

CSE6643 Numerical Linear Algebra HW1

Jiahao Luo 903103970

September 20, 2016

1 Exercise 2.6

If u and v are m -vectors, the matrix $A = I + uv^*$ is known as a rank-one perturbation of the identity. Show that if A is non-singular, then its inverse has the form $A^{-1} = I + \alpha uv^*$ for some scalar α , and give an expression for α . For what u and v is A singular? If it is singular, what is $\text{null}(A)$?

Solution: (1) Suppose A is non-singular and its inverse is $A^{-1} = I + \alpha uv^*$, so

$$\begin{aligned} I &= AA^{-1} \\ &= (I + uv^*)(I + \alpha uv^*) \\ &= I + uv^* + \alpha uv^* + \alpha(v^*u)uv^* \\ &= I + (1 + \alpha + \alpha v^*u)uv^* \\ 0 &= (1 + \alpha + \alpha v^*u)uv^* \end{aligned}$$

Thus if $uv^* = 0$, $A = I$, then $A^{-1} = A = I$
if $uv^* \neq 0$, we can conclude that when

$$\alpha = -\frac{1}{1 + v^*u}$$

A 's inverse has the form $A^{-1} = I + \alpha uv^*$

(2) Suppose A is singular, then $\det(A) = \det(I + uv^*) = \det(uv^* - (-1)I) = 0$, thus the eigenvalue value for uv^* is $\lambda = -1$. Besides, since $\sum_{i=1}^n \lambda_i = \sum_{i=1}^n a_{ii}^2$, thus we know that $-1 = v^*u$. Therefore, when $v^*u = -1$, A is singular.

(3) According to the definition of $\text{Null}(A)$, we set $Ax = 0$, $x \in R^m$.

$$\begin{aligned} (I + uv^*)x &= 0 \\ x + uv^*x &= 0 \end{aligned}$$

Since $v^*u = -1$, then when $x = u$, the equation is satisfied. Thus $\text{Null}(A) = \text{span}\{u\}$.

2 Exercise 3.1

Prove that if W is an arbitrary non-singular matrix, the function $\|x\|_w$ defined by $\|x\|_w = \|Wx\|$ is a vector norm.

Proof: According to the definition of vector norm:

(1) Let $y = Wx$, then y is a vector, such that $\|y\| = \|Wx\| \geq 0$. Let $y = Wx = 0$, since W is a non-singular matrix, then $\text{rank}(W) = n$, that means if and only if $x = 0$, then there is an unique solution for $y = Wx = 0$. So when $x = 0$ satisfied $\|x\|_w = \|Wx\| = \|0\| = 0$.

(2) Since $\|x\|_w = \|Wx\|$, then we can obtain $\|x + y\|_w = \|W(x + y)\| = \|Wx + Wy\|$. Let $A = Wx, B = Wy$, since for any vector, we have $\|A + B\| \leq \|A\| + \|B\|$, thus $\|x + y\|_w = \|Wx + Wy\| \leq \|Wx\| + \|Wy\| = \|x\|_w + \|y\|_w$.

(3) To verify $\|\alpha x\|_w = |\alpha| \|x\|_w$, then let $x = \alpha x$, then according to the defined function, we have:

$$\|\alpha x\|_w = \|W(\alpha x)\| \quad (1)$$

$$= \|\alpha Wx\| \quad (2)$$

Then for vector norm $\|\cdot\|$, we have: $\|\alpha Wx\| = |\alpha| \|Wx\|$. Then we can verify $\|\alpha x\|_w = |\alpha| \|x\|_w$.

Thus, $\|x\|_w$ defined by $\|x\|_w = \|Wx\|$ is a vector norm.

3 Exercise 3.5

Example 3.6 shows that if E is an outer product $E = uv^*$, then $\|E\|_2 = \|u\|_2 \|v\|_2$. Is the same true for the Frobenius norm, i.e., $\|E\|_F = \|u\|_F \|v\|_F$? Prove it or give a counterexample.

Proof: According to the property of Frobenius norm, we have:

$$\|E\|_F = \sqrt{\text{tr}(EE^*)}$$

Besides,

$$E = uv^*$$

$$E^* = (uv^*)^* = vu^*$$

Thus we can combine these three equations to obtain:

$$\|E\|_F = \sqrt{\text{tr}(EE^*)} = \sqrt{\text{tr}(uv^*vu^*)} = \sqrt{(v^*v)\text{tr}(uu^*)} \quad (3)$$

$\|v\|_F = v^*v = \sum_{i=1}^i (v_i)^2$ is a scalar, uu^* is a $n \times n$ matrix, $\text{tr}(uu^*) = \sum_{i=1}^i (u_i)^2 = \|u\|_F$.

