Journal of Humanitarian, Scientific and Social Sciences
Elmergib University – Faculty of Arts and Sciences Kasr Khiar
4th Issue December 2017

An Application Of Inner Product Spaces

" Conditioning And the Rayleigh Quotient "

Lutfia Mohamed Aldali
Elmergib University
Facualty Of Economics & Commerce
Department of Computer and Data Analysis
Latefa.aldally@yahoo.com

الملخص:

في هذه الورقة سندرس تطبيق عملي لفضاء الضرب الداخلي , وكيف يستخدم لحل مجموعة من المعادلات على الصورة :

AX=b

حيث A مصفوفة من الرتبة m×n , b متجه من الرتبة m×1 .

Abstract:

In this paper, we introduce an application on inner product spaces and how used to solve a system of linear equations in the form Ax=b , where A is m×n matrix and b is m×1 vector.

Introduction:

The system of linear equations in the form Ax=b, where A is an m × n matrix and b is an m × 1 vector often arise in applications to the real world. The coefficients in the system are frequently obtained from experimental data, and in many cases, both m and n are so large that a computer must be used in the calculation of the solution. Thus two types of errors must be considered. First , experimental errors arise in the collection of data since no instruments can provide completely accurate measurements. Second, computers introduce roundoff errors. One might intuitively feel that small relative changes in the coefficients of the system cause small relative errors in the solution. A system that has this property is called well-conditioned ,otherwise, the system is called ill conditioned We now consider some examples of these types of errors, concentrating primarily on changes in b rather that on changes in the entries of A . In addition ,we assume that A is a square , complex(or real), invertible matrix since this is the case most frequently encountered in applications.

Definition 1 :(4)

A set **V** is called a vector space over a field **F** if :-

(a) under binary operation called addition (+)

(i) V is closed

(ii) $u + v = v + u$ for all $u, v \in V$ (commutative axiom)

(iii) $u + (v + w) = (u + v) + w$ for all $u, v, w \in V$

(associative axiom)

(iv) there exists an element $O \in V$, called a zero element such that $u + O = O + u$ for all $u \in V$.

(v) for every $v \in V$, there exists an element $(-v) \in V$ called an inverse of v such that $v + (-v) = 0 = (- v) + (v)$.

(b) under the scalar multiplication

(i) V is closed : $\alpha u \in V \ \forall \ \alpha \in F , \forall \ u \in V$.

(ii) $\alpha (u + v) = \alpha u + \alpha v$ for all $u, v \in V$; $\alpha \in F$.

(iii) $(\alpha + \beta) v = \alpha v + \beta v$ for all $v \in V$; $\alpha , \beta \in F$.

(iv) $\alpha (\beta v) = (\alpha \beta) v$ for all $v \in V$; $\alpha , \beta \in F$.

(v) $1 v= v$ for all $v \in V$; $1 \in F$.

Definition 2 : (4)

Let V be a vector space over F . An inner product on vector space V is an operation that assigns to every pair of vectors u and v in V a scalar $<u,v>$ in F such that the following properties hold for all vectors u , v and w in V and all scalars c in F.

(1) $\overline{< u , v >} = < v , u >$, where the bar denotes complex conjugation
(2) $<u + v , w > = < u , w > + < v , w >$
(3) $< cu , v > = c < u , v >$
(4) $< u , u > \geq 0$, and $< u , u > = 0$ if and only if $u = 0$.

Now , we will introduce some examples to apply this application as following :

Example 1

Consider the system

$$x_1 + x_2 = 5$$

$$x_1 - x_2 = 1$$

The solution to this system is

$$\begin{bmatrix} 3 \\ 2 \end{bmatrix}.$$

Now suppose that we change the system somewhat and consider the new system

$$x_1 + x_2 = 5$$

$$x_1 - x_2 = 1.0001$$

This modified system has the solution

$$\begin{bmatrix} 3.00005 \\ 1.99995 \end{bmatrix}.$$

We see that a change of 10^{-4} in one coefficient has caused a change of less than 10^{-4} in each coordinate of the new solution. More generally, the system

$$x_1 + x_2 = 5$$

$$x_1 - x_2 = 1+h$$

has the solution

$$\begin{bmatrix} 3 + \frac{h}{2} \\ 2 - \frac{h}{2} \end{bmatrix}.$$

Hence small changes in b introduce small changes in the solution . Of course, we are really interested in relative changes since a change in the solution of , say , 10. is considered large if the original solution is of the order 10^{-2} , but small if the original solution is of the order 10^{6} .

