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.

Solution. We want to show that

$$PQ = \left(\begin{array}{c|c} I & O \\ \hline O & I \end{array}\right),$$

where the blocks on the diagonal are identity matrices with the same sizes as A and D. First we observe that the top left entry of PQ is

$$\begin{aligned} &A(A^{-1} + A^{-1}BE^{-1}CA^{-1}) + B(-E^{-1}CA^{-1}) \\ &\quad = AA^{-1} + AA^{-1}BE^{-1}CA^{-1} + B(-E^{-1}CA^{-1}) \\ &\quad = I + IBE^{-1}CA^{-1} - BE^{-1}CA^{-1} \\ &\quad = I + BE^{-1}CA^{-1} - BE^{-1}CA^{-1} \\ &\quad = I + O \\ &\quad = I. \end{aligned}$$

The bottom left entry of PQ is

$$\begin{aligned} &C(A^{-1} + A^{-1}BE^{-1}CA^{-1}) + D(-E^{-1}CA^{-1}) \\ &\quad = (CA^{-1}) + CA^{-1}BE^{-1}(CA^{-1}) - DE^{-1}(CA^{-1}) \\ &\quad = (I + CA^{-1}BE^{-1} - DE^{-1})(CA^{-1}) \\ &\quad = (I + (CA^{-1}B - D)E^{-1})(CA^{-1}) \\ &\quad = (I + (-E)E^{-1})(CA^{-1}) \\ &\quad = (I - I)(CA^{-1}) \\ &\quad = O(CA^{-1}) \\ &\quad = O. \end{aligned}$$

The top right entry is

$$\begin{aligned} &A(-A^{-1}BE^{-1}) + BE^{-1} \\ &\quad = -(AA^{-1})BE^{-1} + BE^{-1} \\ &\quad = -BE^{-1} + BE^{-1} \\ &\quad = O. \end{aligned}$$

And the bottom right entry is

$$\begin{aligned} C(-A^{-1}BE^{-1}) &+ DE^{-1} \\ &= -CA^{-1}BE^{-1} + DE^{-1} \\ &= (-CA^{-1}B + D)E^{-1} \\ &= EE^{-1} \\ &= I. \end{aligned}$$

Done.

Remark: According to Wikipedia this formula was reinvented many times and is due to Hans Boltz (1923). Suppose that A, B, C, D are just 1×1 scalars a, b, c, d, with $e = d - bc/a$. Then the block inverse formula becomes the familiar formula for the inverse of a 2×2 matrix:

$$\begin{aligned} \begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} &= \begin{pmatrix} 1/a + bc/a^2e & -b/ae \\ -c/ae & 1/e \end{pmatrix} \\ &= \begin{pmatrix} (ae+bc)/a^2e & -b/(ad-bc) \\ -c/(ad-bc) & a/(ad-bc) \end{pmatrix} \\ &= \begin{pmatrix} (ad)/a(ad-bc) & -b/(ad-bc) \\ -c/(ad-bc) & a/(ad-bc) \end{pmatrix} \\ &= \begin{pmatrix} d/(ad-bc) & -b/(ad-bc) \\ -c/(ad-bc) & a/(ad-bc) \end{pmatrix} \\ &= \frac{1}{ad-bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}. \end{aligned}$$

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 $2P_t\mathbf{x}$.]

(a): See the course notes for a discussion of the geometry.

(b): Using the angle sum formulas gives

$$\begin{aligned} R_s R_t &= \begin{pmatrix} \cos s & -\sin s \\ \sin s & \cos s \end{pmatrix} \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix} \\ &= \begin{pmatrix} \cos s \cos t - \sin s \sin t & -\cos s \sin t - \sin s \cos t \\ \sin t \cos s + \cos s \sin t & -\sin s \sin t + \cos s \cos t \end{pmatrix} \\ &= \begin{pmatrix} \cos(s+t) & -\sin(s+t) \\ \sin(s+t) & \cos(s+t) \end{pmatrix} \\ &= R_{s+t}. \end{aligned}$$

Geometric meaning: Multiplication of matrices is the same as composition of linear functions. Rotating by angle t and then by angle s is the same as rotating once by angle $s+t$. We see from these arguments that rotation matrices commute:

$$R_s R_t = R_{s+t} = R_t R_s.$$

(b): We have

$$\begin{aligned}
F_t^2 &= \begin{pmatrix} \cos t & \sin t \\ \sin t & -\cos t \end{pmatrix} \begin{pmatrix} \cos t & \sin t \\ \sin t & -\cos t \end{pmatrix} \\
&= \begin{pmatrix} \cos^2 t + \sin^2 t & \cos t \sin t - \sin t \cos t \\ \sin t \cos t - \cos t \sin t & \sin^2 t + \cos^2 t \end{pmatrix} \\
&= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\
&= I.
\end{aligned}$$

Geometric meaning: Reflecting twice is the same as doing nothing.

