Centro Universitário de Brasília – UniCEUB

Convergência de Cadeia de Markov Não-Homogênea e Aplicações

Bolsista: Magdal Alves Custódio
Orientador: Juan Alberto Rojas Cruz
Curso: Matemática

Brasília, 1º semestre de 2006

SUMÁRIO

INTRODUÇÃO

A convergência de uma seqüência de matrizes A_1, A_2, ..., A_n, para uma matriz B implica na escolha de uma norma matricial. Assim, falamos de convergência de uma seqüência de matrizes para matriz B, em relação a uma norma $\| \ \|$, quando $\lim_{n\to\infty}\|A_n - B\| = 0$.

O interesse de estudar a convergência de seqüências de matrizes está intimamente ligado a um tipo de processo estocástico chamado *Cadeia de Markov*. O estudo assintótico de uma Cadeia de Markov com espaço de estados finitos ou enumeráveis, pode ser realizado estudando o comportamento das matrizes estocásticas associadas à cadeia. Usualmente é utilizada a norma do supremo $\|A\|_\infty = \operatorname{Sup}_i \sum_{j\geq 1} |a_{ij}|$ para analisar a convergência de produtos de matrizes estocásticas. De fato, a maior parte dos resultados teóricos sobre convergência de produtos de matrizes estocásticas encontradas em Isaacson e Madsem (1973), são estabelecidos em relação a esta norma.

Em trabalho recente, ainda não publicado, Cruz (2006), observa que alguns resultados em produto de matrizes estocásticas podem ser reformulados em relação a normas mais gerais, a saber normas que não são necessariamente completas, porém que satisfazem a desigualdade $\|AB\| \leq \|A\|.\|B\|$. Surge assim o interesse por identificar normas matriciais que satisfaçam a desigualdade acima. Ressaltamos que a identificação de normas mais gerais nas quais possam ser estabelecidos resultados sobre Cadeia de Markov é de grande interesse, uma vez que poderia facilitar as aplicações.

Em relação à pesquisa bibliográfica, destacamos o texto *Álgebra Linear com Aplicações* de Steve J. Leon. Neste livro, encontram-se alguns resultados gerais para gerar normas matriciais que satisfaçam à desigualdade $\|AB\| \leq \|A\|.\|B\|$. A técnica apresentada neste livro consiste em gerar normas matriciais a partir de normas vetoriais; dada uma norma vetorial $\| \ \|_V$ define-se a norma matricial $\| \ \|_M$; da seguinte forma:

$\|A\|_M = \max\limits_{x \neq 0} \dfrac{\|Ax\|_V}{\|x\|_V}$ e mostra-se que estas normas satisfazem a desigualdade desejada. Se considerarmos a norma vetorial $\|A\|_1 = \max\limits_{1 \leq j \leq n} \left(\sum\limits_{i=1}^{m} a_{ij} \right)$, obtemos a norma matricial $\| \; \|_\infty$.

Um outro resultado interessante também encontrado nesse livro é a norma de Frobenius $\|A\|_F = \left(\sum\limits_{j=1}^{n} \sum\limits_{i=1}^{m} a_{ij}^2 \right)^{\frac{1}{2}}$, a qual é induzida pelo produto interior canônico em R^n e satisfaz a desigualdade $\|AB\|_F \leq \|A\|_F . \|B\|_F$.

Capítulo 1

NORMAS E PRODUTOS INTERIORES

1.1 – Espaços com produto interno

Definição: Seja V um espaço vetorial real. Suponha que a cada par de vetores u,v $\in$ V associamos um número real, denotado por <u,v>. Essa função é denominada produto interno (real) em V, se satisfaz às propriedades abaixo:

(i) $<u,u> \geq 0$; e $<u,u> = 0$ se e só se $u = 0$.

(ii) $<u,v> = <v,u>$ quaisquer que sejam u e v em V.

