Solutions for Practice Final Exam

November 30th, 2021

Problems:

1. True or False:

   (a) **2.5 pt** – If $\|\cdot\|$ is a norm on a vector space V, then $d(\underline{u}, \underline{v}) = \|\underline{u} - \underline{v}\|$ is a metric on V.
   Solution: True.

   (b) **2.5 pt** – If A is an $n \times n$ matrix over F whose columns span the space F^n, then A is invertible.
   Solution: True. This holds because span implies basis due to dimension.

   (c) **2.5 pt** – Let the matrix $A \in \mathbb{C}^{m\times n}$ define a linear mapping between from $\mathbb{C}^n$ to $\mathbb{C}^m$. Using the standard inner product to define the four fundamental subspaces associated with A, it follows that $\mathcal{N}(A) = \mathcal{R}(A^H)^\perp$.
   Solution: True. To see this, we write

$$\begin{aligned}\mathcal{N}(A) &\triangleq \{\underline{v} \in \mathbb{C}^n \mid A\underline{v} = \underline{0}\} \\ &= \{\underline{v} \in \mathbb{C}^n \mid \underline{v}^H A^H = \underline{0}^H\} \\ &= \{\underline{v} \in \mathbb{C}^n \mid \underline{v}^H A^H \underline{u} = 0, \underline{u} \in \mathbb{C}^m\} \\ &= \{\underline{v} \in \mathbb{C}^n \mid \underline{v}^H \underline{w} = 0, \underline{w} \in \mathcal{R}(A^H)\} \\ &= \mathcal{R}(A^H)^\perp.\end{aligned}$$

   (d) **2.5 pt** – Consider the Banach space $V = \mathbb{R}^n$ with norm $\|\underline{v}\| = \sum_{i=1}^n |v_i|$ and subspace W. Then, the vector $\arg\min_{\underline{v}' \in W} \|\underline{v} - \underline{v}'\|$ is called the orthogonal projection of $\underline{v}$ onto W.
   Solution: False. It is called the best approximation of $\underline{v}$ by vectors in W. The orthogonal projection is not well-defined without an inner product.

   (e) **2.5 pt** – Let V be a vector space and $f\colon V \to \mathbb{R}$ be a convex functional on V. Then, for all $\underline{u}, \underline{v} \in V$, we have $\alpha f(\underline{u}) + (1-\alpha) f(\underline{v}) \le f(\alpha \underline{u} + (1-\alpha)\underline{v})$.
   Solution: False, the inequality is backwards.

   (f) **2.5 pt** – Let $P \in \mathbb{C}^{n\times n}$ be a projection matrix. Then, $P\underline{v} = \underline{v}$ for all $\underline{v} \in \mathcal{R}(P)$.
   Solution: True. This holds because $P\underline{v} = P(P\underline{u}) = P^2\underline{u} = P\underline{u} = \underline{v}$.

2. Short answer questions.

   (a) **2.5 pt** – For $A \in \mathcal{C}^{m\times n}$ with compact SVD $A = U_1 \Sigma_1 V_1^H$, give an expression for the orthgonal projection matrix onto the range of A.
   Solution: Since the columns of U_1 form an orthonormal basis for the range of A, we have $P_A = U_1 U_1^H$.

(b) **2.5 pt** – Let V be an inner product space and $\underline{v}_1, \underline{v}_2 \in V$ be linearly independent vectors. Give an expression for a vector $\underline{w}_2$ such that $\underline{v}_1$ and $\underline{w}_2$ form an orthogonal basis for $\mathrm{span}(\underline{v}_1, \underline{v}_2)$.

Solution: This requires one step of Gram-Schmidt orhtogonalization and gives

$$\underline{w}_2 = \underline{v}_2 - \frac{\langle \underline{v}_2 | \underline{v}_1 \rangle}{\|\underline{v}_1\|^2} \underline{v}_1.$$

(c) **2.5 pt** – What do we call an $n \times n$ matrix A where SAS^{-1} is diagonal for some invertible S?

Solution: Diagonalizable.

(d) **2.5 pt** – Let T be a linear transformation from V to W. If V is finite-dimensional, then what does the rank-nullity theorem say about the range and nullspace of T.

Solution: It says $\mathrm{rank}(T) + \mathrm{nullity}(T) = \dim(V)$.