Thus we have:

$$\|E\|_F = \|u\|_F \|v\|_F$$

It is same true for Frobenius norm.

4

Let A and B be $n \times n$ matrices. Show that:

$$(1) \|A\|_2 \leq \|A\|_F \leq \sqrt{n} \|A\|_2$$

$$(2) \text{ Show that } \|AB\|_2 \leq \|A\|_2 \|B\|_2$$

Proof: (1) Let λ_i be the eigenvalue of $n \times n$ matrix A^*A , i denotes the i th eigenvalue of the matrix.

Assume there is the k th eigenvalue ($1 \leq k \leq n$) that is greater or equal to any of the $n - 1$ eigenvalues. Then we can denote $\lambda_{max} = \lambda_k$. Thus:

$$\|A\|_2 = \sqrt{\lambda_{max}} \quad (4)$$

$$\|A\|_F = \sqrt{\text{tr}(A^*A)} = \sqrt{\sum_{i=1}^n \lambda_i} = \sqrt{\lambda_1 + \lambda_2 + \dots + \lambda_{max} + \dots + \lambda_n} \quad (5)$$

$$\sqrt{n} \|A\|_2 = \sqrt{n \lambda_{max}} = \sqrt{\lambda_{max} + \lambda_{max} + \dots + \lambda_{max}} \quad (6)$$

Obviously, from equation (10), (11), (12), we can obtain that:

$$\|A\|_2 \leq \|A\|_F \leq \sqrt{n} \|A\|_2$$

When there is only one eigenvalue (no duplicate root), then the $\|A\|_2 = \|A\|_F = \sqrt{n} \|A\|_2$.

(2) According to the induced matrix norm definition, $\|A\|_{(m,n)}$ is the smallest number C for which the following inequality holds for all $x \in C^n$:

$$\|Ax\|_m \leq C \|x\|_n$$

And $\|A\|_{(m,n)} = \sup_{x \neq 0} \frac{\|Ax\|_m}{\|x\|_n}$, assume $x \in R^n$, we can obtain:

$$\|Ax\|_m \leq \|A\|_{(m,n)} \|x\|_n$$

Thus in this question:

$$\|ABx\|_2 \leq \|A\|_2 \|Bx\|_2 \leq \|A\|_2 \|B\|_2 \|x\|_2$$

Dividing by $\|x\|_2$:

$$\frac{\|ABx\|_2}{\|x\|_2} \leq \|A\|_2 \|B\|_2$$

Thus, $\|AB\|_2 = \sup_{x \neq 0} \frac{\|ABx\|_2}{\|x\|_2} \leq \|A\|_2 \|B\|_2$

5

Let $\|\cdot\|$ be a norm on R^n and also the corresponding induced matrix norm. Let X be an $n \times n$ invertible matrix.

- (1) Show that $|||x||| := \|Xx\|$ is a norm on R^n .
- (2) For the induced matrix norm $\|\cdot\|$, show that $|||A||| = \|XAX^{-1}\|$ for every $n \times n$ matrix A .
- (3) Let $J(\lambda)$ be the $k \times k$ Jordan cell

$$J(\lambda) = \begin{bmatrix} \lambda & 1 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \cdots & \lambda & 1 \\ 0 & \cdots & \cdots & \lambda \end{bmatrix}$$

Calculate $\|J(\lambda)\|_1$, the matrix norm induced by the 1-norm.

- (4) Let $\eta > 0$ and X be the $k \times k$ diagonal matrix X

$$X = \begin{bmatrix} \eta & 0 & \cdots & 0 \\ 0 & \eta^2 & \ddots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & \eta^k \end{bmatrix}$$

Calculate $\|XJ(\lambda)X^{-1}\|_1$. Conclude that for every $\varepsilon > 0$, there exists a matrix norm so that the norm of $J(\lambda)$ is less than $\lambda + \varepsilon$.

(1)Solution: According to the definition of vector norm, we have to verify the following conditions:

(i)Non-negativity:

For $\|\cdot\|$ vector norm definition, $|||x||| = \|Xx\| \geq 0$, besides X is non-singular, then when $x = 0$, $\|Xx\| = 0$.

