Math 510 **Fall 2022**
Homework 2 **Drew Armstrong**

---

1. Matrix Arithmetic. You will practice matrix arithmetic by examining the formula for *block matrix inversion.* Consider a block matrix

$$P = \left(\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right),$$

where A and D are square, and where the inverse matrices A^{-1} and $(D - CA^{-1}B)^{-1}$ exist. To save notation, let's write $E = D - CA^{-1}B$. In this case we consider the block matrix

$$Q = \left(\begin{array}{c|c} A^{-1} + A^{-1}BE^{-1}CA^{-1} & -A^{-1}BE^{-1} \\ \hline -E^{-1}CA^{-1} & E^{-1} \end{array}\right).$$

Check that $PQ = I$. It is also true that $QP = I$ but please don't check this.

2. Special 2×2 Matrices. For any real number t we define the following matrices:

$$R_t = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix}, \quad F_t = \begin{pmatrix} \cos t & \sin t \\ \sin t & -\cos t \end{pmatrix}, \quad P_t = \begin{pmatrix} \cos^2 t & \cos t \sin t \\ \cos t \sin t & \sin^2 t \end{pmatrix}.$$

(a) Describe what each matrix does geometrically. [Hint: Rotate, reFlect, Project.]
(b) Check that $R_s R_t = R_{s+t}$. What does this mean geometrically?
(c) Check that $F_t^2 = I$. What does this mean geometrically?
(d) Check that $P_t^2 = P_t$. What does this mean geometrically?
(e) Check that $F_{2t} + I = 2P_t$. Draw a picture to show what this means geometrically. [For example, maybe take $t = \pi/3$ and $\mathbf{x} = (1, 0)$. Draw the line $y = \sqrt{3}x$ and the four points $\mathbf{x}$, $P_t\mathbf{x}$, $F_{2t}\mathbf{x}$, and $2P_t\mathbf{x}$.]

3. Examples of Matrix Groups. Consider the following sets of matrices:

$$\mathrm{GL}_n(\mathbb{R}) = \{\text{matrices } A \in \mathbb{R}^{n\times n} \text{ such that } A^{-1} \text{ exists}\},$$
$$\mathrm{O}_n(\mathbb{R}) = \{\text{matrices } A \in \mathbb{R}^{n\times n} \text{ such that } A^{-1} = A^T\}.$$

(a) Check that each of these sets is a *group*. That is, it contains the identity matrix, it is closed under under taking inverses, and it is closed under taking products.
(b) The equation $A^TA = I$ tells us that the columns of A are an orthonormal set of vectors. Use this fact to show that every matrix in $\mathrm{O}_2(\mathbb{R})$ is equal to R_t or F_t from Problem 2. [Hint: Since the first column has length 1 it equals $(\cos t, \sin t)$ for some angle t. The second column must be a unit vector that is perpendicular to the first column.]

4. Frobenius Norm. For any complex matrix $A = (a_{ij})$ we define the Frobenius norm:

$$\|A\|_F := \sqrt{\sum_{i,j} |a_{ij}|^2}.$$

We already know that $\|\cdot\|_F$ is a norm on the vector space $\mathbb{C}^{m\times n}$ of $m \times n$ matrices under addition and scalar multiplication. In this problem you will show that $\|AB\|_F \leq \|A\|_F\|B\|_F$ for any matrices A, B where the product is AB defined.

(a) If $\mathbf{a}_1, \ldots, \mathbf{a}_m \in \mathbb{R}^\ell$ are the columns of $A \in \mathbb{C}^{\ell\times m}$, show that

$$\|A\|_F = \sqrt{\|\mathbf{a}_1\|_F^2 + \cdots + \|\mathbf{a}_m\|_F^2}.$$

(b) For any column vectors $\mathbf{a}, \mathbf{b} \in \mathbb{C}^\ell$, show that $\|\mathbf{a}\mathbf{b}^T\|_F = \|\mathbf{a}\|_F \|\mathbf{b}\|_F$.

(c) For any real numbers $x_1, \ldots, x_m$ and $y_1, \ldots, y_m$ use Cauchy-Schwarz to show that

$$x_1 y_1 + \cdots + x_m y_m \leq \sqrt{x_1^2 + \cdots + x_m^2} \cdot \sqrt{y_1^2 + \cdots y_m^2}.$$

(d) Let $A \in \mathbb{C}^{\ell \times m}$ have column vectors $\mathbf{a}_1, \ldots, \mathbf{a}_m \in \mathbb{C}^\ell$ and let $B \in \mathbb{C}^{m \times n}$ have row vectors $\mathbf{b}_1^T, \ldots, \mathbf{b}_m^T \in \mathbb{C}^n$. Combine (abc) with the usual triangle inequality to show that $\|AB\|_F \leq \|A\|_F \|B\|_F$. Hint: Apply $\|\cdot\|_F$ to both sides of the formula

$$AB = \mathbf{a}_1 \mathbf{b}_1^T + \cdots + \mathbf{a}_m \mathbf{b}_m^T.$$

5. Geometric Series of Matrices (Optional). Let A be a square matrix with $\|A\|_F < 1$. In this problem you will show that $I - A$ is invertible, with a power series expansion that converges with respect to the Frobenius norm:[1]

$$(I - A)^{-1} = I + A^2 + A^3 + \cdots = \sum_{k \geq 0} A^k.$$

(a) Show that $\|A^n\|_F \leq \|A\|_F^n$. Use this to show that A^n converges to the zero matrix.

(b) Let $S_n = \sum_{k=0}^n A^k$, and show that $\|S_n\|_F \leq \sum_{k=0}^n \|A\|_F^k$. Then the usual geometric series implies that S_n is a Cauchy sequence, hence S_n converges to some matrix T.

(c) Observe that $(I - A)S_n = I - I - A^{n+1}$. Use (a) to show that the right side converges to I and use (b) to show that the left side converges to $(I - A)T$. Hence $(I - A)T = I$.

(d) **Application.** Consider a partitioned matrix

$$P = \left(\begin{array}{c|c} I & R \\ \hline 0 & Q \end{array}\right),$$

where I is an identity matrix, R is any rectangular matrix and Q is a square matrix satisfying $\|Q\|_F < 1$. Use the geometric series for matrices to show that

$$P^n \to \left(\begin{array}{c|c} I & R(I - Q)^{-1} \\ \hline 0 & 0 \end{array}\right) \qquad \text{as } n \to \infty.$$

[Hint: Compute the first few powers of P and observe a pattern.]

[1] To be precise, for any sequence of matrices $X_1, X_2, \ldots$ and for any matrix Y, we say that X_n converges to Y if and only if $\|X_n - Y\|_F$ converges to the number zero. It follows from the completeness of the complex numbers that if $\|X_n - X_m\|_F$ gets arbitrarily small for n and m arbitrarily large (i.e., if X_n is a *Cauchy sequence*) then there exists some matrix Y such that $X_n \to Y$.