(c): We have

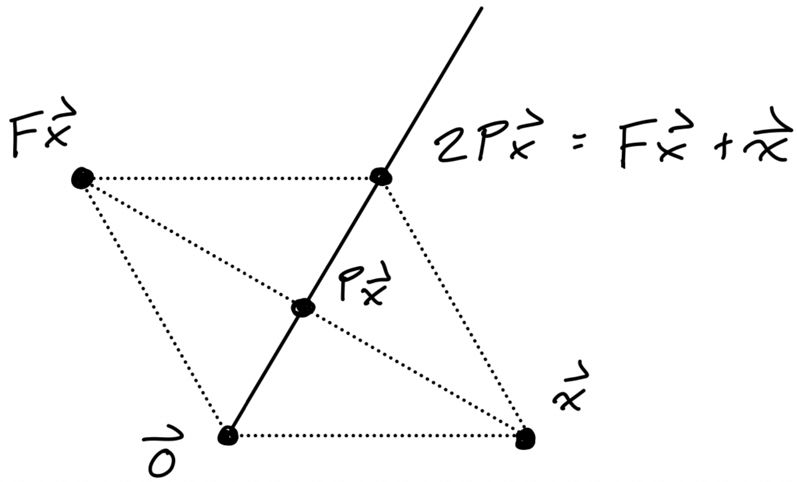
$$\begin{aligned}
P_t^2 &= \begin{pmatrix} \cos^2 t & \cos t \sin t \\ \cos t \sin t & \sin^2 t \end{pmatrix} \begin{pmatrix} \cos^2 t & \cos t \sin t \\ \cos t \sin t & \sin^2 t \end{pmatrix} \\
&= \begin{pmatrix} \cos^4 t + \cos^2 t \sin^2 t & \cos^3 t \sin t + \cos t \sin^3 t \\ \cos^3 t \sin t + \cos t \sin^3 t & \cos^2 t \sin^2 t + \sin^4 t \end{pmatrix} \\
&= \begin{pmatrix} \cos^2 t(\cos^2 t + \sin^2 t) & \cos t \sin t(\cos^2 t + \sin^2 t) \\ \cos t \sin t(\cos^2 t + \sin^2 t) & \sin^2 t(\cos^2 t + \sin^2 t) \end{pmatrix} \\
&= \begin{pmatrix} \cos^2 t & \cos t \sin t \\ \cos t \sin t & \sin^2 t \end{pmatrix} \\
&= P_t.
\end{aligned}$$

Geometric meaning: Projecting twice is the same as projecting once.

(d): Using the double angle formulas gives

$$\begin{aligned}
F_{2t} + I &= \begin{pmatrix} \cos(2t) & \sin(2t) \\ \sin(2t) & -\cos(2t) \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\
&= \begin{pmatrix} \cos(2t) + 1 & \sin(2t) \\ \sin(2t) & -\cos(2t) + 1 \end{pmatrix} \\
&= \begin{pmatrix} 2\cos^2 t & 2\cos t \sin t \\ 2\cos t \sin t & 2\sin^2 t \end{pmatrix} \\
&= 2\begin{pmatrix} \cos^2 t & \cos t \sin t \\ \cos t \sin t & \sin^2 t \end{pmatrix} \\
&= 2P_t.
\end{aligned}$$

Geometric meaning: This identity expresses the relationship between projection and reflection. Recall that F_{2t} (resp. P_t) are the reflection across (resp. projection onto) the line in $\mathbb{R}^2$ with positive angle t from the x-axis. Here is a picture:



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 $A^T A = I$ 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.]

(a): The identity matrix is invertible with $I^{-1} = I$. If A^{-1} exists then $(A^{-1})^{-1}$ exists (and is equal to A). If A^{-1} and B^{-1} exist then we saw in class that $(AB)^{-1}$ exists (and is equal to $B^{-1}A^{-1}$). Thus we have shown that $\mathrm{GL}_n(\mathbb{R})$ is a group, called the *general linear group*.