We use the notation δb to denote the vector $b' - b$, where b is the vector in the original system and b' is the vector in the modified system . Thus we have :

$$\delta \mathrm{b} = \begin{bmatrix} 5 \\ 1+\mathrm{h} \end{bmatrix} - \begin{bmatrix} 5 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ \mathrm{h} \end{bmatrix}$$

We now define the relative change in b to be the scalar $\|\delta b\| / \|b\|$, where $\|.\|$ denotes the standard norm on C^n or R^n ; that is , $\|b\| = \sqrt{<b,b>}$. Most of what follows, however, is true for any norm. Similar definitions hold for the relative change in x.

In this example :

$$\frac{\|\delta b\|}{\|b\|}=\frac{|h|}{\sqrt{26}} \text{ and}$$

$$\frac{\|\delta x\|}{\|x\|}=\frac{\left\|\begin{bmatrix}3+(h/2)\\2-(h/2)\end{bmatrix}-\begin{bmatrix}3\\2\end{bmatrix}\right\|}{\left\|\begin{bmatrix}3\\2\end{bmatrix}\right\|}=\frac{|h|}{\sqrt{26}}$$

Thus the relative change in x equals, coincidentally, the relative change in b; so the system is well-conditioned.

Example 2

Consider the system

$$x_1 + x_2 = 3$$

$$x_1 + 1.00001\, x_2 = 3.00001 \,,$$

which has $\begin{bmatrix}2\\1\end{bmatrix}$

as its solution . The solution to the related system

$$x_1 + x_2 = 3$$

$$x_1 + 1.00001\, x_2 = 3.00001 + h$$

is

$$\begin{bmatrix}2-(10^5)h\\1+(10^5)h\end{bmatrix}$$

Hence,

$$\frac{\|\delta x\|}{\|x\|}=\frac{\left\|\begin{bmatrix}2-(10^5)h\\1+(10^5)h\end{bmatrix}-\begin{bmatrix}2\\1\end{bmatrix}\right\|}{\left\|\begin{bmatrix}2\\1\end{bmatrix}\right\|}$$

$$=10^5\sqrt{2/5}\,|h| \quad \geq \quad 10^4\,|h|,$$

While

$$\frac{\|\delta b\|}{\|b\|} \cong \frac{\|h\|}{3\sqrt{2}}$$

Thus the relative change in x is at least 10^4 times the relative change in b! This system is very ill-conditioned . Observe that the lines defined by the two equations are nearly coincident. So a small change in either line could greatly alter the point of intersection, that is , the solution to the system.

To apply the full strength of the theory of self-adjoint matrices to the study of conditioning, we need the notion of the norm of a matrix .

Definition 3 : (2)

An n × n real or complex matrix A is self-adjoint (Hermitian) if $A = A^*$

(i.e., such that $(A)_{ij} = \overline{A_{ji}}$ for all i , j) .

Definition 4 : (3)

Let A be a complex (0r real) n × n matrix. Define the Euclidean norm of A by

$$\|A\| = \max_{x \neq 0} \frac{\|Ax\|}{\|x\|},$$

where $x \in C^n$ or $x \in R^n$.

Intuitively, $\|A\|$ represents the maximum magnification of a vector by the matrix A. The question of whether or not this maximum exists, as well as the problem of how to compute it, can be answered by the use of the so-colled Rayleigh quotient.

Definition 5 : (2)

Let B be an n × n self- adjoint matrix. The Rayleigh quotient for $x \neq 0$ is defined to be the scalar $R(x) = \langle Bx, x \rangle / \|x\|^2$.