(iii) $<\alpha u + \beta v, w> = \alpha<u,w> + \beta<v,w>$ quaisquer que sejam u, v, w em V e α, β escalares.

Exemplos de produtos interiores

1º) Produto interno usual do R^n

Se $u = (x_1, x_2, \ldots, x_n)$ e $v = (y_1, y_2, \ldots, y_n)$ são vetores genéricos do R^n, então $<u,v> = x_1y_1 + x_2y_2 + \ldots + x_ny_n$, define um produto interior.

2º) Espaço dos polinômios P(t)

O espaço vetorial P(t) de todos os polinômios é um produto interno definido definido pôr:

$$<f,g> = \int_0^1 f(t)g(t)dt$$

Por exemplo: considere $f(t) = 3t – 5$ e $g(t) = t^2$ no espaço dos polinômios P(t) então <f,g>,temos que $f(t)g(t) = 3t^3 – 5t^2$. Assim

$$<f,g> = \int_0^1 (3t^3 - 5t^2)dt = \frac{3}{4}t^4 - \frac{5}{3}t^3 \Big|_0^1 = \frac{3}{4} - \frac{5}{3} = -\frac{11}{12}$$

3º) Espaço de Hilbert

Seja V o espaço vetorial de todas as seqüências infinitas de números reais (a_1, a_2, a_3, . . .) satisfazendo a

$$\sum_{i=1}^{\infty} a_i^2 = a_1^2 + a_2^2 + a_3^2 + \ldots < \infty$$

isto é, a soma converge. A adição e a multiplicação por escalar são definidas em V componente a componente, isto é, se u = (a_1, a_2, . . .) e v = (b_1, b_2, . . .)
então, u + v = ($a_1 + b_1$, $a_2 + b_2$, . . .) e αu = (αa_1, αa_2, . . .),e temos o produto interno em V definido pôr,

$\langle u,v \rangle = a_1b_1 + a_2b_2 + \ldots$

A soma acima converge absolutamente para qualquer par de pontos em V. Desta forma, o produto está bem definido. Este espaço com produto interno é chamado de espaço ζ_2 ou espaço de Hilbert.

4º) Espaço das matrizes M = M_{mxn}

Seja M = M_{mxn} o espaço vetorial de todas as matrizes reais do tipo mxn. Um produto interno é definido em M por:

$\langle A,B \rangle = tr\,(B^tA)$ onde, tr(   ) é o traço, isto é, a soma dos elementos da diagonal principal. Se A = [a_{ij}] e B = [b_{ij}], então

$$\langle A,B \rangle = tr\,(B^tA) = \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}b_{ij}$$

Isto é, <A,B> é a soma dos elementos correspondentes a A e B.

Por exemplo: Considerando as matrizes $A = \begin{pmatrix} 1 & 2 \\ 3 & 5 \end{pmatrix}$ e $B = \begin{pmatrix} 1 & 0 \\ -1 & 2 \end{pmatrix}$

$\langle A,B \rangle = tr\,(B^tA)$

$B^t = \begin{pmatrix} 1 & -1 \\ 0 & 2 \end{pmatrix} \Rightarrow B^tA = \begin{pmatrix} 1 & -1 \\ 0 & 2 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 \\ 3 & 5 \end{pmatrix} = \begin{pmatrix} -2 & -3 \\ 6 & 10 \end{pmatrix}$

$\langle A,B \rangle = tr(B^tA) = tr\begin{pmatrix} -2 & -3 \\ 6 & 10 \end{pmatrix} = -2 + 10 = 8$.

1.2 – Propriedade num espaço com produto interno

Dois vetores u e v são ditos **ortogonais** se $<u,v> = 0$.