3. Let V be a real inner product space and $\underline{w}_1, \underline{w}_2 \in V$ be vectors. Let $\hat{\underline{v}}$ be the best approximation of $\underline{v} \in V$ by vectors in $W = \mathrm{span}(\underline{w}_1, \underline{w}_2)$.

(a) **5 pt** – State the normal equations that s_1, s_2 must satisfy for $\hat{\underline{v}} = s_1 \underline{w}_1 + s_2 \underline{w}_2$

Solution: The normal equations state that the error $\underline{v} - \hat{\underline{v}}$ must be orthogonal to the vectors $\underline{w}_1, \underline{w}_2$ spanning the subspace. Thus, we have

$$\begin{bmatrix} \langle \underline{v} | \underline{w}_1 \rangle \\ \langle \underline{v} | \underline{w}_2 \rangle \end{bmatrix} = \begin{bmatrix} \langle \underline{w}_1 | \underline{w}_1 \rangle & \langle \underline{w}_2 | \underline{w}_1 \rangle \\ \langle \underline{w}_1 | \underline{w}_2 \rangle & \langle \underline{w}_2 | \underline{w}_2 \rangle \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \end{bmatrix}$$

Now, consider the collection of points $\{(x_i, y_i)\}$ in $\mathbb{R}^2$ given by

$$\{(0,2), (1,2), (2,4), (3,8)\}.$$

The goal is to fit these points to a line by solving

$$\min_{s_1, s_2} \sum_{i=1}^{4} (y_i - s_1 - s_2 x_i)^2.$$

(b) **5 pt** – Let $V = \mathbb{R}^4$ be the standard real inner product space. Find vectors $\underline{v}, \underline{w}_1, \underline{w}_2 \in V$ such that the optimal s_1, s_2 are given by the normal equations from part (a).

Solution: If we choose $\underline{v} = (y_1, y_2, y_3, y_4)^T$, $\underline{w}_1 = (1,1,1,1)^T$, and $\underline{w}_2 = (x_1, x_2, x_3, x_4)^T$, then we find that

$$\|\underline{v} - s_1 \underline{w}_1 - s_2 \underline{w}_2\|^2 = \sum_{i=1}^{4} (y_i - s_1 - s_2 x_i)^2.$$

Thus, the best projection (given by minimizing over s_1, s_2) is defined by the normal equations.

(c) **5 pt** – Find optimal values for s_1 and s_2.

Solution: Computing the inner products, one finds that the normal equations are given by

$$\begin{bmatrix} 16 \\ 34 \end{bmatrix} = \begin{bmatrix} 4 & 6 \\ 6 & 14 \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \end{bmatrix}$$

Solving for s_1, s_2 gives $s_1 = 1$ and $s_2 = 2$.

(d) **5 pt** – Find the squared error achieved by the optimal values of s_1, s_2.

Solution: Using the optimal s_1, s_2, we see that the best approximation is $\hat{\underline{v}} = (1,3,5,7)^T$. Thus, we have $\|\underline{v} - \hat{\underline{v}}\|^2 = \|(1,-1,-1,1)^T\|^2 = 4$.

4. The real matrix A and its compact SVD decomposition are defined by

$$A = \begin{bmatrix} \frac{1}{3} & -\frac{1}{2} & \frac{2}{3} & \frac{2}{3} \\ \frac{1}{3} & -\frac{1}{2} & \frac{2}{3} & \frac{2}{3} \\ \frac{1}{3} & \frac{5}{6} & 0 & -\frac{2}{3} \\ \frac{1}{3} & \frac{5}{6} & 0 & -\frac{2}{3} \end{bmatrix} = U_1 \Sigma_1 V_1^T = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & \frac{2}{3} \\ -\frac{2}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & 0 \end{bmatrix}^T .$$

(a) **5 pt** – Give two orthonormal bases, one for the range of A and the other for the range of A^T.

Solution: The columns of U_1 provide an orthonormal basis for the range of A while the columns of V_1 provide an orthonormal basis for the range of A^T.

(b) **5 pt** – Let $\hat{\underline{v}}$ denote the projection of $\underline{v} = [8\ 4\ 0\ -4]^T$ onto the range of A. Compute $\hat{\underline{v}}$.

