Practice Final Exam

November 30th, 2021

"I have adhered to the Duke Community Standard in completing this assignment."

Signature:

Name:

You may use the following list of mathematical expressions.

Theorem. Every rank-r matrix $A \in \mathbb{C}^{m\times n}$ has a singular value decomposition

$$A = U\Sigma V^H = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_1^H \\ V_2^H \end{bmatrix} = U_1 \Sigma_1 V_1^H,$$

where $U \in \mathbb{C}^{m\times m}$ and $V \in \mathbb{C}^{n\times n}$ are unitary and $U_1 \in \mathbb{C}^{m\times r}$, $U_2 \in \mathbb{C}^{m\times m-r}$, $V_1 \in \mathbb{C}^{n\times r}$, and $V_2 \in \mathbb{C}^{n\times n-r}$ have orthonormal columns. The diagonal matrix $\Sigma_1 \in \mathbb{R}^{r\times r}$ contains the non-zero singular values

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0.$$

Definition. For an $m \times n$ complex matrix A, the *pseudo-inverse* $A^\dagger$ maps any vector $\underline{w} \in \mathbb{C}^m$ to the minimum-norm vector $\underline{v}^* \in \mathbb{C}^n$ that minimizes $\|A\underline{v} - \underline{w}\|$. In general, the pseudo-inverse can be written as $A^\dagger = V_1 \Sigma_1^{-1} U_1^H$ in terms of the compact SVD. The orthogonal projection onto the range of A is given by $P_{\mathcal{R}(A)} = AA^\dagger$. For full-rank A, if $m > n$, then $A^\dagger = (A^H A)^{-1} A^H$ and, if $n > m$, then $A^\dagger = A^H (AA^H)^{-1}$.

Definition. For any $p \in [1, \infty)$ and $\underline{v} \in \mathbb{C}^m$, the *vector p-norm* is defined to be $\|\underline{v}\|_p = \left(\sum_{i=1}^m |v_i|^p\right)^{1/p}$. By continuity, $\|\underline{v}\|_\infty = \max_{i\in\{1,\ldots,m\}} |v_i|$. For a matrix $A \in \mathbb{C}^{m\times n}$, the induced *matrix p-norm* is

$$\|A\|_p \triangleq \sup_{\underline{v}\in\mathbb{R}^n : \|\underline{v}\|_p = 1} \|A\underline{v}\|_p = \begin{cases} \max_{j\in\{1,\ldots,n\}} \sum_{i=1}^m |A_{i,j}| & \text{if } p = 1 \\ \max_{i\in\{1,\ldots,m\}} \sum_{j=1}^n |A_{i,j}| & \text{if } p = \infty. \end{cases}$$

Fact. The inverse of a 2×2 matrix over a field is given in closed form by

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

Definition. For linearly independent vectors $\underline{v}_1, \ldots, \underline{v}_n$ in an inner-product space, *Gram-Schmidt* returns

$$\underline{w}_j = \underline{v}_j - \sum_{i=1}^{j-1} \frac{\langle \underline{v}_j \,|\, \underline{w}_i \rangle}{\|\underline{w}_i\|^2} \underline{w}_i, \quad j \in \{1, 2, \ldots, n\}.$$

Definition. A *Markov chain* with n states is a discrete random process $X_1, X_2, \ldots$ defined by an $n \times n$ matrix P satisfying $[P^m]_{i,j} = \Pr(X_{t+m} = j | X_t = i)$ for all $m \in \mathbb{N}$ and $i, j \in \{1, 2, \ldots, n\}$. A state distribution, when represented a column vector $\underline{\pi} = (\pi_1, \ldots, \pi_n)^T$, is called *stationary* if $\underline{\pi}^T P = \underline{\pi}^T$.

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.

   (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.

   (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 $\mathcal{N}(A) = \mathcal{R}(A^H)^\perp$.

   (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.

   (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 $\alpha f(\underline{u}) + (1-\alpha) f(\underline{v}) \leq f(\alpha \underline{u} + (1-\alpha)\underline{v})$.

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

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.

   (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 $\mathrm{span}(\underline{v}_1, \underline{v}_2)$.

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

   (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.

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 $W = \operatorname{span}(\underline{w}_1, \underline{w}_2)$.

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

   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).

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

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

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.

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

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

   (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.

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

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

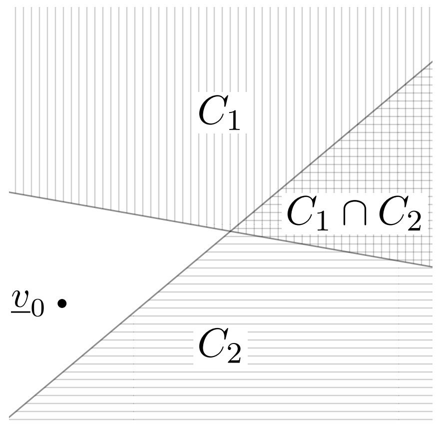
   (a) **5 pt** – Suppose $m = 2$, $A = \operatorname{span}\{\underline{a}\}$, and $B = \operatorname{span}\{\underline{b}\}$, where $\underline{a} = (1,2)^T$ and $\underline{b} = (2,1)^T$. Draw picture illustrating these subspaces and compute the projection matrices P_A, P_B, and P_C.

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

(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?

(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 $P_B = \underline{b}\,\underline{b}^T/\|\underline{b}\|^2$.

(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 $X_t = 1$?

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

(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 $\mathbb{R}^3$.

(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 $\underline{\pi}^T P = \underline{\pi}^T$.