The following result characterizes the extreme values of the Rayleigh quotient of a self-adjoint matrix.

Theorem 1 : (1)

For a self-adjoint matrix $B \in M_{nxn}(F)$, we have that $\max_{x \neq 0} R(x)$ is the largest eigenvalue of B

and $\min_{x \neq 0} R(x)$ is the smallest eigenvalue of B.

Remark 1

Let T be a linear operator on a vector space V. A nonzero vector $v \in V$ is called an eigenvector of T if there exist a scalar λ such that $T(v) = \lambda v$. The scalar λ is called eigenvalue corresponding to the eigenvector v.

Corollary 1: (1)

For any square matrix A, $\|A\|$ is finite and , in fact, equals $\sqrt{\lambda}$ where λ is the largest eigenvalue of A*A.

Lemma 1 : (1)

For any square matrix A, λ is an eigenvalue of A*A if and only if λ is an eigenvalue of AA*.

Corollary 2 : (1)

Let A be an invertible matrix . Then $\|A^{-1}\| = 1/\sqrt{\lambda}$, where λ is the smallest eigenvalue of A*A.

For many applications, it is only the largest and smallest eigenvalues that are of interest.

For example, in the case of vibration problems, the smallest eigenvalue represents the lowest frequency at which vibrations can occur.

We see the role of both of these eigenvalues in our study of conditioning.

Example 3

Let $A = \begin{bmatrix} 1 & 0 & 1 \\ -1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$

Then

$$B = A*A = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ -1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 2 & -1 & 1 \\ -1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}$$

The eigen values of B are found as follows :

We solve the equation:

Det (B- tI_3) = 0

$$B - tI_3 = \begin{bmatrix} 2 & -1 & 1 \\ -1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix} - \begin{bmatrix} t & 0 & 0 \\ 0 & t & 0 \\ 0 & 0 & t \end{bmatrix}$$

$$= \begin{bmatrix} 2-t & -1 & 1 \\ -1 & 2-t & 1 \\ -1 & 1 & 2-t \end{bmatrix}$$

$$\det(B - tI_3) = \begin{vmatrix} 2-t & -1 & 1 \\ -1 & 2-t & 1 \\ 1 & 1 & 2-t \end{vmatrix}$$

= (2-t) (4-4t+t^2) – 6+3t-2

= 8 – 8t + 2t^2 – 4t + 4t^2 –t^3 -6 + 3t -2

= -t^3+ 6t^2 – 9t

= t [-t^2+ 6t – 9]

Det(B- tI_3) = 0 , then

t[t^2-6t + 9] = 0

Hence,

t = 0 , t^2 – 6 t + 9 = 0

t= 0 (t-3 $)^2$ = 0

The eigen values of B are 3 , 3 and 0

Therefore

$$\|A\| = \max_{x \neq 0} \frac{\|AX\|}{\|x\|} \text{ for any } x = \begin{bmatrix} a \\ b \\ c \end{bmatrix} \neq 0$$

Then , by corollary $\|A\| = \sqrt{3}$

We may compute R(x) for the matrix B as $R(x)=\dfrac{<Bx,x>}{\|x\|^2}$

$$Bx=\begin{bmatrix}2 & -1 & 1\\ -1 & 2 & 1\\ 1 & 1 & 2\end{bmatrix}\begin{bmatrix}a\\ b\\ c\end{bmatrix}$$

$$=\begin{bmatrix}2a-b+c\\ -a+2b+c\\ a+b+2c\end{bmatrix}$$

$$<Bx,x>=a(2a-b+c)+b(-a+2b+c)+c(a+b+2c)$$

$$=2a^2-ab+ac-ba\_ab^2+bc+ca+cb+2c^2$$

$$=2a^2+2b^2+2c^2+2ac+2bc-2ba$$

$$\|x\|^2=a^2+b^2+c^2$$

$$R(x)=\frac{<Bx,x>}{\|x\|^2}=\frac{2(a^2+b^2+c^2-ab+ac+bc)}{a^2+b^2+c^2}\leq 3$$

Definition 6 : (1)

The number $\|A\|.\|A^{-1}\|$ is called the condition number of A and is denoted cond (A).