Solution: The projection onto the range of A is given by the matrix $P_{\mathcal{R}(A)} = U_1 U_1^T$ and thus

$$\hat{\underline{v}} = P_{\mathcal{R}(A)}[8\ 4\ 0\ -4]^T = U_1 U_1^T [8\ 4\ 0\ -4]^T = U_1 [8\ 4]^T = [6\ 6\ -2\ -2]^T.$$

(c) **5 pt** – Find the minimum-norm solution $\underline{v}^*$ of the linear system $A\underline{v} = [12\ 12\ -6\ -6]^T$.

Solution: The minimum-norm solution can be found by multiplying on the left by the pseudo-inverse $A^\dagger = V_1 \Sigma_1^{-1} U_1^T$. This gives

$$\underline{v}^* = A^\dagger [12\ 12\ -6\ -6]^T = V_1 \Sigma_1^{-1} U_1^T [12\ 12\ -6\ -6]^T = V_1 \Sigma_1^{-1} [18\ 6]^T = V_1 [9\ 6]^T = [4\ -4\ 7\ 6]^T.$$

(d) **5 pt** – Using the four fundamental subspaces, give a formula based on the compact SVD for the matrix that projects orthogonally onto the nullspace of A^T.

Solution: From the four fundamental subspaces, we have $\mathcal{N}(A^T) = \mathcal{R}(A)^\perp$. Thus, the projection is given by $P_{\mathcal{N}(A^T)} = I - P_{\mathcal{R}(A)} = I - U_1 U_1^T$.

5. Let $V = \mathbb{R}^m$ be the standard inner-product space with subspaces $A, B \subseteq V$ whose intersection is $C = A \cap B$. Let the matrices P_A, P_B, P_C define orthogonal projections onto A, B, C. Starting from $\underline{v}_0 \in V$, the alternating projection algorithm generates the sequence

$$\underline{v}_{n+1} = \begin{cases} P_A \underline{v}_n & \text{if } n \text{ even} \\ P_B \underline{v}_n & \text{if } n \text{ odd.} \end{cases}$$

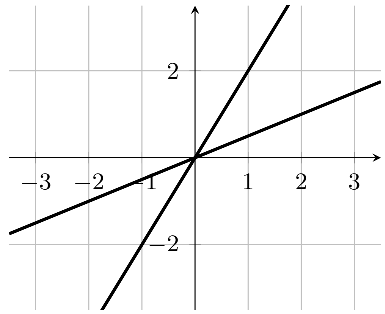
(a) **5 pt** – Suppose $m = 2$, $A = \text{span}\{\underline{a}\}$, and $B = \text{span}\{\underline{b}\}$, where $\underline{a} = (1,2)^T$ and $\underline{b} = (2,1)^T$. Draw picture illustrating these subspaces and compute the projection matrices P_A, P_B, and P_C.

Solution: Since $C = \{\underline{0}\}$, this gives

$$P_A = \frac{\underline{a}\underline{a}^T}{\|\underline{a}\|^2} = \frac{1}{5}\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$$

$$P_B = \frac{\underline{b}\underline{b}^T}{\|\underline{b}\|^2} = \frac{1}{5}\begin{bmatrix} 4 & 2 \\ 2 & 1 \end{bmatrix}$$

$$P_C = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$



(b) **5 pt** – Using the above setup, compute $\underline{v}_n$ for $n = 1, 2, 3$ starting from $\underline{v}_0 = 25\underline{b} = (50, 25)^T$.

Solution: Projecting $\underline{v}_0$ onto $\underline{a}$ gives

$$\underline{v}_1 = \frac{\langle \underline{v}_0 | \underline{a} \rangle}{\|\underline{a}\|^2}\underline{a} = 20\underline{a} = (20, 40)^T.$$

Projecting $\underline{v}_1$ onto $\underline{b}$ gives

$$\underline{v}_2 = \frac{\langle \underline{v}_1 | \underline{b} \rangle}{\|\underline{b}\|^2}\underline{b} = \frac{80}{5} = 16\underline{b} = (32, 16)^T.$$

Projecting $\underline{v}_2$ onto $\underline{a}$ gives

$$\underline{v}_3 = \frac{\langle \underline{v}_2 | \underline{a} \rangle}{\|\underline{a}\|^2}\underline{a} = \frac{64}{5}\underline{a} = \left(\frac{64}{5}, \frac{128}{5}\right)^T.$$

(c) **5 pt** – What do you observe about $\underline{v}_n$? Compare $\underline{v}_n$ to $\underline{v}_{n-1}$ and $\underline{v}_{n-2}$. How are they related?

