# Math 5610/6860: Assignment 1

Introduction to Numerical Analysis-1, Fall 2022

Due date: September, 26

*Instructions:*

1. *You are free to discuss the homework questions with your classmates. Sharing of solutions is not allowed.*
2. *Write your own codes, do not copy.*
3. *Submit your code files separately, but discuss the results in your main solution file.*
4. *Show the work done to obtain answers/solutions.*
5. *Questions marked with [$G^*$] are required for students registered for* 6860. *Others registered for* 5610 *need not answer the question.*

## Problem 1 (12 points)

Two norms $\|\cdot\|_p$ and $\|\cdot\|_q$ on a vector space $V$ are said to be *equivalent* if there exist two positive constants $c$ and $C$ such that,

$$c\,\|\boldsymbol{x}\|_p \leq \|\boldsymbol{x}\|_q \leq C\,\|\boldsymbol{x}\|_p \quad \text{for all } x \in V \tag{1}$$

If the vector space $V$ is $\mathbb{R}^n$ show that, the norms $\|\cdot\|_\infty$, $\|\cdot\|_1$ and $\|\cdot\|_2$ are equivalent.
More precisely, show that, the following inequalities hold:

a) $\|\boldsymbol{x}\|_\infty \leq \|\boldsymbol{x}\|_2 \leq \sqrt{n}\,\|\boldsymbol{x}\|_\infty$ for all $\boldsymbol{x} \in \mathbb{R}^n$

b) $\dfrac{\|\boldsymbol{x}\|_1}{\sqrt{n}} \leq \|\boldsymbol{x}\|_2 \leq \|\boldsymbol{x}\|_1$ for all $\boldsymbol{x} \in \mathbb{R}^n$

## Problem 2: (13 points)

An example of induced matrix norm is the *p-norm* of a matrix defined as,

$$\|\boldsymbol{A}\|_p := \sup_{\|\boldsymbol{x}\|_p \neq 0} \frac{\|\boldsymbol{A}\boldsymbol{x}\|_p}{\|\boldsymbol{x}\|_p} \tag{2}$$

If $\sigma_1$ is the largest singular value of $\boldsymbol{A}$ then show that,

a) $\|\boldsymbol{A}\|_2 = \sigma_1$

b) If $\boldsymbol{A}$ is hermitian (or real and symmetric) then $\|\boldsymbol{A}\|_2 = \rho(A)$

c) If $\boldsymbol{A}$ is unitary then $\|\boldsymbol{A}\|_2 = 1$

Also, in each case find the relative condition number $K_2(\boldsymbol{A})$

## Problem 3: (15 points)

Let $\boldsymbol{A}$ be a $6 \times 6$ matrix and $\boldsymbol{b}$ be a $6 \times 1$ vector. Partition the matrix $\boldsymbol{A}$ into blocks of $4$ $3 \times 3$ matrices $\boldsymbol{C}, \boldsymbol{D}, \boldsymbol{E}, \boldsymbol{F}$ and the vector $\boldsymbol{b}$ into $2$ $3 \times 1$ blocks $\boldsymbol{g}, \boldsymbol{h}$.

$$A = \begin{bmatrix} 6 & 4 & 5 & 9 & 4 & 3 \\ 3 & 7 & 5 & 8 & 10 & 4 \\ 8 & 8 & 5 & 7 & 9 & 5 \\ 2 & 1 & 4 & 4 & 6 & 3 \\ 7 & 10 & 6 & 9 & 7 & 9 \\ 2 & 8 & 6 & 6 & 6 & 2 \end{bmatrix} \tag{3}$$

$$\boldsymbol{b} = \begin{bmatrix} 103 \\ 138 \\ 142 \\ 80 \\ 170 \\ 102 \end{bmatrix} = \left[\frac{g}{h}\right] \tag{4}$$

$$A = \left(\begin{array}{c|c} C & D \\ \hline E & F \end{array}\right)$$

a) Obtain the vector $\boldsymbol{p}$ by direct matrix-vector multiplication $\boldsymbol{A}\boldsymbol{b} = \boldsymbol{p}$ and verify that the same result is obtained using block partitioned matrices

$$\left[\begin{array}{c} Cg + Dh \\ \hline Eg + Fh \end{array}\right] \equiv \boldsymbol{p} \tag{5}$$

b) Now, consider a linear system given by $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{p}$. Partition the matrix $\boldsymbol{A}$ as above, and partition $\boldsymbol{p}$ and $\boldsymbol{x}$ as

$$\boldsymbol{p} = \left[\begin{array}{c} g \\ \hline h \end{array}\right], \quad \boldsymbol{x} = \left[\begin{array}{c} y \\ \hline z \end{array}\right] \tag{6}$$

Find the expressions for $\boldsymbol{y}$ and $\boldsymbol{z}$ in terms of the partitioned block matrices. Compute the solutions for $\boldsymbol{y}$ and $\boldsymbol{z}$ and verify with the solution obtained using $\boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{p}$

*(Note: Solving problems with large coefficient matrices of sizes in millions× millions can be challenging when solved serially. Block partitioning can speed up the evaluation and allow to employ multiple cores. in a parallel computing environment.)*

## Problem 4: (40 points)

Consider the linear system $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ with $\boldsymbol{A} \in \mathbb{R}^{n\times n}$, $\boldsymbol{b} \in \mathbb{R}^n$ and $\boldsymbol{x} \in \mathbb{R}^n$

$$\begin{aligned} \boldsymbol{A} \equiv a_{ij} &= \frac{1}{i+j-1} \quad \text{for all } i,j = 1,...,n \\ \boldsymbol{b} \equiv b_i &= \sum_{j=1}^{n} \frac{1}{i+j-1} \end{aligned} \tag{7}$$

a) For $n = 2$, use GEM to find the solution $\boldsymbol{x}$

b) Write a code that uses GEM to find the solution $\hat{\boldsymbol{x}}$. Verify the code with your solution from a). Compute $\boldsymbol{x}$ for 5 cases, $n = 5, 8, 10, 12, 15$.

c) In each case ($n = 5, 8, 10, 12$), find the condition number of the matrix $\boldsymbol{A}$ and using the relation (9) comment on the singular behavior of the $\boldsymbol{A}$ matrix in each case.