Theorem 2 : (1)

For the system Ax= b , where A is invertible and $b\neq 0$ the following statement are true .

(a) For any norm $\|.\|$, we have

$$\frac{1}{cond\,(A)}\frac{\|\delta b\|}{\|b\|} \le \frac{\|\delta x\|}{\|x\|} \le cond\,(A)\frac{\|\delta b\|}{\|b\|}$$

(b) If $\|.\|$ is the Euclidean norm, then cond $(A) = \sqrt{\lambda_1 / \lambda_n}$, where λ_1 and λ_n are the largest and smallest eigenvalues , respectively , of A^*A .

Proof:

For a given δb , let δx are the vector that satisfies

$A(x + \delta x) = b + \delta b$

$Ax + A(\delta x) = b + \delta b$

$A(\delta x) = \delta b$ (1)

$A^{-1}A(\delta x) = A^{-1}(\delta b)$

$\delta x = A^{-1}(\delta b)$ (2)

a) (i) $\|b\| = \|Ax\| \le \|A\|\|x\|$, so we get

$$\frac{1}{\|b\|} \ge \frac{1}{\|A\|\,\|x\|}$$

$$\frac{1}{\|x\|} \le \frac{\|A\|}{\|b\|}$$

Now , from (2) we get

$$\|\delta x\| = \|A^{-1}(\delta b)\| \le \|A^{-1}\|\,\|\delta b\|$$

$$\frac{\|\delta x\|}{\|x\|} \le \frac{\|A\|}{\|b\|} \,.\, \|A^{-1}\|\,\|\delta b\| = \|A\|\,\|A^{-1}\|\,\frac{\|\delta b\|}{\|b\|} = cond\ (A)\ \frac{\|\delta b\|}{\|b\|}$$

(ii) x = Ab , then

$$\|x\| = \|A^{-1}\|\|A^{-1}b\| \le \|A^{-1}\|\,\|b\|$$

Hence ,

$$\frac{1}{||x||} \geq \frac{1}{||A^{-1}||\;||b||}$$

$$\frac{1}{||A^{-1}||||b||} \leq \frac{1}{||x||}$$

Now , from (1) , we get

$$\|\delta b\| \leq \|A\| \|\delta b\|$$

Hence ,

$$\frac{||\delta b||}{||A^{-1}||\;||b||} \leq \frac{||A||\;||\delta x||}{||x||}$$

So we get

$$\frac{1}{||A||||A^{-1}||}\;\frac{||\delta b||}{||b||} \leq \frac{||\delta x||}{||x||}$$

Hence ,

$$\frac{1}{\text{cond (A)}}\;\frac{||\delta b||}{||b||} \leq \frac{||\delta x||}{||x||}$$

(b) From corollary 3.1 and corollary we have $\|A\| = \sqrt{\lambda_1}$, $\|A^{-1}\| = \dfrac{1}{\sqrt{\lambda_n}}$ where λ_1 and λ_n are the smallest and largest eigenvalues, respectively of A*A .

Then $cond\,(A) = \|A\| . \|A^{-1}\| = \sqrt{\dfrac{\lambda_1}{\lambda_n}}$.

Conclusion :

The solution of linear equations is of great importance in linear algebraic science .So we used Rayleigh Quotient to solve the linear equations of the form Ax = b , where A is a m×n matrix and b is a m × 1 vector .

References :-

(1) – Friedberg , s , H , Insel , A, J and Spence , l , E - " Linear Algebra " Pearson Education, Inc (2003). Illinois State University .
(2) Morris, A, O - "Linear algebra An introduction**" - (1982) -** Vnr New Mathematics Library9 .
(3) Lay David C " Linear Algebra and Its Applications " **– (2006) -** university of mary land _ college park.
(4) Bretscher, O - "Linear Algebra with Applications" - Pearson Education Inc (2005) - Colby College.