Solution: First, we see that v_n is either a scalar mutiple of $\underline{a}$ or $\underline{b}$, depending on whether n is odd or even. Next, we observe that the coefficients are $25, 20, 16, 64/5$. Thus, the next coefficient is multiplied by $4/5$ each time.

(d) **5 pt** – For all $n \geq 0$, prove that, for some $\alpha_{n+1} \in \mathbb{R}$, we have $\underline{v}_{n+1} = \alpha_{n+1}\underline{a}$ if n is even and $\underline{v}_{n+1} = \alpha_{n+1}\underline{b}$ if n is odd. Assuming that $\underline{v}_0 = \alpha_0\underline{b}$, find a recursive formula for α_n and evaluate it numerically. Hint: Try using the formulas $P_A = \underline{a}\,\underline{a}^T/\|\underline{a}\|^2$ and $P_B = \underline{b}\,\underline{b}^T/\|\underline{b}\|^2$.

Solution: Since $\underline{v}_{n+1} = P_A\underline{v}_n$ for n even, it follows that $\underline{v}_{n+1} \in \text{span}(A)$ when n is even. Additionally, $\underline{v}_{n+1} = c_{n+1}\underline{a}$ because A is one dimensional. Similarly, since $\underline{v}_{n+1} = P_B\underline{v}_n$ for n odd, it follows that $\underline{v}_{n+1} \in \text{span}(B)$ and $\underline{v}_{n+1} = c_{n+1}\underline{b}$ when n is odd.

To find a formula, we assume that n is even and $\underline{v}_n = \alpha_n\underline{b}$. Applying the update rule gives

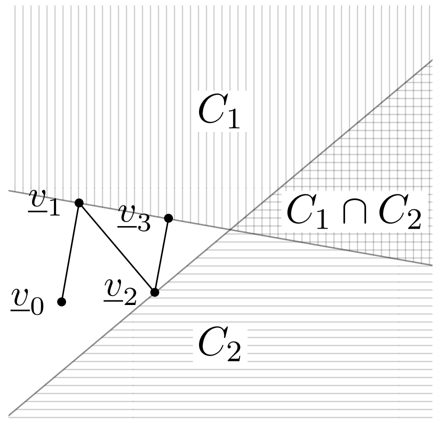
$$\underline{v}_{n+1} = P_A\underline{v}_n = \alpha_n\frac{\underline{a}\,\underline{a}^T}{\|\underline{a}\|^2}\underline{b} = \alpha_n\frac{\underline{a}^T\underline{b}}{\|\underline{a}\|^2}\underline{a} = \alpha_{n+1}\underline{a}.$$

Similarly, for n odd, we have

$$\underline{v}_{n+1} = P_B\underline{v}_n = \alpha_n\frac{\underline{b}\,\underline{b}^T}{\|\underline{b}\|^2}\underline{a} = \alpha_n\frac{\underline{b}^T\underline{a}}{\|\underline{b}\|^2}\underline{b} = \alpha_{n+1}\underline{b}.$$

Numerically, we have $\underline{a}^T\underline{b} = \underline{b}^T\underline{a} = 4$ and $\|\underline{a}\| = \|\underline{b}\| = 5$. Thus, $\alpha_{n+1} = \frac{4}{5}\alpha_n$.

(e) **5 pt** – Now, consider the alternating projection algorithm for the convex sets C_1 and C_2 below. These two half-spaces are defined by vertical/horizontal hatch lines. On the figure, draw 3 steps of alternating projection: $\underline{v}_1 = P_{C_1}(\underline{v}_0)$, $\underline{v}_2 = P_{C_2}(\underline{v}_1)$, and $\underline{v}_3 = P_{C_1}(\underline{v}_2)$ starting from $\underline{v}_0$.



6. Consider a Markov chain $X_1, X_2, \ldots$ with states $\mathcal{S} = \{1, 2, 3\}$ and transition probability matrix

$$P = \begin{bmatrix} \frac{3}{4} & 0 & \frac{1}{4} \\ \frac{1}{4} & \frac{3}{4} & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix},$$

where $P_{i,j} = \Pr(X_{t+1} = j \mid X_t = i)$ for $i, j \in \mathcal{S}$.

(a) **5 pt** – Compute the probability that $X_{t+2} = 3$ given that $X_t = 1$?