**Definition:** The relative distance of a matrix $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ from the set of singular matrices with respect to the p-norm is given as,

$$dist_p(\boldsymbol{A}) = \min\left\{\frac{\|\delta\boldsymbol{A}\|_p}{\|\boldsymbol{A}\|_p} : \boldsymbol{A} + \delta\boldsymbol{A} \text{is singular}\right\}, \tag{8}$$

and it can be shown that,

$$dist_p(\boldsymbol{A}) = \frac{1}{K(\boldsymbol{A})} \tag{9}$$

d) Now, compute solutions for the cases $n = 5, 6, \ldots, 15$. Given, that the exact solution in each case is $\boldsymbol{x} = [1, \ldots, 1]^T_{n\times 1}$.

Compute the 2-norm of the error for all $n$ values.

Plot the "error vs. n" curve. and comment on the observed behavior of error and the possible reasons for it?

## Problem 5: (10 points)

a) Show that given any consistent matrix norm $\|\cdot\|$, then $\rho(\boldsymbol{A}) \leq \|\boldsymbol{A}\|$ for all $\boldsymbol{A} \in \mathbb{R}^{n\times n}$.

b) The following property holds for consistent matrix norms: Given any matrix $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ and $\epsilon > 0$, then there exists a consistent matrix norm $\|\cdot\|_\epsilon$ such that,

$$\|\boldsymbol{A}\|_\epsilon \leq \rho(\boldsymbol{A}) + \epsilon \tag{10}$$

Combining your proof along with this property we can state that,

$$\rho(\boldsymbol{A}) = \min_{\|\cdot\|} \|\boldsymbol{A}\| \tag{11}$$

Using (11) show that if $\boldsymbol{A}$ is a square matrix then,

$$\lim_{k\to\infty} \boldsymbol{A}^k = 0 \iff \rho(\boldsymbol{A}) < 1 \tag{12}$$

c) [$\mathbf{G}^*$] Continuing with the result you have proved in b), the geometric series $\sum_{k=0}^{\infty} \boldsymbol{A}^k$ is convergent iff $\rho(\boldsymbol{A}) < 1$. In this case show that,

$$\sum_{k=0}^{\infty} \boldsymbol{A}^k = (\boldsymbol{I} - \boldsymbol{A})^{-1} \tag{13}$$

d) [Bonus] Show that:

$$\frac{1}{1+\|\boldsymbol{A}\|} \leq \left\|(\boldsymbol{I}-\boldsymbol{A})^{-1}\right\| \leq \frac{1}{1-\|\boldsymbol{A}\|} \tag{14}$$

where $\|\cdot\|$ is an induced matrix norm such that $\|\boldsymbol{A}\| < 1$

*(Note: The results of this proof are useful in deriving the bounds on the change in solution when the system $\boldsymbol{Ax} = \boldsymbol{b}$ is perturbed to $(\boldsymbol{A} + \delta\boldsymbol{A})\boldsymbol{x} = \boldsymbol{b} + \delta\boldsymbol{b}$. The corresponding bound was discussed in class and you can find it in the lecture notes.)*

## Problem 6: (10 points)

Given 2 matrices, $\boldsymbol{A} = \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix}$ and $\boldsymbol{B} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. Find whether they have a unique LU factorization. In case of unique factorization, show that $\det \boldsymbol{A} = \det \boldsymbol{U}$

# Homework #1:

**Problem 1 (12 points)**

Two norms $\|\cdot\|_p$ and $\|\cdot\|_q$ on a vector space $V$ are said to be *equivalent* if there exist two positive constants $c$ and $C$ such that,

$$c\,\|\boldsymbol{x}\|_p \leq \|\boldsymbol{x}\|_q \leq C\,\|\boldsymbol{x}\|_p \quad \text{for all } x \in V \tag{1}$$

If the vector space $V$ is $\mathbb{R}^n$ show that, the norms $\|\cdot\|_\infty$, $\|\cdot\|_1$ and $\|\cdot\|_2$ are equivalent.
More precisely, show that, the following inequalities hold:

a) $\|\boldsymbol{x}\|_\infty \leq \|\boldsymbol{x}\|_2 \leq \sqrt{n}\,\|\boldsymbol{x}\|_\infty$ for all $\boldsymbol{x} \in \mathbb{R}^n$

b) $\dfrac{\|\boldsymbol{x}\|_1}{\sqrt{n}} \leq \|\boldsymbol{x}\|_2 \leq \|\boldsymbol{x}\|_1$ for all $\boldsymbol{x} \in \mathbb{R}^n$

a)

$$\|x\|_2 = \sum_{i=1}^{n} |x_i|^2 = x_1^2 + x_2^2 + x_3^2 + \cdots + x_n^2$$

$$\|x\|_\infty = \max|x| = x_{max}$$

Show $\|x\|_\infty \leq \|x_2\|$:

$$x_{max} \leq \sqrt{x_1^2 + x_2^2 + \cdots + x_{max}^2 + \cdots + x_n^2}$$

$$x_{max}^2 \leq x_1^2 + x_2^2 + \cdots + x_{max}^2 + \cdots x_n^2$$

$$0 \leq x_1^2 + x_2^2 + \cdots + x_n^2$$

if $x = \emptyset$ then $\|x\|_\infty = \|x_2\|$
and if $x \neq \emptyset$ then $\|x\|_\infty < \|x_2\|$ ✓

