ECE 174 Sample Midterm Question Solutions

1. The definitions can be found in the lecture notes and textbook. Note that this list is not exhaustive and *other* definitions can also be asked for on the exam (such as the definitions of field; onto; one-to-one; rank; etc.).

2. Geometry of Least Squares and the Projection Theorem.

   (a) Domain = $\mathcal{X} = \mathbb{C}^n$. Codomain = $\mathcal{Y} = \mathbb{C}^m$. Geometry induced by A is $\mathcal{Y} = \mathcal{R}(A) \oplus \mathcal{N}(A^*)$, where $\mathcal{N}(A^*) = \mathcal{R}(A)^\perp$, and $\mathcal{X} = \mathcal{R}(A^*) \oplus \mathcal{N}(A)$, where $\mathcal{R}(A^*) = \mathcal{N}(A)^\perp$. A^* is the adjoint operator of A. Dim $\mathcal{R}(A) = r = \text{rank}(A)$, $\dim \mathcal{N}(A^*) = m - r$, $\dim \mathcal{R}(A^*) = r$, $\dim \mathcal{N}(A) = \nu = n - r$.

   (b) (i) A has rank m (A is onto). (ii) A has rank n (A is one-to-one). (iii) Either $b \neq \mathcal{R}(A)$, in which case *no* solution exists, or $b \in \mathcal{R}(A)$, in which case an infinity of solution exists.

   (c) (i) The solution, $\widehat{x}$, is such that $(b - A\widehat{x}) \perp \mathcal{R}(A)$. (ii) $(b - A\widehat{x}) \in \mathcal{R}(A)^\perp = \mathcal{N}(A^*) \Rightarrow A^*(b - A\widehat{x}) = 0 \Rightarrow A^*A\widehat{x} = A^*b$. A unique optimal solution exists because A is full column rank ($r = n$) so that the null space of A is trivial, $\mathcal{N}(A) = \{0\}$. In this case (A^*A) is invertible and the normal equations can be solved explicitly as $\widehat{x} = (A^*A)^{-1}A^*b$. Note that we have shown that $A^+ = (A^*A)^{-1}A^*$ when A is one-to-one (A has rank n).

   (d) (i) In order to insure that the solution, $\widehat{x}$, is the unique minimum norm solution we require $\widehat{x} \perp \mathcal{N}(A)$ to ensure that $\widehat{x}$ has *no* component in the nullspace of A. (ii) $\widehat{x} \in \mathcal{N}(A)^\perp = \mathcal{R}(A^*) \Rightarrow \widehat{x} = A^*\lambda$ for some $\lambda \in \mathcal{Y} \Rightarrow AA^*\lambda = b$ (from $Ax = b$) $\Rightarrow \lambda = (AA^*)^{-1}b$ (since (AA^*) is invertible) $\Rightarrow x = A^*\lambda = A^*(AA^*)^{-1}b$. Note that we have shown that $A^+ = A^*(AA^*)^{-1}$ when A in onto (A has rank m).

   (e) (i) and (ii) has been shown above. (iii) Here A is assumed to be square and full rank ($r = n = m$), and hence invertible. Furthermore A^* must also be square and full rank (since $r(A^*) = r(A)$) and is therefore also invertible. We have for case (i) that $A^+ = (A^*A)^{-1}A^* = A^{-1}(A^*)^{-1}A^* = A^{-1}$, and for case (ii) that $A^+ = A^*(AA^*)^{-1} = A^*(A^*)^{-1}A^{-1} = A^{-1}$.

3. Operator Adjoints and Quadratic Optimization.

   (a) Using the definition of the adjoint, first one determines that $A^* = A^H W$. Then, using the form of the pseudoinverse derived above for the one-to-one case, one obtains $A^+ = (A^H W A)^{-1} A^H W$.

   (b) Using the definition of the adjoint, first one determines that $A^* = \Omega^{-1}A^H$. Then, using the form of the pseudoinverse derived above for the onto case, one obtains $A^+ = \Omega^{-1}A^H(A\Omega^{-1}A^H)^{-1}$.

4. Simple Applications.

(a) We have $y_i = sx_i$, $i = 1, \cdots, m$, where s is the unknown slope. We can restate this in vector-matrix form as $y = xs$. Note that x viewed as a $m \times 1$ matrix operator is one-to-one and that its adjoint is given by $x^* = x^T$ (assuming the standard unweighted inner product on R^2). The least squares estimate for s is then given by

$$\widehat{x} = (x^T x)^{-1} x^T y = \frac{\sum_{i=1}^m x_i y_i}{\sum_{i=1}^m x_i^2} \,.$$