(ii)Triangle inequality:

$|||x + y||| = \|X(x + y)\| = \|Xx + Xy\|$, and we can utilize the triangle inequality of vector norm, so that we can obtain: $|||x + y||| = \|Xx + Xy\| \leq \|Xx\| + \|Xy\| = |||x||| + |||y|||$

(iii)Absolute homogeneity:

$|||\alpha x||| = \|\alpha Xx\|$, for vector norm $\|\cdot\|$, we have : $\|\alpha Xx\| = |\alpha| \|Xx\| = |\alpha| |||x|||$.

Therefore, according to the above verification, we can say $|||x|||$ is a norm.

(2)Solution: According to the vector norm definition from (1). Let $y = X^{-1}x$:

$$|||A||| = \sup_{y \neq 0} \frac{|||Ay|||}{|||y|||} = \sup_{y \neq 0} \frac{||X Ay||}{||X y||} = \sup_{x \neq 0} \frac{||X A X^{-1} x||}{||X X^{-1} x||} = \sup_{x \neq 0} \frac{||X A X^{-1} x||}{||x||}$$

And by induced matrix norm definition, we can obtain:

$$||X A X^{-1}|| = \sup_{x \neq 0} \frac{||X A X^{-1} x||}{||x||}$$

Such that, $||A|| = ||X A X^{-1}||$

(3)Solution: Given vector $x \in R^k$, according to the definition of induced matrix norm, we can obtain:

$$||J(\lambda)||_1 = \sup_{x \neq 0} \frac{||J(\lambda)x||_1}{||x||_1}$$

$$J(\lambda)x = \begin{bmatrix} \lambda x_1 & x_2 & \cdots & \cdots & 0 \\ 0 & \lambda x_2 & \ddots & \ddots & 0 \\ \vdots & \cdots & \ddots & \ddots & \vdots \\ \vdots & \cdots & \cdots & \lambda x_{k-1} & x_k \\ 0 & \cdots & \cdots & \cdots & \lambda x_k \end{bmatrix}$$

Thus,

$$\begin{aligned} ||J(\lambda)||_1 &= \sup_{x \neq 0} \frac{||J(\lambda)x||_1}{||x||_1} = \sup_{x \neq 0} \frac{(|\lambda||x_1| + |\lambda||x_2| + \dots + |\lambda||x_k| + |x_2| + \dots + |x_k|)}{|x_1| + |x_2| + \dots + |x_k|} \\ &= \sup_{x \neq 0} \frac{(|\lambda| + 1)(|x_1| + |x_2| + \dots + |x_k|) - |x_1|}{|x_1| + |x_2| + \dots + |x_k|} \\ &= \sup_{x \neq 0} \left(|\lambda| + 1 - \frac{|x_1|}{|x_1| + |x_2| + \dots + |x_k|} \right) \\ &= \max_{|x_1|=0, x \neq 0} |\lambda| + 1 = |\lambda| + 1 \end{aligned}$$

(4)Solution:

$$\begin{aligned} X J(\lambda) X^{-1} &= \begin{bmatrix} \eta & 0 & \cdots & \cdots & 0 \\ 0 & \eta^2 & \ddots & \ddots & 0 \\ \vdots & \cdots & \ddots & \ddots & \vdots \\ \vdots & \cdots & \cdots & \eta^{k-1} & 0 \\ 0 & \cdots & \cdots & \cdots & \eta^k \end{bmatrix} \begin{bmatrix} \lambda & 1 & \cdots & \cdots & 0 \\ 0 & \lambda & \ddots & \ddots & 0 \\ \vdots & \cdots & \ddots & \ddots & \vdots \\ \vdots & \cdots & \cdots & \lambda & 1 \\ 0 & \cdots & \cdots & \cdots & \lambda \end{bmatrix} \begin{bmatrix} \frac{1}{\eta} & 0 & \cdots & \cdots & 0 \\ 0 & \frac{1}{\eta^2} & \ddots & \ddots & 0 \\ \vdots & \cdots & \ddots & \ddots & \vdots \\ \vdots & \cdots & \cdots & \frac{1}{\eta^{k-1}} & 0 \\ 0 & \cdots & \cdots & \cdots & \frac{1}{\eta^k} \end{bmatrix} \\ &= \begin{bmatrix} \lambda & \frac{1}{\eta} & \cdots & \cdots & 0 \\ 0 & \lambda & \ddots & \ddots & 0 \\ \vdots & \cdots & \ddots & \ddots & \vdots \\ \vdots & \cdots & \cdots & \lambda & \frac{1}{\eta} \\ 0 & \cdots & \cdots & \cdots & \lambda \end{bmatrix} \end{aligned}$$