A matrix satisfying $A^{-1} = A^T$ is called *orthogonal*. The identity matrix is orthogonal because $I^{-1} = I = I^T$. If A is orthogonal then so is A^{-1} because

$$(A^{-1})^{-1} = A = (A^T)^T = (A^{-1})^T.$$

If A and B are orthogonal then so is AB because

$$(AB)^{-1} = B^{-1}A^{-1} = B^T A^T = (AB)^T.$$

Thus we have shown that $\mathrm{O}_n(\mathbb{R})$ is a group, called the *orthogonal group*.

(b): The condition $A^{-1} = A^T$ is equivalent to $A^T A = AA^T = I$.[1] The statement $A^T A = I$ says that A has orthonormal columns and the statement $AA^T = I$ says that A has orthonormal rows. Indeed, let $\mathbf{a}_1, \ldots, \mathbf{a}_n$ be the columns of A then we have

$$\begin{aligned}(ij \text{ entry of } A^T A) &= (i\text{th row of } A^T)(j\text{th col of } A)\\ &= \mathbf{a}_i^T \mathbf{a}_j\end{aligned}$$

[1] And it follows from the Fundamental Theorem that the statements $A^T A = I$ and $AA^T = I$ are equivalent to each other.

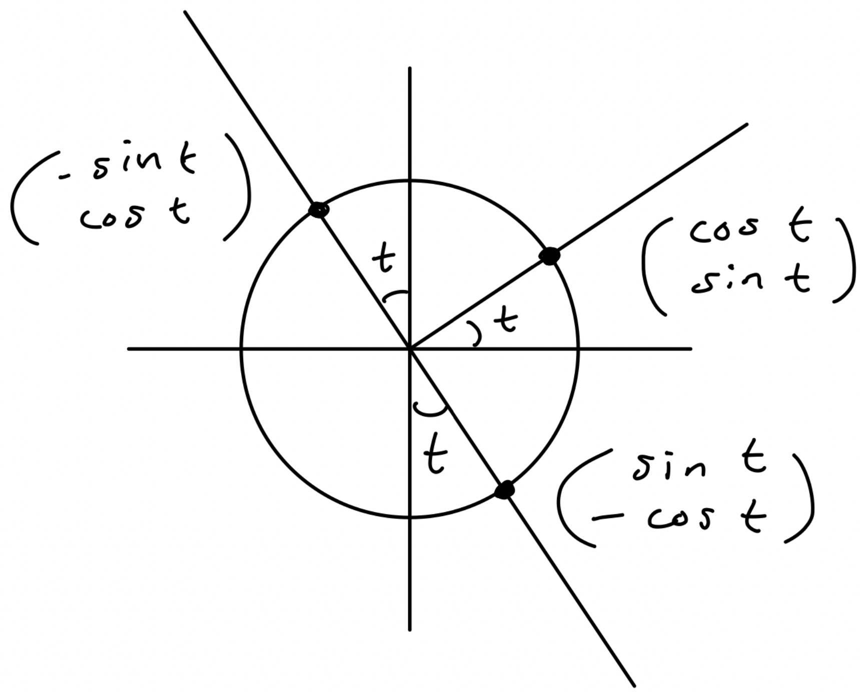
$$= (\text{dot product of } \mathbf{a}_i \text{ and } \mathbf{a}_j).$$

But the ij entry of the identity matrix is the Kronecker delta δ_{ij}. Hence if $A^T A = I$ then

$$(\text{dot product of } \mathbf{a}_i \text{ and } \mathbf{a}_j) = \delta_{ij}$$

and we see that the columns of A are orthonormal.

Now let $A \in \mathrm{O}_2(\mathbb{R})$ with columns $\mathbf{a}_1, \mathbf{a}_2 \in \mathbb{R}^2$. Since $\|\mathbf{a}_1\|^2$ we can write $\mathbf{a}_1 = (\cos t, \sin t)$ for some unique angle $t \in [0, 2\pi)$. Then since $\mathbf{a}_1 \bullet \mathbf{a}_2 = 0$ and $\|\mathbf{a}_2\| = 1$ we must have $\mathbf{a}_2 = (-\sin t, \cos t)$ or $\mathbf{a}_2 = (\sin t, -\cos t)$. Picture:



Hence the group $\mathrm{O}_2(\mathbb{R})$ consists of rotations and reflections. Remark: Continuing from Problem 2, one can check the identities:

- $R_s R_t = R_{s+t}$,
- $R_s F_t = F_{s+t}$,
- $F_s R_t = F_{s-t}$,
- $F_s F_t = R_{s-t}$.

