Assignment 7

Due Friday, 11/15/19

Reading:

- Required: Course Notes 4.1-4.4
- Required: LADR 6.C

Problems:

1. (MMA: 2.13.67) (5 pts) Let

$$\mathbf{p}_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} \quad \mathbf{p}_2 = \begin{bmatrix} 4 \\ -2 \\ -6 \\ -7 \end{bmatrix} \quad \mathbf{p}_3 = \begin{bmatrix} 3 \\ 4 \\ -2 \\ 1 \end{bmatrix}$$

   and

$$\mathbf{x} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 7 \end{bmatrix}.$$

   Determine the best approximation, $\hat{\mathbf{x}}$, of the vector $\mathbf{x}$ by vectors in $W = \text{span}[\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3]$. Also, determine the projection of $\mathbf{x}$ onto the orthogonal complement of W.

2. (MMA: 3.8.4) (5 pts) Formulate and solve the abstract least-squares regression problem

$$\min_{c_1,c_2,c_3 \in \mathbb{R}} \sum_{i=1}^{n} |y_i - c_1 - c_2 x_i - c_3 x_i^2|^2$$

   in its linear form based on $\underline{e} = A\underline{c} - \underline{y}$. Use this to fit a parabola to the (x_i, y_i) data points $(-2,2)(-1,-10)(0,0)(1,2)(2,1)$. Make a plot showing the data on top of the fitted curve.

3. (EF: 4.3.1) (5 pts) Let V be the inner-product space of continuous functions $f\colon [0,1] \to \mathbb{R}$, where

$$\langle f \mid g \rangle \triangleq \int_0^1 f(x)g(x)\mathrm{d}x.$$

   Let W be the subspace of polynomials with degree at most 2. Find the best approximation, $\hat{f}(x)$, of the continuous function $f(x) = e^x \in V$ by the polynomials in W. Plot the error $f(x) - \hat{f}(x)$ on $[0,1]$ to see the quality of the approximation.

4. (MMA: 2.13.72) (5 pts) If P is an orthogonal projection matrix, show that $I-P$ is a orthogonal projection matrix. Determine the range and nullspace of $I - P$.

5. (MMA: 3.10.14) (5 pts each) Consider a zero-mean random vector $\mathbf{x} = (x_1, x_2, x_3)$ with covariance

$$\mathrm{cov}(\mathbf{x}) \triangleq E\left[\mathbf{x}\mathbf{x}^T\right] - E\left[\mathbf{x}\right] E\left[\mathbf{x}\right]^T = \begin{bmatrix} 1 & .7 & .5 \\ .7 & 4 & .2 \\ .5 & .2 & 3 \end{bmatrix}.$$

   (a) Determine the optimal coefficients of the predictor of x_1 in terms of x_2 and x_3,

$$\hat{x}_1 = c_1 x_2 + c_2 x_3.$$

   (b) Determine the minimum mean-squared error.

   (c) How should this be estimator modified if the mean of $\mathbf{x}$ is $E[\mathbf{x}] = (1, 2, 3)^T$?

6. (MMA: 2.13.68) (5 pts) Let A be an $m \times n$ matrix that can be factored as

$$A = U \Sigma V^H, \tag{1}$$

such that U is an $m \times k$ matrix satisfying $U^H U = I$, V is an $n \times k$ matrix satisfying $V^H V = I$, and Σ is a $k \times k$ matrix with positive real entries on the diagonal and zeros elsewhere. While the conditions $U^H U = I$ and $V^H V = I$ imply that $m \geq k$ and $n \geq k$, they *do not* imply that $UU^H = I$ or $VV^H = I$. The factorization in (1) is called the compact singular-value decomposition. Show that the projection P_A onto the range of A is equal to the projection P_U onto the range of U. Give an expression for the projection matrix onto the range of A in terms of U.

Hint: Show the range of A equals the range of U but avoid the projection-matrix formula because $A^H A$ is not invertible when $n > k$.