Solution: This probability is given by $[P^2]_{1,3} = \frac{3}{4}\frac{1}{4} + \frac{1}{4}\frac{1}{2} = \frac{5}{16}$.

(b) **5 pt** – This Markov chain has a unique stationary distribution $\underline{\pi} = [\pi_1\ \pi_2\ \pi_3]^T$. Find it.

Solution: The stationary distribution of the Markov chain satisfies $\underline{\pi}^T P = \underline{\pi}^T$ and $\pi_1 + \pi_2 + \pi_3 = 1$. Thus, we start by finding the null space of

$$(I - P)^T = \begin{bmatrix} \frac{1}{4} & -\frac{1}{4} & 0 \\ 0 & \frac{1}{4} & -\frac{1}{2} \\ -\frac{1}{4} & 0 & \frac{1}{2} \end{bmatrix}.$$

After row reduction, this gives

$$\begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Arbitrarily fixing $v_3 = 1$, this gives $v_2 = 2v_3 = 2$, and $v_1 = v_2 = 2$. Normalizing the sum to one gives $\underline{\pi} = \frac{1}{5}\underline{v} = (\frac{2}{5}, \frac{2}{5}, \frac{1}{5})^T$.

(c) **5 pt** – Assume this Markov chain converges to the same stationary distribution $\underline{\pi}$ from any initial state distribution $\underline{u}$ (i.e., $\|\underline{u}^T P^n - \underline{\pi}^T\| \to 0$ for some norm and all $\underline{u}$). Use this to show that the matrix sequence P^n converges to the rank-1 matrix $P^* = \underline{1}\,\underline{\pi}^T$, where $\underline{1}$ is a column vector of ones and $\underline{\pi}$ is a column vector containing the stationary distribution (i.e., each row of P^* equals the stationary distribution). Hint: Try using the standard basis for $\mathbb{R}^3$.

Solution: Following the hint, we consider $\underline{u} = \underline{e}_i$ for $i \in \{1, 2, 3\}$ and use the given convergence to see that

$$\|\underline{e}_i^T P^n - \underline{\pi}^T\| \to 0 \text{ for } i \in \{1, 2, 3\}.$$

Using this norm, we can define a metric on the space of 3×3 matrices

$$d(A, B) = ([\underline{a}_1\, \underline{a}_2\, \underline{a}_3]^T, [\underline{b}_1\, \underline{b}_2\, \underline{b}_3]^T) = \|\underline{a}_1 - \underline{b}_1\| + \|\underline{a}_2 - \underline{b}_2\| + \|\underline{a}_3 - \underline{b}_3\|,$$

where $\underline{a}_i$ and $\underline{b}_i$ are the i-th row vectors of A and B. Since $\underline{e}_i^T$ extracts the i-th row of a matrix, it follows that each row of P^n converges to $\underline{\pi}^T$. Hence,

$$d(P^n, P^*) = \|\underline{e}_1^T P^n - \underline{\pi}^T\| + \|\underline{e}_2^T P^n - \underline{\pi}^T\| + \|\underline{e}_3^T P^n - \underline{\pi}^T\| \to 0$$

and P^n converges to a matrix where all rows equal $\underline{\pi}^T$, which is precisely the definition of P^*.

(d) **5 pt** – Use induction on n to prove that $(P - \underline{1}\,\underline{\pi}^T)^n = P^n - \underline{1}\,\underline{\pi}^T$ for $n \geq 1$.
Hint: Use the facts $P\underline{1} = \underline{1}$ and $\underline{\pi}^T P = \underline{\pi}^T$.

Solution: The base case of $n = 1$ is trivial. The induction from n to $n + 1$ is given by

$$\begin{aligned}
(P - \underline{1}\,\underline{\pi}^T)^{n+1} &= (P^n - \underline{1\pi}^T)(P - \underline{1\pi}^T) \\
&= P^{n+1} - P^n \underline{1}\,\underline{\pi}^T - \underline{1}\,\underline{\pi}^T P + \underline{1}\,\underline{\pi}^T \underline{1}\,\underline{\pi}^T \\
&= P^{n+1} - 2 \cdot \underline{1}\,\underline{\pi}^T + \underline{1}\,\underline{\pi}^T \\
&= P^{n+1} - \underline{1}\,\underline{\pi}^T.
\end{aligned}$$