to compute the inverse Laplace transforms

$$\mathcal{L}^{-1}\left(\frac{1}{(s-\lambda)^k}\right)(t), \;\; k = 1, 2.$$

4. Use the results derived above to assemble the solution $x(t)$.
5. Under what conditions $\lim_{t\to\infty} x(t) = \mathbf{0}$ for all $x(0)$?

Solution:

(1) Noting that $\mathcal{L}(x')(s) = sX(s) - x(0)$, and applying the Laplace transform to both sides of the equation, we obtain

$$sX(s) - x(0) = BX(s) \quad \Rightarrow \quad (sI-B)X(s) = x_0 \quad \Rightarrow \quad X(s) = (sI-B)^{-1}x_0.$$

(2) To find the inverse of $sI - B$,

$$\left(\begin{array}{ccc|ccc} s-\lambda_1 & 0 & 0 & 1 & 0 & 0 \\ 0 & s-\lambda_2 & -1 & 0 & 1 & 0 \\ 0 & 0 & s-\lambda_2 & 0 & 0 & 1 \end{array}\right) \to \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{1}{s-\lambda_1} & 0 & 0 \\ 0 & 1 & \frac{-1}{s-\lambda_2} & 0 & \frac{1}{s-\lambda_2} & 0 \\ 0 & 0 & 1 & 0 & 0 & \frac{1}{s-\lambda_2} \end{array}\right)$$

$$\to \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{1}{s-\lambda_1} & 0 & 0 \\ 0 & 1 & 0 & 0 & \frac{1}{s-\lambda_2} & \frac{1}{(s-\lambda_2)^2} \\ 0 & 0 & 1 & 0 & 0 & \frac{1}{s-\lambda_2} \end{array}\right)$$

$$R(s) = \frac{1}{s-\lambda_1}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \frac{1}{s-\lambda_2}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \frac{1}{(s-\lambda_2)^2}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

Hence,

$$X(s) = R(s)x_0 = \frac{1}{s-\lambda_1}\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \frac{1}{s-\lambda_2}\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} + \frac{1}{(s-\lambda_2)^2}\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

(Okay to collect the 3 terms into one vector.)

(3)

$$\mathcal{L}^{-1}\left(\frac{1}{s-\lambda}\right)(t) = res\left(\frac{e^{st}}{s-\lambda}, \lambda\right) = \lim_{s\to\lambda} e^{st} = e^{\lambda t}$$

$$\mathcal{L}^{-1}\left(\frac{1}{(s-\lambda)^2}\right)(t) = res\left(\frac{e^{st}}{(s-\lambda)^2}, \lambda\right) = \lim_{s\to\lambda} \frac{d}{ds} e^{st} = te^{\lambda t}$$

(4) From the above,

$$x(t) = \mathcal{L}^{-1}(X)(t) = e^{\lambda_1 t}\begin{pmatrix}1\\0\\0\end{pmatrix} + e^{\lambda_2 t}\begin{pmatrix}0\\1\\1\end{pmatrix} + te^{\lambda_2 t}\begin{pmatrix}0\\1\\0\end{pmatrix} = \begin{pmatrix}e^{\lambda_1 t}\\(1+t)e^{\lambda_2 t}\\e^{\lambda_2 t}\end{pmatrix}$$

(5) $x(t) \to 0$ for all $x(0)$ if and only if $Re(\lambda_j) < 0, j = 1, 2$.

Problem 4 (20 points)

It is known that every square matrix A has a Schur decomposition $A = UTU^*$ where T is upper triangular ($t_{ij} = 0, i > j$) and U is unitary ($U^*U = I$). If T is diagonal, then A is diagonalizable by a unitary transformation.

Now we consider skew-symmetric matrices that satisfy $A^T = -A$.

1. (a) Construct a nonzero, skew-symmetric matrix $A \in \mathbb{R}^{2\times 2}$ using only ± 1 and 0, find its eigenvalues λ_1, λ_2 and corresponding unit eigenvectors u_1 and u_2, and then form its eigen-decomposition $A = UDU^*$ where D is diagonal and U is unitary.
(b) Write down the 2 by 2 matrix e^A (either in an analytic or a numeric form) and determine whether it is unitary or not.
2. (a) Show that any skew-symmetric $A \in \mathbb{R}^{n\times n}$ is diagonalizable by a unitary transformation; i.e., $A = UDU^*$ where D is diagonal and U is unitary. Moreover, the eigenvalues of A are all purely imaginary.
(b) Prove that for any skew-symmetric A the matrix exponential function $\exp(At)$ is unitary for $t \in \mathbb{R}$.