The product of two rotations is a rotation. The product of a rotation and a reflection is a reflection. The product of two reflections is a rotation. Rotations commute with each other, but they don't commute with reflections. Reflections do not commute with each other.

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

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

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

(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

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

(a): The jth column of A is $\mathbf{a}_j = (a_{1j}, a_{2j}, \ldots, a_{\ell j})$, hence

$$\|\mathbf{a}_j\|_F^2 = \|\mathbf{a}_j\|^2 = |a_{1j}|^2 + \cdots + |a_{\ell j}|^2 = \sum_{i=1}^{\ell} |a_{ij}|^2.$$

Then the Frobenius norm of A satisfies

$$\|A\|_F^2 = \sum_{i,j} |a_{ij}|^2 = \sum_{i=1}^{\ell} \sum_{j=1}^{m} |a_{ij}|^2 = \sum_{j=1}^{m} \left(\sum_{i=1}^{\ell} |a_{ij}|^2 \right) = \sum_{j=1}^{m} \|\mathbf{a}_j\|^2.$$

(b): For any column vectors $\mathbf{a} = (a_1, \ldots, a_\ell)$ and $\mathbf{b} = (b_1, \ldots, b_\ell)$ note that the ij entry of the $\ell \times \ell$ matrix $\mathbf{a}\mathbf{b}^T$ is $a_i b_j$. Recall that the absolute value of complex numbers is multiplicative: $|a_i b_j| = |a_i||b_j|$. Hence we have

$$\begin{aligned}
\|\mathbf{a}\mathbf{b}^T\|_F^2 &= \sum_{i,j} |a_i b_j|^2 \\
&= \sum_{i,j} |a_i|^2 |b_j|^2 \\
&= \left(\sum_i |a_i|^2 \right) \left(\sum_j |b_j|^2 \right) \\
&= \|\mathbf{a}\|_F^2 \|\mathbf{b}\|_F^2.
\end{aligned}$$

(c): Given **real** vectors $\mathbf{x} = (x_1, \ldots, x_m)$ and $\mathbf{y} = (y_1, \ldots, y_m)$, the Cauchy-Schwarz inequality for the dot product in $\mathbb{R}^m$ tells us that

$$\begin{aligned}
|\mathbf{x} \bullet \mathbf{y}|^2 &\leq (\mathbf{x} \bullet \mathbf{x})(\mathbf{y} \bullet \mathbf{y}), \\
|\mathbf{x} \bullet \mathbf{y}| &\leq \sqrt{\mathbf{x} \bullet \mathbf{x}} \sqrt{\mathbf{y} \bullet \mathbf{y}}, \\
\mathbf{x} \bullet \mathbf{y} &\leq \sqrt{\mathbf{x} \bullet \mathbf{x}} \sqrt{\mathbf{y} \bullet \mathbf{y}}, \\
x_1 y_1 + \cdots + x_m y_m &\leq \sqrt{x_1^2 + \cdots + x_m^2} \cdot \sqrt{y_1^2 + \cdots y_m^2}.
\end{aligned}$$

(d): Finally, 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 $\mathbf{b}_1^T, \ldots, \mathbf{b}_m^T \in \mathbb{C}^n$. From the definition of matrix multiplication we have

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

Now we apply the Frobenius norm to both sides:

$$\begin{aligned}
\|AB\|_F &= \|\mathbf{a}_1\mathbf{b}_1^T + \cdots + \mathbf{a}_m\mathbf{b}_m^T\|_F \\
&\leq \|\mathbf{a}_1\mathbf{b}_1^T\|_F + \cdots + \|\mathbf{a}_m\mathbf{b}_m^T\|_F && \text{triangle inequality} \\
&= \|\mathbf{a}_1\|\|\mathbf{b}_1\| + \cdots + \|\mathbf{a}_m\|\|\mathbf{b}_m\| && \text{part (b)} \\
&\leq \sqrt{\|\mathbf{a}_1\|^2 + \cdots + \|\mathbf{a}_m\|^2}\sqrt{\|\mathbf{b}_1\|^2 + \cdots + \|\mathbf{b}_m\|^2} && \text{part (c)} \\
&= \|A\|_F\|B\|_F. && \text{part (a)}
\end{aligned}$$

□

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:[2]