Show $\|x_2\| \leq \sqrt{n}\,\|x\|_\infty$:

$$\sqrt{x_1^2 + x_2^2 + \cdots + x_{max}^2 + \cdots + x_n^2} \leq \sqrt{n}\, x_{max}$$

$$x_1^2 + x_2^2 + \cdots + x_{max}^2 + \cdots + x_n^2 \leq (\sqrt{n}\, x_{max})^2 = n\, x_{max}^2$$

$$\frac{x_1^2 + x_2^2 + \cdots + x_{max}^2 + \cdots + x_n^2}{x_{max}^2} \leq n$$

$$\frac{x_1^2}{x_{max}^2} + \frac{x_2^2}{x_{max}^2} + \cdots + \frac{x_{max}^2}{x_{max}^2} + \cdots + \frac{x_n^2}{x_{max}^2} \leq n$$

$$0 \leq \frac{x_i^2}{x_{max}^2} \leq 1 \quad \text{so} \quad \sum_{i=1}^{n} \frac{x_i^2}{x_{max}^2} \leq \sum_{i=1}^{n} 1 = n$$ ✓

b)

$$\|x\|_2 = \sum_{i=1}^{n} |x_i|^2 = \sqrt{x_1^2 + x_2^2 + x_3^2 + \cdots + x_n^2}$$

$$\|x\|_1 = \sum_{i=1}^{n} |x_i| = |x_1| + |x_2| + \cdots + |x_n|$$

Show $\|x\|_2 \le \|x\|_1$ :

$$\sqrt{x_1^2 + x_2^2 + \cdots + x_n^2} \le |x_1| + |x_2| + \cdots + |x_n|$$

$$x_1^2 + x_2^2 + \cdots + x_n^2 \le \left(|x_1| + |x_2| + \cdots + |x_n|\right)^2$$

sum of squares is always less than square of sums

proof as follows:

Let $n=2$ so $(x_1+x_2)^2 = x_1^2 + 2x_1x_2 + x_2^2 \ge x_1^2 + x_2^2$

Inductive step

$$(x_1 + x_2 + \cdots + x_n + x_{n+1})^2 = \left((x_1 + x_2 + \cdots + x_n) + x_{n+1}\right)^2$$

$$* \quad = (x_1 + x_2 + \cdots + x_n)^2 + 2(x_1 + x_2 + \cdots + x_n)x_{n+1} + x_{n+1}^2$$

$$\ge (x_1 + x_2 + \cdots + x_n)^2 + x_{n+1}^2$$

$$\ge x_1^2 + x_2^2 + \cdots x_n^2 + x_{n+1}^2$$

$*$ $2(x_1 + x_2 + \cdots + x_n)(x_{n+1}) > 0$ ✓ by induction

Show $\frac{\|x\|_1}{\sqrt{n}} \le \|x\|_2$

$$\frac{|x_1| + |x_2| + \cdots + |x_n|}{\sqrt{n}} \le \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$$

$$\frac{(|x_1| + |x_2| + \cdots + |x_n|)^2}{n} \le x_1^2 + x_2^2 + \cdots + x_n^2$$

The Cauchy-Schwartz inequality states

$$\left(\sum_{i=1}^{n} a_i b_i\right)^2 \le \left(\sum_{i=1}^{n} a_i^2\right)\left(\sum_{i=1}^{n} b_i^2\right)$$

so let $a_i = |x_i|$ and $b_i = 1$, then

$$\left(\sum_{i=1}^{n} |x_i| \cdot 1\right)^2 \le \left(\sum_{i=1}^{n} x_i^2\right)\left(\sum_{i=1}^{n} 1^2\right)$$

$$\Rightarrow (|x_1| + |x_2| + \cdots |x_n|)^2 \le (x_1^2 + x_2^2 + \cdots + x_n^2)\, n$$

$$\Rightarrow \frac{(|x_1|+|x_2|+\cdots+|x_n|)^2}{n} \le x_1^2+x_2^2+\cdots+x_n^2 \quad \blacksquare$$

**Problem 2: (13 points)**
An example of induced matrix norm is the *p-norm* of a matrix defined as,

$$\|\boldsymbol{A}\|_p := \sup_{\|\boldsymbol{x}\|_p \neq 0} \frac{\|\boldsymbol{A}\boldsymbol{x}\|_p}{\|\boldsymbol{x}\|_p}$$

If $\sigma_1$ is the largest singular value of $\boldsymbol{A}$ then show that,
a) $\|\boldsymbol{A}\|_2 = \sigma_1$
b) If $\boldsymbol{A}$ is hermitian (or real and symmetric) then $\|\boldsymbol{A}\|_2 = \rho(A)$
c) If $\boldsymbol{A}$ is unitary then $\|\boldsymbol{A}\|_2 = 1$
Also, in each case find the relative condition number $K_2(\boldsymbol{A})$

a)

$\boxed{\text{Prove } \|A\|_2 = \sigma_1}$

We know that $A^H A = A$ & $\exists U$ s.t. $U^H A^H A U = \text{diag}(\mu_1, \mu_2, \dots \mu_n)$
where $\mu_n$ = eigenvalues of $A^H A$

and given $\|A\|_2 = \sup\limits_{\|x\|_2 \neq} \frac{\|Ax\|_2}{\|x\|_2}$, then

if $A_{n\times n}$ is symmetric, $U^T U = I$ and $U^T A U = D = \text{diag}(\mu_1, \mu_2, \dots \mu_n)$, so

$$\|A\|_2 = \sup_{\|x\|_2 \neq} \frac{\|Ax\|_2}{\|x\|_2} = \cdots$$

$$\cdots = \sup \frac{\sqrt{(Ax, Ax)}}{\sqrt{(x,x)}} = \sup \frac{\sqrt{(A^H A x, x)}}{\sqrt{(x,x)}} = \sup \frac{\sqrt{(A^H A U U^H x, x)}}{\sqrt{(U U^H x, x)}} = \cdots$$