Thus,

$$\|XJ(\lambda)X^{-1}\|_1 = |\lambda| + \frac{1}{\eta}$$

According to the induced matrix norm $\|\cdot\|$ from previous question(2), the matrix 1-norm of $J(\lambda)$ is

$$\|J(\lambda)\|_1 = \|XJ(\lambda)X^{-1}\|_1 = |\lambda| + \frac{1}{\eta}$$

Therefore, in order to satisfy this inequality: $\lambda + \frac{1}{\eta} < \lambda + \epsilon$ for every $\epsilon > 0$, we just let $\eta > \frac{1}{\epsilon}$ for the matrix X .

6

(You can use Matlab or any other programming language of your choice.)

Let $f(x) = \frac{\log(x+1)}{x}$ and the goal of the exercise is to plot the function near 0. This function can be calculated in (at least) two different ways:

1.

$$g(x) = \frac{\log(x+1)}{x}$$

2.

$$h(x) = \frac{\log(x+1)}{(x+1)-1}$$

Mathematically they are equivalent. Plot both functions around $x=0$ (use nine points on each side of 0) with step sizes 10^{-16} , 10^{-15} , 10^{-13} (and other values if necessary). Discuss what you observed and explain why.

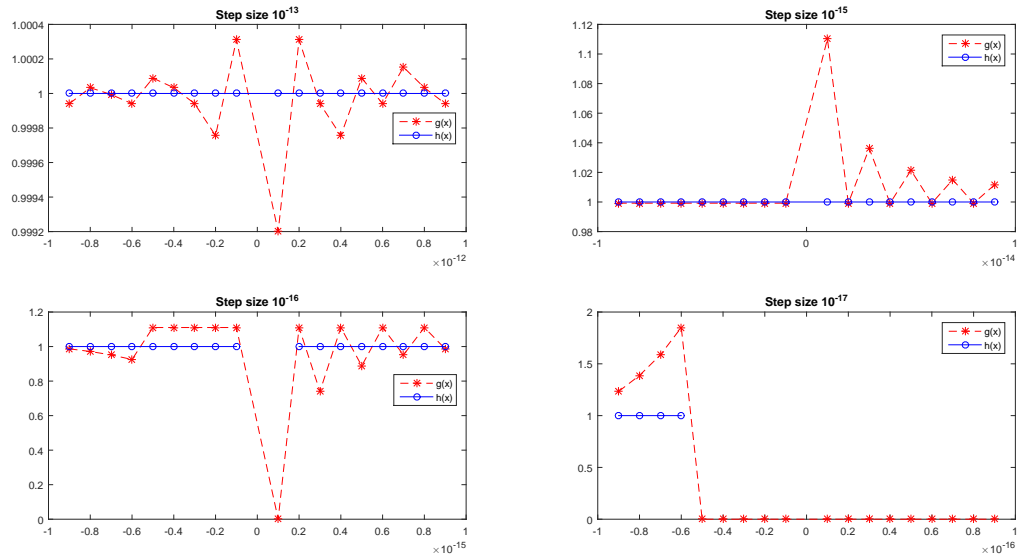


Figure 1: Function comparison with different step sizes

Answer: Figure 1 shows function $g(x)$ and $f(x)$ in different time steps. According to the figure, we can observe and conclude that:

(1) When it comes to the x which is very close to 0, the result of function $h(x)$ is basically equal to 0. While the result of function $g(x)$ is in a constant oscillation and it has a larger amplitude when x is closer to 0. It is basically a limit as x approaches to 0 problem. For the function $g(x)$, the nominator $\log(x + 1) = \log(1)$, when x is close to 0, but the denominator x remains to a infinitesimal. For the function $f(x)$, the denominator $(x + 1) - 1 \neq x$. Therefore we can see the curve is in oscillation around 1.

(2) According to the last two graphs, when the step size is less or equal to 10^{-16} , the plot of function $h(x)$ is discontinuous, while the plot of function $g(x)$ is continuous. In $h(x)$, when x is less than the step size 10^{-16} , $x + 1$ will ignore the infinitesimal part, such that the denominator will be $(x + 1) - 1 \approx 1 - 1 = 0$. Therefore that would be a *NaN* in Matlab which explains why the plot is discontinuous when x is less or equal to 10^{-16} .