Solution:

1a) By routine calculations,

$$A = \begin{pmatrix}0 & 1\\-1 & 0\end{pmatrix} \quad \Rightarrow \quad D = \begin{pmatrix}i & 0\\0 & -i\end{pmatrix}, U = \frac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\\i & -i\end{pmatrix}.$$

Other possibilities include:

$$A = \begin{pmatrix}0 & -1\\1 & 0\end{pmatrix} \quad \text{and/or} \quad D = \begin{pmatrix}-i & 0\\0 & i\end{pmatrix} \quad \text{and/or} \quad U = \frac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\\i & -i\end{pmatrix}\begin{pmatrix}c_1 & 0\\0 & c_2\end{pmatrix}$$

where $c_1, c_2 \in \mathbb{C}$ can be arbitrary as long as $|c_1| = |c_2| = 1$ (for example, $c_1 = i$ and $c_2 = -1$).

1b) Substituting $e^i = \cos(1) + i\sin(1)$ and $e^{-i} = \cos(1) - i\sin(1)$ into the following

$$\exp(A) = U \begin{pmatrix} e^i & 0 \\ 0 & e^{-i} \end{pmatrix} U^* = \begin{pmatrix} \cos(1) & \sin(1) \\ -\sin(1) & \cos(1) \end{pmatrix}$$

(If you fliped the negative sign in A, the sign in e^A would be flipped corrspondingly). It is easy to verify that e^A is indeed unitary (in fact orthogonal).

2a) From $A = U^*TU$, we have $T = U^*AU$ and $T^* = U^*A^TU = -U^*AU = -T$ since $A^* = A^T = -A$. SInce T is upper triangular and T^* is lower triangular, the triangular matrix T must be diagonal, i.e., $T = D \equiv \text{diag}(\lambda_j)$, and

$$D = -\bar{D} \quad \Rightarrow \quad \lambda_j = -\bar{\lambda}_j \quad \Rightarrow \quad Re(\lambda_j) = 0 \quad \Rightarrow \quad \lambda_j = \beta_j i.$$

2b) Since $e^{Dt} = \text{diag}(e^{i\beta_i t})$, we have $e^{Dt}(e^{Dt})^* = \text{diag}(e^{i\beta_i t}e^{-i\beta_i t}) = I$. As a result,

$$e^{At}(e^{At})^* = (Ue^{Dt}U^*)(Ue^{Dt}U^*)^* = Ue^{Dt}U^*U(e^{Dt})^*U^* = Ue^{Dt}(e^{Dt})^*U^* = UU^* = I.$$

Hence, e^{At} is unitary.

Problem 5 (20 points) Let $A \in \mathbb{R}^{10\times 5}$ have singular value decomposition $A = U\Sigma V^T$ where the first 3 diagonal entries of Σ are $\sigma_1 = 3$, $\sigma_2 = 2$ and $\sigma_3 = 1$, and the rest of the entries in Σ are all zeros. Let the columns of U be $u_1, u_2, \cdots, u_{10}$, and the columns of V be $v_1, v_2, \cdots, v_5$. Let $A^\dagger$ be the pseudo-inverse of A and I be an identity matrix of appropriate size.

Answer the following questions (justification unnecessary).

1. What is the dimension of $\mathcal{R}(A^T)$? rank(A) = # of singular values = 3
2. What is the dimension of $\mathcal{N}(A^T)$? $10 - 3 = 7$
3. What is the value of $u_1^T AA^T u_1$? $\sigma_1^2 = 9$
4. What is the value of $u_4^T AA^T u_4$? 0
5. Give an orthonormal basis for $\mathcal{N}(A)$. $\{v_4, v_5\}$
6. What is the trace of AA^T? $\sigma_1^2 + \sigma_2^2 + \sigma_3^2 = 14$
7. What is the trace of $(A^TA + I)^{-1}$? $\frac{1}{\sigma_1^2+1} + \frac{1}{\sigma_2^2+1} + \frac{1}{\sigma_3^2+1} + \frac{1}{0+1} + \frac{1}{0+1} = 14/5$
8. What is the Frobenius norm $\|A\|_F$? $\sqrt{\sigma_1^2 + \sigma_2^2 + \sigma_3^2} = \sqrt{14}$
9. What is the Frobenius norm $\|A^\dagger\|_F$? $\sqrt{1/\sigma_1^2 + 1/\sigma_2^2 + 1/\sigma_3^2} = \sqrt{49/36} = 7/6$
10. Express $I - A^\dagger A$ using the columns of V only. $v_4v_4^T + v_5v_5^T$