$$\cdots = \sup \frac{\sqrt{(A^H A U U^H x, U U^H x)}}{\sqrt{(U^H x, U^H x)}}$$ let $y = U^H x$, then

$$= \sup_{\|y\| \neq 0} \frac{\sqrt{(A^H A U y, U y)}}{\sqrt{(y,y)}} = \sup_{\|y\| \neq 0} \frac{\sqrt{(U^H A^H A U y, y)}}{\sqrt{(y,y)}} = \sup_{\|y\| \neq 0} \frac{\sqrt{\sum_{i=1}^{n} \mu_i |y_i|^2}}{\sqrt{\sum_{i=1}^{n} |y_i|^2}}$$

$$= \max(\sqrt{\mu_i}) = \sqrt{\lambda_1 (A^H A)} = \sigma_1 (A) \quad \blacksquare$$

b) Given $A$ is hermitian, then $A^H A = A^2$

So $\|A\|_2 = \sqrt{\lambda_1(A^H A)} = \sqrt{\lambda_1(A^2)} = \sqrt{\lambda_1^2} = |\lambda_1| = \rho(A)$ $\blacksquare$

c) Given $A$ is unitary, then $A^H A = A^{-1} A = I$

So $\|A\|_2 = \sqrt{\lambda_1(A^H A)} = \sqrt{\lambda(I)} = \sqrt{1} = 1$ $\blacksquare$

$K_2(A)$ For each problem: $K_2(A) = \|A\|_2 \|A^{-1}\|_2$

a) $\|A\|_2 = \sigma_1$ $\quad \|A^{-1}\|_2 = \frac{1}{\sigma_n}$ $\quad$ so $K_2(A) = \frac{\sigma_1}{\sigma_n}$

b) $\|A\|_2 = |\lambda_1|$ $\quad \|A^{-1}\|_2 = \frac{1}{\lambda_n}$ $\quad$ so $K_2(A) = \frac{\lambda_1}{\lambda_n}$

c) $\|A\|_2 = 1$ $\quad \|A^{-1}\|_2 = \frac{1}{1}$ $\quad$ so $K_2(A) = 1$

## Problem 3: (15 points)

Let $\boldsymbol{A}$ be a $6 \times 6$ matrix and $\boldsymbol{b}$ be a $6 \times 1$ vector. Partition the matrix $\boldsymbol{A}$ into blocks of $4$ $3 \times 3$ matrices $\boldsymbol{C}, \boldsymbol{D}, \boldsymbol{E}, \boldsymbol{F}$ and the vector $\boldsymbol{b}$ into $2$ $3 \times 1$ blocks $\boldsymbol{g}, \boldsymbol{h}$.

$$A = \left[\begin{array}{ccc|ccc} 6 & 4 & 5 & 9 & 4 & 3 \\ 3 & 7 & 5 & 8 & 10 & 4 \\ 8 & 8 & 5 & 7 & 9 & 5 \\ \hline 2 & 1 & 4 & 4 & 6 & 3 \\ 7 & 10 & 6 & 9 & 7 & 9 \\ 2 & 8 & 6 & 6 & 6 & 2 \end{array}\right] \tag{3}$$

$$\boldsymbol{b} = \left[\begin{array}{c} 103 \\ 138 \\ 142 \\ \hline 80 \\ 170 \\ 102 \end{array}\right] = \left[\begin{array}{c} g \\ \hline h \end{array}\right] \tag{4}$$

$$A = \left(\begin{array}{c|c} C & D \\ \hline E & F \end{array}\right)$$



a) Obtain the vector $\boldsymbol{p}$ by direct matrix-vector multiplication $\boldsymbol{Ab} = \boldsymbol{p}$ and verify that the same result is obtained using block partitioned matrices

$$\left[\begin{array}{c} Cg + Dh \\ \hline Eg + Fh \end{array}\right] \equiv \boldsymbol{p} \tag{5}$$

b) Now, consider a linear system given by $\boldsymbol{Ax} = \boldsymbol{p}$. Partition the matrix $\boldsymbol{A}$ as above, and partition $\boldsymbol{p}$ and $\boldsymbol{x}$ as

$$\boldsymbol{p} = \left[\begin{array}{c} g \\ \hline h \end{array}\right], \quad \boldsymbol{x} = \left[\begin{array}{c} y \\ \hline z \end{array}\right] \tag{6}$$

Find the expressions for $\boldsymbol{y}$ and $\boldsymbol{z}$ in terms of the partitioned block matrices. Compute the solutions for $\boldsymbol{y}$ and $\boldsymbol{z}$ and verify with the solution obtained using $\boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{p}$

*(Note: Solving problems with large coefficient matrices of sizes in millions× millions can be challenging when solved serially. Block partitioning can speed up the evaluation and allow to employ multiple cores. in a parallel computing environment.)*

b) $Cg + Dh = g_p$        a-b On coded notebook

$g = C^{-1}(g_p - Dh)$

## Problem 4: (40 points)

Consider the linear system $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ with $\boldsymbol{A} \in \mathbb{R}^{n\times n}$, $\boldsymbol{b} \in \mathbb{R}^n$ and $\boldsymbol{x} \in \mathbb{R}^n$

$$\begin{aligned} \boldsymbol{A} \equiv a_{ij} &= \frac{1}{i+j-1} \quad \text{for all } i,j = 1,...,n \\ \boldsymbol{b} \equiv b_i &= \sum_{j=1}^{n} \frac{1}{i+j-1} \end{aligned} \tag{7}$$

a) For $n = 2$, use GEM to find the solution $\boldsymbol{x}$

b) Write a code that uses GEM to find the solution $\hat{\boldsymbol{x}}$. Verify the code with your solution from a). Compute $\boldsymbol{x}$ for 5 cases, $n = 5, 8, 10, 12, 15$.