Practice Problems (do not hand in):

1. (MMA: 4.2.10) Let $\|\cdot\|$ be a matrix norm satisfying the submultiplicative property. For a square matrix F satisfying $\|F\| < 1$, show that

$$\|(I - F)^{-1}\| \leq \frac{1}{1 - \|F\|}.$$

Hint: Use the Neumann expansion

2. (EF: 5.3.2) Suppose an imaging system measures $\underline{y} = A\underline{x}$ and wants to recover $\underline{x}$, where A is an unknown invertible matrix. During calibration, the system estimates $A \approx \tilde{A} = A + E$ where $\tilde{A}$ is invertible. Use the induced operator norm $\|\cdot\|_{\mathrm{op}}$, assuming $\|E\|_{\mathrm{op}} < 1/\|\tilde{A}^{-1}\|_{\mathrm{op}}$, to upper bound the error $\|\tilde{A}^{-1}\underline{y} - A^{-1}\underline{y}\|$ incurred by reconstructing with $\tilde{A}$ rather than A. Write your answer in terms of $\|E\|_{\mathrm{op}}$ and computable quantities like $\|\tilde{A}\|_{\mathrm{op}}$, $\|\tilde{A}^{-1}\|_{\mathrm{op}}$, $\|\tilde{A}^{-1}y\|$.

Hint: First show $\tilde{A}^{-1} - (\tilde{A} - E)^{-1} = (I - (I - \tilde{A}^{-1}E)^{-1})\tilde{A}^{-1}$, then use the Neumann expansion.

3. (LA: 8.2.10) Let V be the vector space of all $n \times n$ matrices over C, with the inner product $(A|B) = \mathrm{tr}(AB^H)$. Find the orthogonal complement of the subspace of diagonal matrices.

4. (MMA: 3.8.8) Consider the linear regression (i.e., least-squares fit) of n points $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$ to the line $y = ax + b$ defined by

$$\min_{a,b \in \mathbb{R}} \sum_{i=1}^{n} |y_i - ax_i - b|^2.$$

Set this problem up in matrix form and perform the computations to verify the slope and intercept are given by

$$\begin{aligned} a &= \frac{n \sum_{i=1}^{n} x_i y_i - \left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} y_i\right)}{n \sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}, \\ b &= \frac{\left(\sum_{i=1}^{n} x_i^2\right)\left(\sum_{i=1}^{n} y_i\right) - \left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} x_i y_i\right)}{n \sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}. \end{aligned}$$

5. (MMA: 4.3.28) Let H, G be Hilbert spaces and $A : H \to G$ be a bounded linear transform. The *adjoint* of A abstracts the transpose and is defined as the unique linear transform $A^* \colon G \to H$ satisfying $\langle A\mathbf{u} | \mathbf{w} \rangle_G = \langle \mathbf{u} | A^* \mathbf{w} \rangle_H$ for all $\mathbf{u} \in H$ and $\mathbf{w} \in G$. Assuming $G = H$, show that:

   (a) The adjoint operator A^* is linear.

   (b) The adjoint operator A^* is bounded (using the induced norm).

   (c) $\|A\| = \|A^*\|$.

6. (FSSP1: 8.12) For a real $\mathbf{H}$ with invertible $\mathbf{H}^T\mathbf{H}$, the projection matrix is given by

$$\mathbf{P} = \mathbf{H}\left(\mathbf{H}^T\mathbf{H}\right)^{-1}\mathbf{H}^T.$$

   Prove the following properties.

   (a) $\mathbf{P}$ is idempotent.

   (b) $\mathbf{P}$ is positive semidefinite.

   (c) The eigenvalues of $\mathbf{P}$ are either 1 or 0.

   (d) If $\mathbf{H}^T\mathbf{H}$ is $p \times p$, then the rank of $\mathbf{P}$ is p. Use the fact that the trace of a matrix is equal to the sum of its eigenvalues and $\mathrm{tr}(AB) = \mathrm{tr}(BA)$.