(b) Because of Ohm's Law, $V_i = I_i R_i$, $i = 1, 2, 3$, we can take as the definition of the input space (domain) *either* the set of possible values of the three resistor voltages *or* of the three resistor currents. Here, we choose the latter case. Taking $x = (I_1,\, I_2,\, I_3)^T$ and $y = I$, we obtain the constraint condition $y = Ax$, with $A = (1,\, 1,\, 1)$. A is obviously onto, so the choice of inner product in the output space (codomain) is irrelevant. Therefore use the simple scalar product as the inner product. For the input space (domain), the choice of inner product weighting is determined from the power dissipation formula ("Joule's Law", $P_i = I_i V_i$) and Ohm's law to be $\Omega = \text{diag}(R_1,\, R_2,\, R_3)$. Using the form of the pseudoinverse appropriate for this case derived in problem 2b above, we obtain,

$$\widehat{x} = \frac{1}{A\Omega^{-1}A^T}\, \Omega^{-1} A^T y \,.$$

This is equivalent to,

$$\begin{aligned} I_1 &= \frac{R_2 R_3\, I}{R_2 R_3 + R_1 R_3 + R_1 R_2} \,, \\ I_2 &= \frac{R_1 R_3\, I}{R_2 R_3 + R_1 R_3 + R_1 R_2} \\ I_3 &= \frac{R_1 R_2\, I}{R_2 R_3 + R_1 R_3 + R_1 R_2} \,. \end{aligned}$$

Note that the currents sum to I as required. To obtain the optimal voltages, we use Ohm's Law, $V_i = I_i R_i$, $i = 1, 2, 3$, yielding,

$$V_1 = V_2 = V_3 = \frac{R_1 R_2 R_3\, I}{R_2 R_3 + R_1 R_3 + R_1 R_2} \,.$$

Call this voltage V. The optimal (minimal) power dissipation is given by,

$$P_{\text{opt}} = I_1 V_1 + I_2 V_2 + I_3 V_3 = (I_1 + I_2 + I_3) V = IV = \frac{R_1 R_2 R_3\, I^2}{R_2 R_3 + R_1 R_3 + R_1 R_2} \,.$$

Note that when all resistors have the same value R then $P_{\text{opt}} = \frac{1}{3} R I^2$, which is $\frac{1}{3}$ the value of the power dissipation which would occur if only a simple resistor is used in the circuit.

(c) We want to minimize $\sqrt{x^2+y^2}$ subject to the linear constraint $y - ax = b$. Let $z = (x,\, y)^T$. Then the problem can be recast as minimize $\|z\|$ subject to $Az = b$ where $A = (-a,\, 1)$ is onto. Assuming the standard inner product, the adjoint of A is just A^T and we obtain the optimal z as

$$\widehat{z} = \frac{A^T b}{AA^T} = \frac{b}{1+a^2}\begin{pmatrix} -a \\ 1 \end{pmatrix}.$$

This yields the minimum distance of $\|\widehat{z}\| = \frac{|b|}{\sqrt{1+a^2}}$.

(d) In order to fit the abstract data, we want to find the parameters $x = (\alpha, \beta)^2 \in \mathbb{R}^2$ which will allow the model to best "explain" I in the least-squares sense. The inverse problem to be solve is,

$$y = \begin{pmatrix} I_1 \\ \vdots \\ I_m \end{pmatrix} = \begin{pmatrix} 1 & V_1^3 \\ \vdots & \vdots \\ 1 & V_m^3 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = A\, x\,.$$

Note that A has full column rank provided that $m \geq 2$. Assuming the standard inner product, the least-squares solution is given by,

$$\hat{x} = \begin{pmatrix} \hat{\alpha} \\ \hat{\beta} \end{pmatrix} = (A^T A)^{-1} A^T y = \cdots = \begin{pmatrix} 1 & \langle V^3 \rangle \\ \langle V^3 \rangle & \langle V^6 \rangle \end{pmatrix}^{-1} \begin{pmatrix} \langle I \rangle \\ \langle I V^3 \rangle \end{pmatrix},$$

where the indicated sample averages are given by $\langle V \rangle = \frac{1}{m}\sum_{k=1}^{m} V_k$, $\langle V^3 \rangle = \frac{1}{m}\sum_{k=1}^{m} V_k^3$, $\langle V^6 \rangle = \frac{1}{m}\sum_{k=1}^{m} V_k^6$, $\langle I \rangle = \frac{1}{m}\sum_{k=1}^{m} I_k$, and $\langle I V^3 \rangle = \frac{1}{m}\sum_{k=1}^{m} I_k V_k^3$. Using the fact that

$$\begin{pmatrix} a & b \\ b & c \end{pmatrix}^{-1} = \frac{1}{ac - b^2}\begin{pmatrix} c & -b \\ -b & a \end{pmatrix}$$

we obtain

$$\hat{\alpha} = \frac{\langle I \rangle \langle V^6 \rangle - \langle I V^3 \rangle \langle V^3 \rangle}{\langle V^6 \rangle - \langle V^3 \rangle^2},$$

and

$$\hat{\beta} = \frac{\langle I V^3 \rangle - \langle I \rangle \langle V^3 \rangle}{\langle V^6 \rangle - \langle V^3 \rangle^2}.$$

As a check, note that if the model is perfectly correct, so that $I = \alpha_0 + \beta_0 V^3$ with *no* modelling error for the specific parameter values α_0 and β_0, then the above formulas will yield $\hat{\alpha} = \alpha_0$ and $\hat{\beta} = \beta_0$.[1]

[1] Using the facts that $\langle I \rangle = \alpha_0 + \beta_0 \langle V^3 \rangle$, $\langle I V^3 \rangle = \alpha_0 \langle V^3 \rangle + \beta_0 \langle V^6 \rangle$, *et cetera.*