c) In each case ($n = 5, 8, 10, 12$), find the condition number of the matrix $\boldsymbol{A}$ and using the relation (9) comment on the singular behavior of the $\boldsymbol{A}$ matrix in each case.
**Definition:** The relative distance of a matrix $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ from the set of singular matrices with respect to the p-norm is given as,

$$dist_p(\boldsymbol{A}) = \min\left\{\frac{\|\delta \boldsymbol{A}\|_p}{\|\boldsymbol{A}\|_p} : \boldsymbol{A} + \delta\boldsymbol{A} \text{is singular}\right\}, \tag{8}$$

and it can be shown that,

$$dist_p(\boldsymbol{A}) = \frac{1}{K(\boldsymbol{A})} \tag{9}$$

d) Now, compute solutions for the cases $n = 5, 6, \ldots, 15$. Given, that the exact solution in each case is $\boldsymbol{x} = [1, \ldots, 1]^T_{n\times 1}$.
Compute the 2-norm of the error for all $n$ values.
Plot the "error vs. n" curve. and comment on the observed behavior of error and the possible reasons for it?

a) $\underset{2}{A} = \begin{bmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{3} \end{bmatrix} \quad \underline{b} = \begin{bmatrix} 3/2 \\ 5/6 \end{bmatrix}$

$A^{(1)} \vdots \underline{b}^{(1)} = \left[\begin{array}{cc:c} 1 & 1/2 & 3/2 \\ \frac{1}{2} & 1/3 & 5/6 \end{array}\right]$

Step 1: $m_{21} = \frac{1/2}{1} = \frac{1}{2}$

$a_{21}^{(2)} = a_{21}^{(1)} - m_{21} \cdot a_{11}^{(1)} = 0$

$a_{22}^{(2)} = a_{22} - m_{21} \cdot a_{12} = \frac{1}{3} - \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{3} - \frac{1}{4} = \frac{1}{12}$

$b_2^{(2)} = b_2^{(1)} - m_{21}^{(1)} \cdot b_1^{(1)} = \frac{5}{6} - \frac{1}{2} \cdot \frac{3}{2} = \frac{5}{6} - \frac{3}{4} = \frac{10}{12} - \frac{9}{12} = \frac{1}{12}$

$\underline{A}^{(2)} \vdots \underline{b}^{(2)} = \left[\begin{array}{cc:c} 1 & 1/2 & 3/2 \\ 0 & 1/12 & 1/12 \end{array}\right]$

$x_1 + \frac{1}{2} x_2 = \frac{3}{2}$
$1/12\, x_2 = \frac{1}{12}$
$\Longrightarrow \quad x_1 = 1, \quad x_2 = 1$

b-d) Jupyter Notebook

**Problem 5: (10 points)**

a) Show that given any consistent matrix norm $\|\cdot\|$, then $\rho(\boldsymbol{A}) \leq \|\boldsymbol{A}\|$ for all $\boldsymbol{A} \in \mathbb{R}^{n\times n}$.

b) The following property holds for consistent matrix norms: Given any matrix $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ and $\epsilon > 0$, then there exists a consistent matrix norm $\|\cdot\|_\epsilon$ such that,

$$\|\boldsymbol{A}\|_\epsilon \leq \rho(\boldsymbol{A}) + \epsilon \tag{10}$$

Combining your proof along with this property we can state that,

$$\rho(\boldsymbol{A}) = \min_{\|\cdot\|} \|\boldsymbol{A}\| \tag{11}$$

Using (11) show that if $\boldsymbol{A}$ is a square matrix then,

$$\lim_{k\to\infty} \boldsymbol{A}^k = 0 \iff \rho(\boldsymbol{A}) < 1 \tag{12}$$

c) [G*] Continuing with the result you have proved in b), the geometric series $\sum_{k=0}^{\infty} \boldsymbol{A}^k$ is convergent iff $\rho(\boldsymbol{A}) < 1$. In this case show that,

$$\sum_{k=0}^{\infty} \boldsymbol{A}^k = (\boldsymbol{I} - \boldsymbol{A})^{-1} \tag{13}$$

d) [Bonus] Show that:

$$\frac{1}{1 + \|\boldsymbol{A}\|} \leq \|(\boldsymbol{I} - \boldsymbol{A})^{-1}\| \leq \frac{1}{1 - \|\boldsymbol{A}\|} \tag{14}$$

where $\|\cdot\|$ is an induced matrix norm such that $\|\boldsymbol{A}\| < 1$

*(Note: The results of this proof are useful in deriving the bounds on the change in solution when the system $\boldsymbol{Ax} = \boldsymbol{b}$ is perturbed to $(\boldsymbol{A} + \delta\boldsymbol{A})\boldsymbol{x} = \boldsymbol{b} + \delta\boldsymbol{b}$. The corresponding bound was discussed in class and you can find it in the lecture notes.)*

a) $\underline{A}\underline{v} = \lambda \underline{v}$ so $\|\underline{A}\underline{v}\| = \|\lambda v\| = |\lambda| \|v\|$

$\|\underline{A}\underline{v}\| \leq \|\underline{A}\| \|\underline{v}\|$ because matrix is consistent

$\therefore |\lambda| \|\underline{v}\| \leq \|\underline{A}\| \|\underline{v}\|$

$\Rightarrow |\lambda| \leq \|\underline{A}\|$

$\rho(\underline{A}) = |\lambda|$ so $\rho(A) \leq \|A\|$ $\boxtimes$

b)