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

(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 $(I - A)S_n = I - A^{n+1}$. 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

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

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

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

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

(a): It follows directly from Problem 4(d) that $0 \leq \|A^n\|_F \leq \|A\|_F^n$ for any square matrix. If $\|A\|_F < 1$ then this implies that $\|A\|_F^n \to 0$ and hence $\|A^n - O\|_F = \|A^n\|_F \to 0$ as $n \to \infty$. **By definition** this means that $A^n \to O$ with respect to the Frobenius norm.[3]

(b): I guess I should have asked you to prove that S_n is a Cauchy sequence with respect to the Frobenius norm. So consider integers $0 < N \leq m < n$ and let $\lambda := \|A\|_F$. Assuming $\lambda \neq 1$ we have[4]

$$0 \leq \|S_n - S_m\|_F = \left\|\sum_{k=m+1}^n A^k\right\| \leq \sum_{k=m+1}^n \|A\|_F^k = \sum_{k=m+1}^n \lambda^k = \frac{\lambda^{m+1} - \lambda^{n+1}}{1 - \lambda}.$$

[2]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 $X_n \to Y$.

[3]In fact, $A^n \to O$ with respect to any norm on $\mathbb{C}^{n\times n}$ because any two norms on a finite dimensional vector space are topologically equivalent. Never mind what that means.

[4]The last equality uses $\sum_{k=0}^m \lambda^k = (\lambda^{m+1} - 1)/(\lambda - 1)$ and $\sum_{k=0}^n \lambda^k = (\lambda^{n+1} - 1)/(\lambda - 1)$.

If $0 \le \lambda < 1$ then since $n - m \ge 1$ we have $0 \le \lambda^{n-m} < 1$ and hence $0 < 1 - \lambda^{n-m} \le 1$. But we also have $0 < 1 - \lambda$ and $0 < \lambda^{m+1}/(1-\lambda)$, hence

$$0 \le \frac{\lambda^{m+1} - \lambda^{n+1}}{1-\lambda} = \frac{\lambda^{m+1}}{1-\lambda}(1 - \lambda^{n-m}) \le \frac{\lambda^{m+1}}{1-\lambda}.$$

Since $\lambda^{m+1} \to 0$ as $m \to \infty$, it follows that $(\lambda^{m+1} - \lambda^{n+1})/(1-\lambda) \to 0$ as $N \to \infty$.[5] We have shown that S_n is a Cauchy sequence with respect to the Frobenius norm. By the completeness of Hermitian space it follows that S_n converges to some matrix T.

(c): We have a straightforward algebraic identity:

$$(I - A)S_n = \sum_{k=0}^{n} A^k - \sum_{k=1}^{n+1} A^k = I - A^{n+1}.$$

Since $A^{n+1} \to O$, the right hand side converges to I. To be precise, we have $\|I - (I - A^{n+1})\|_F = \|A^{n+1}\|_F \to 0$, so $I - A^{n+1} \to I$ **by definition** of convergence with respect to the Frobenius norm. On the other hand, since S_n converges to T, the left hand side converges to $(I - A)T$. To be precise, since $\|S_n - T\|_F \to 0$ **by definition**, we have

$$0 \le \|(I - A)S_n - (I - A)T\|_F = \|(I - A)(S_n - T)\|_F \le \|I - A\|_F \|S_n - T\|_F \to 0.$$

Thus we have shown that $(I - A)T = I$. We could also show that $T(I - A) = I$ by using the identity $S_n(I - A) = I - A^{n+1}$, but this is not necessary because $(I - A)T = I$ implies $T(I - A) = I$ by the Fundamental Theorem. $\square$

Remark: As I mentioned above, the Frobenius norm is not special. Essentially the same proof works for any matrix norm $\| - \|$ satisfying $\|A\| < 1$. The furthest we can push this is to say that the geometric series $\sum_k A^k$ converges if and only if the maximum of the absolute values of the eigenvalues of A is < 1. That result is harder to prove and requires something like the Jordan Canonical Form.

Reminder: Problem 5 was **optional**.

(d): **Application.** Let Q be a square matrix with $\|Q\|_F < 1$ and consider the matrix

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

By induction we can show that

$$P^n = \left(\begin{array}{c|c} I & R(I + Q + \cdots + Q^{n-1}) \\ \hline O & Q^n \end{array}\right),$$

whence

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

In class we will apply this result to Markov chains.

[5] Isn't analysis fun?