$\boxed{\text{Show } \lim_{k\to\infty} A^k = 0 \implies \rho(A) < 1}$

Let $\lim_{k\to\infty} A^k = 0$ and $Av = \lambda v$

So, $\left(\lim_{k\to\infty} A^k\right)\underline{v} = 0\underline{v} = 0$

$\implies \lim_{k\to\infty} (A^k v) = 0$

given $Av = \lambda v$

we can say $A(Av) = A(\lambda v) = \lambda(Av) = \lambda(\lambda v)$ so $A^2 v = \lambda^2 v$

which implys $A^k v = \lambda^k v$

$\implies \lim_{k\to\infty} (\lambda^k v) = 0$

$\implies \lim_{k\to\infty} \lambda^k = 0$ therefore $1 > |\lambda_1| > |\lambda_2| > \cdots > |\lambda_n|$

and given $\rho(A) = |\lambda_1|$

then $\rho(A) < 1$

$\blacksquare$

(didn't know how to use $\rho(A) = \min_{\|\cdot\|} \|A\|$ for this direction)

$\boxed{\text{Show } \rho(A) < 1 \implies \lim_{k\to\infty} A^k = 0}$

$\rho(A) = \min_{\|\cdot\|} \|A\|$ and $\rho(A) = |\lambda| < 1$, so $|\lambda|^k \to 0$

$1 > |\lambda|^k = \|A\|_{min}^k \geq \|A^k\|_{min}$ so $\lim_{k\to\infty} \|A^k\|_{min} = 0$ which

implys $\lim_{k\to\infty} A^k = 0$ $\blacksquare$

c) $\boxed{\text{Prove } \sum_{k=0}^{\infty} A^k = (I-A)^{-1} \text{ where } \rho(A)<1}$

As stated above: $A^k \underline{v} = \lambda^k \underline{v}$

$$\text{So } \sum_{k=0}^{\infty} A^k = \sum_{k=0}^{\infty} \lambda^k$$

Because we have a geometric series, we know that $\sum_{k=0}^{\infty} ax^k = \frac{a}{1-x} \quad \forall |x|<1$

$\therefore \sum_{k=0}^{\infty} \lambda^k = \sum_{k=0}^{\infty} 1\cdot\lambda^k$ and because $\rho(A)<1$, then $|\lambda_1|<1$ and $|\lambda_1|>|\lambda_2|>\cdots>|\lambda_n|$

$$\text{So } \sum_{k=0}^{\infty} 1\cdot\lambda^k = \frac{1}{1-\lambda}$$

given $Av=\lambda v$, then $(1-\lambda)=(I-A)$

$$\text{so } \frac{1}{1-\lambda} = (1-\lambda)^{-1} = (I-A)^{-1}$$

$\boxtimes$

d) $\boxed{\text{Show that } \frac{1}{1+\|A\|} \le \|(I-A)^{-1}\| \le \frac{1}{1-\|A\|} \text{ where } \|\cdot\| \text{ is a matrix norm s.t. } \|A\|<1}$

Given that we need a $\|\cdot\|$ s.t. $\|A\|<1$, we know that $\rho(A) \le \|A\| < 1$

$$\implies \frac{1}{1+\|A\|} \le \frac{1}{1+\rho(A)} \le \|(I-A)^{-1}\| \le \frac{1}{1-\rho(A)} \le \frac{1}{1-\|A\|}$$

$$\implies \frac{1}{1+\|A\|} \le \frac{1}{1-(-|\lambda_1|)} \le \|(I-A)^{-1}\| \le \frac{1}{1-|\lambda_1|} \le \frac{1}{1-\|A\|}$$

From the previous problem we know that $(I-A)^{-1} = \frac{1}{1-\lambda}$

$$\text{So } \frac{1}{1+\|A\|} \le \frac{1}{1-(-|\lambda_1|)} \le \frac{1}{1-\lambda} \le \frac{1}{1-|\lambda_1|} \le \frac{1}{1-\|A\|}$$

we know that $-|\lambda_1| \le \lambda \le |\lambda_1|$

so the inequality is true $\boxtimes$

**Problem 6: (10 points)**

Given 2 matrices, $\boldsymbol{A} = \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix}$ and $\boldsymbol{B} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. Find whether they have a unique LU factorization. In case of unique factorization, show that $\det \boldsymbol{A} = \det \boldsymbol{U}$

$A_{(1,1)} = \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix} = 2 \qquad A_{(1,2)} = \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix} = 1$

$A_{(2,1)} = \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix} = 2 \qquad A_{(2,2)} = \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix} = 1$

A has unique LU Factorization

$B_{(1,1)} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = [0]$

the $\det([0]) = 0$ so the principal submatrix $B_{(1,1)}$ is singular so B does not have a unique LU Factorization

$m_{jk} = \frac{a_{jk}^{(k)}}{a_{kk}} \qquad A = LU$

$L^{-1}A = U \qquad L = (M_1)$

$L_A = M_1 = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$

$L_A^{-1} = \frac{1}{1-0}\begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}$

$L_A^{-1} A = U_A = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix}$

$A^{(1)} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix}$

$\det(A) = 0$

$\det(U) = 0$ ✓

# hw1

September 25, 2022

## 1 Numerical Homework 1

Code as follow for problems that require it

**Original work created on 25/09/2022**

**Author: Terry Cox**

[1]:
```
import numpy as np
import pandas as pd
```

#### 1.0.1 Problem 3

[2]:
```
A = np.matrix([[6, 4, 5, 9, 4, 3],
            [3, 7, 5, 8, 10, 4],
            [8, 8, 5, 7, 9, 5],
            [2, 1, 4, 4, 6, 3],
            [7, 10, 6, 9, 7, 9],
            [2, 8, 6, 6, 6, 2]])

b = np.array([103, 138, 142, 80, 170, 102])

g = b[:3]
h = b[-3:]

C = A[:3,:3]
D = A[:3, -3:]
E = A[-3:, :3]
F = A[-3:, -3:]

I = np.array([[1,0,0],[0,1,0],[0,0,1]])
```

**3a)**

[3]:
```
# original p
p = np.matmul(A, b)
p
```

[3]: matrix([[3586, 4733, 5238, 2558, 5781, 3866]])

[4]:
```
# partitioned p --- p_hat
p_hat = np.zeros(b.shape)
p_hat[:3] = np.matmul(C, g) + np.matmul(D, h)
p_hat[-3:] = np.matmul(E, g) + np.matmul(F, h)
p_hat
```

[4]: array([3586., 4733., 5238., 2558., 5781., 3866.])

[5]:
```
p == p_hat
```

[5]: matrix([[ True,  True,  True,  True,  True,  True]])

**3b)**

[6]:
```
# original x
x = np.matmul(np.linalg.inv(A), p.T)
x.T
```

[6]: matrix([[103., 138., 142.,  80., 170., 102.]])

[7]:
```
# partitioned x ---- y and z
y = np.matmul(np.linalg.inv(C), (p[0,:3]-np.matmul(D, h)).T)
z = np.matmul(np.linalg.inv(F), (p[0,-3:]-np.matmul(E, g)).T)
y, z
```

[7]:
```
(matrix([[103.],
         [138.],
         [142.]]),
 matrix([[ 80.],
         [170.],
         [102.]]))
```

### 1.0.2 Problem 4

**4b)**

[8]:
```
def GEM_step(A, b):
    operations = 0
    A_temp = A.copy()
    b_temp = b.copy()
    for i in range(1, A.shape[0]):
        #print(A)
        for j in range(0, A.shape[1]):
            A_temp[i,j] = A[i,j] - A[i,0]/A[0,0] * A[0,j]
            operations +=1
        b_temp[i] = b[i] - A[i,0]/A[0,0] * b[0]
```

```
        operations +=1
    return A_temp, b_temp, operations

def back_substitute(A,b):
    n = len(b)
    x = np.zeros(n)

    for i in range(n-1, -1, -1):
        temp = b[i]
        for j in range(n-1, i, -1):
            temp -= x[j]*A[i,j]

        x[i] = temp/A[i,i]
    return x

def GEM(A, b):
    o = 0
    for i in range(A.shape[0]-1):
        A_step, b_step, operations = GEM_step(A[i:,i:], b[i:])
        A[i:,i:] = A_step
        b[i:] = b_step
        o += operations
        #print(i, A, b)
    x = back_substitute(A,b)
    return A, b, x, o
```

[9]:
```
def build_A_and_b(n):
    A = np.zeros((n,n))
    b = np.zeros(n)
    s = 0
    for i in range(n):
        for j in range(n):
            A[i,j] = 1/(i+1 +j+1 -1)
            s += A[i,j]
        b[i] = s
        s = 0

    return A, b
```

[10]:
```
xs = {n : None for n in [5,8,10,12,15]}
os = xs.copy()
for n in xs.keys():
    A, b, x, o = GEM(*build_A_and_b(n))
    xs[n] = x
    os[n] = o
xs
```

[10]:
```
{5: array([1., 1., 1., 1., 1.]),
 8: array([1.        , 1.        , 0.99999998, 1.00000013, 0.99999966,
       1.00000049, 0.99999965, 1.0000001 ]),
 10: array([1.        , 1.00000011, 0.99999774, 1.00002048, 0.99990264,
       1.00026691, 0.99956309, 1.0004214 , 0.99977914, 1.0000485 ]),
 12: array([0.99999998, 1.00000272, 0.99991614, 1.00112492, 0.99185897,
       1.03538471, 0.90231481, 1.17541948, 0.79576776, 1.14866257,
       0.93852547, 1.01102248]),
 15: array([ 0.99999997,  1.00000242,  0.99999492,  0.99864026,  1.02566811,
        0.78193349,  2.06684093, -2.2790367 ,  7.53239313, -7.35504757,
        7.38066706, -1.12904142,  0.42574875,  1.73328423,  0.81795234])}
```

[ ]:

**4c)**

[11]:
```python
for n in xs.keys():
    A, b = build_A_and_b(n)
    dist_p = 1/(np.linalg.norm(A, 1)*np.linalg.norm(np.linalg.inv(A), 1))
    print(n, dist_p)
```

```
5 1.0597081987512313e-06
8 2.9522222003903674e-11
10 2.8286183568254468e-14
12 2.550908719679527e-17
15 9.144065894767732e-19
```

As the size of n increases, the relative distance of the matrix $A_{n,n}$ decreases towards 0 rappedly

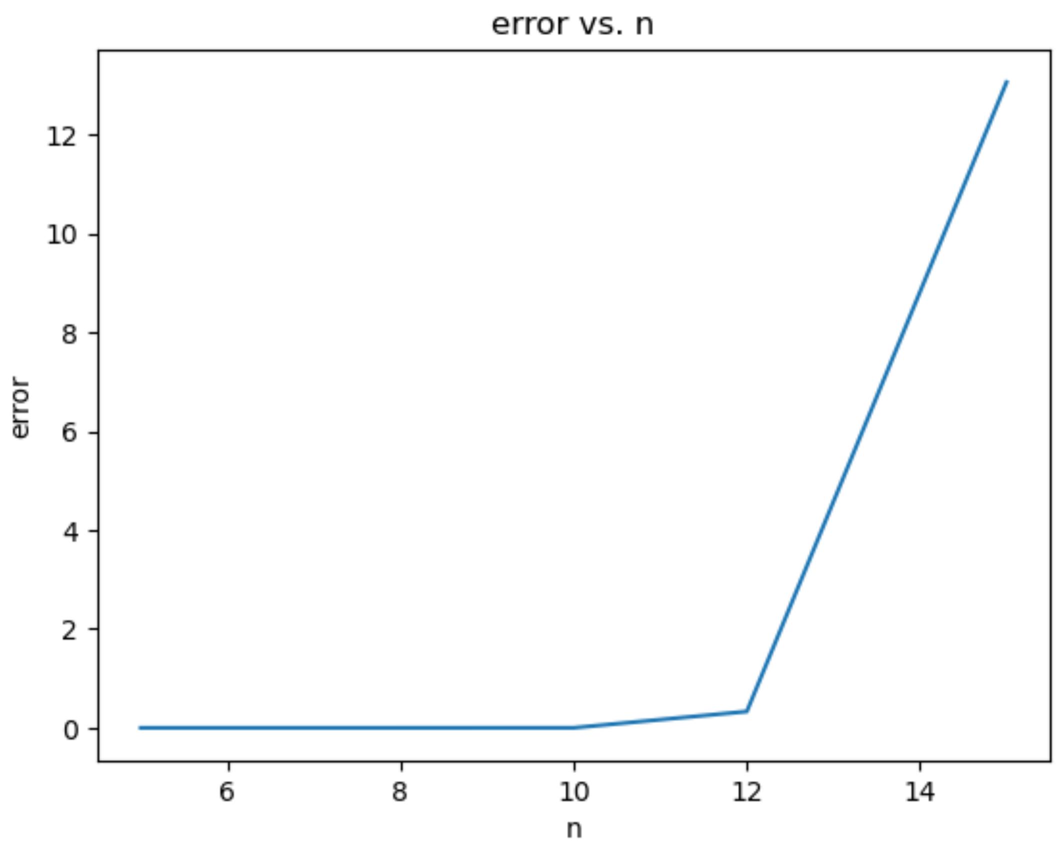
**4d)**

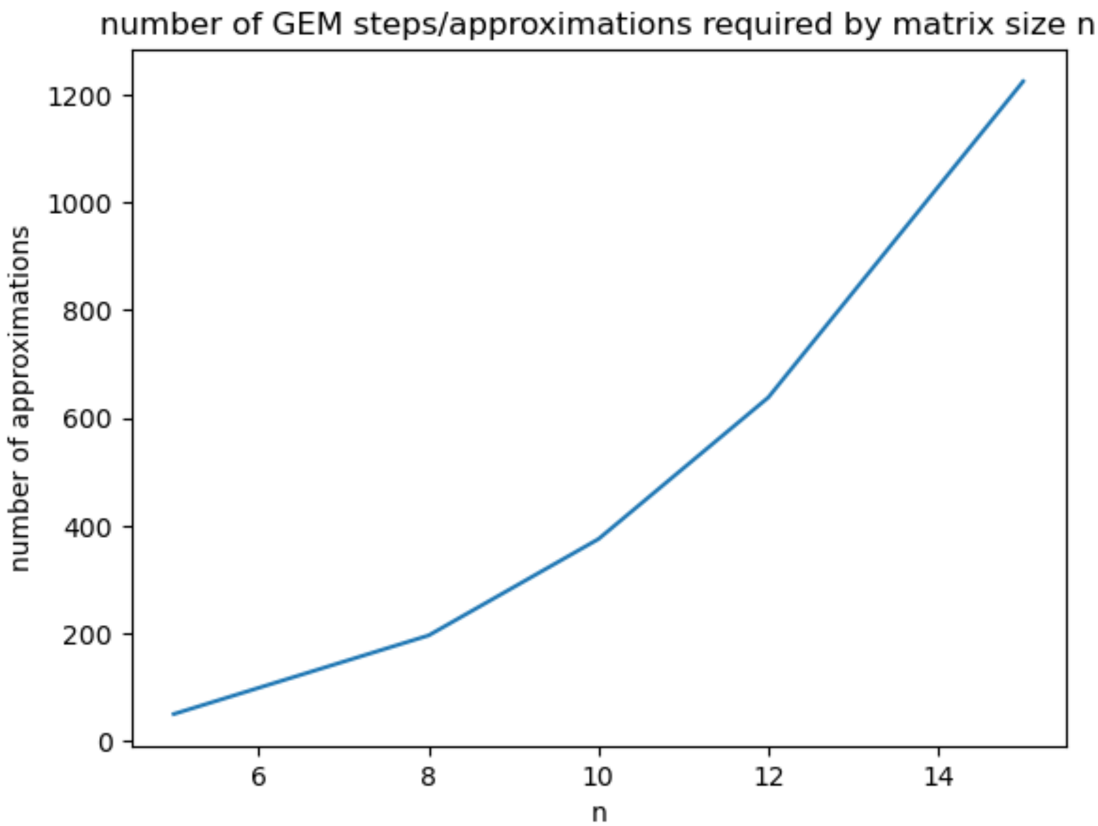
[12]:
```python
error_l2_norm = {}
for n in xs.keys():
    error_l2_norm[n] = np.linalg.norm(np.ones(len(xs[n]))-xs[n], 2)

error_l2_norm
```

[12]:
```
{5: 3.4726644211552867e-12,
 8: 7.08726214592578e-07,
 10: 0.0007076326965799963,
 12: 0.3306750623968309,
 15: 13.06000635092709}
```

[13]:
```python
import matplotlib.pyplot as plt
plt.plot(error_l2_norm.keys(), error_l2_norm.values())
plt.xlabel('n')
plt.ylabel('error')
plt.title('error vs. n')
```

```
plt.show()

plt.plot(os.keys(), os.values())
plt.xlabel('n')
plt.ylabel('number of approximations')
plt.title('number of GEM steps/approximations required by matrix size n')
plt.show()
```






As the size of the matrix increases, the error increases what appears to be exponentially. I would say the reason for this has to do with the number of variables we are working with, which would mean we are approximating more variables. These approximations are happening by a factor of $O(x^3)$ for the number of approximations happening using GEM. The number of GEM estimations by matrix size $A_{n,n}$ can be seen in the figure above as it appears to be increasing at $x^3$.

```
[ ]:
```