Por exemplo:

(a) O vetor 0 é ortogonal a todos os vetores em R^2;

(b) Os vetores $\begin{pmatrix} 3 \\ 2 \end{pmatrix}$ e $\begin{pmatrix} -4 \\ 6 \end{pmatrix}$ são ortogonais em R^2;

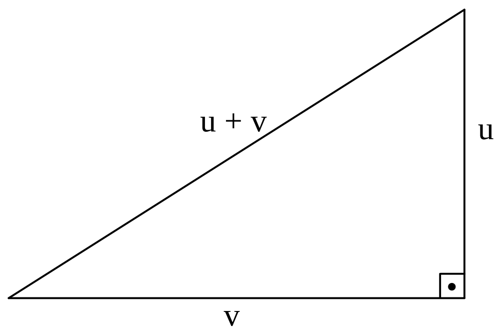
(c) Os vetores $\begin{pmatrix} 2 \\ -3 \\ 1 \end{pmatrix}$ e $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ são ortogonais em R^3.

Propriedade 1: Teorema de Pitágoras

Se u e v são vetores ortogonais em um espaço vetorial V munido de um produto interno, então

$<u + v,u + v> = <u,u> + <v,v>$.

Demonstração:



$<u + v,u + v>^2 = <u,u>^2 + 2<u,v> + <v,v>^2$

como $<u,v> = 0$, temos que

$$<u+v,u+v>^2 = <u,u>^2 + <v,v>^2$$

Propriedade 2: Desigualdade de Cauchy-Schwartz

Se u e v estão em um espaço vetorial V com produto interno, então

$$|<u,v>| \le <u,u>^{\frac{1}{2}} . <v,v>^{\frac{1}{2}} \text{ ou } <u,v>^2 \le <u,u> . <v,v>$$

Demonstração:

Para qualquer número real t temos,

$$<tu+v,tu+v> = t^2 <u,u> + 2t <u,v> + <v,v>$$

Sejam $a = <u,u>, b = 2<u,v>, c = <v,v>$.

Como $<tu+v,tu+v> \ge 0,$ temos

$$at^2 + bt + c \geq 0 \text{ para todo t.}$$

Isso significa que o polinômio de 2º grau não pode ter duas raízes reais, o que implica que

$$b^2 - 4ac \leq 0 \text{ ou que } b^2 \leq 4ac.$$

Desta forma,

$$4 < u,v >^2 \leq 4 < u,u > . < v,v > .$$

Dividindo por 4 temos o resultado desejado.

1.3 – Produto interior e normas

Definição:

Seja V um espaço vetorial. Suponha-se que a cada v $\in$ V associemos um número real, denotado por $\|v\|$. Esta função $\| . \|$ é denominada de norma em V, se satisfaz as propriedades abaixo:

i) $\|v\| \geq 0$; e $\|v\| = 0$ se, e somente se, v = 0;

ii) $\|kv\| = |k|.\|v\|$, quaisquer que sejam u, v $\in$ V, k escalar;

iii) $\|u + v\| \leq \|u\| + \|v\|$, para todo u, v $\in$ V.

Um espaço vetorial com uma norma é chamado de **espaço vetorial normado**.

1.4 – Normas induzidas por produto interiores

Teorema 1.Se < , > é um produto interior em V, então $\|v\| = < v, v >^{\frac{1}{2}}$ é uma norma.

Demonstração:

Verificação de i)

Se v $\neq$ 0, então <v,v> >0 e, assim, $\|v\| = \sqrt{< v, v >} > 0$. Se v = 0, então <0,0> = 0. Desta forma, temos que $\|0\| = \sqrt{0} = 0$. Assim i) é verdadeira.

Verificação de ii)

Temos que $\|kv\|^2 = < kv, kv > = k^2 < v, v > = k^2 \|v\|^2$. Tomando a raiz quadrada dos dois lados da igualdade obtemos ii

Verificação de iii)

Usando a desigualdade de Cauchy-Schwartz temos:

$\|u+v\|^2 =< u+v, u+v >=< u,u > +2 < u,v > + < v,v >\leq \|u\|^2 + 2\|u\|.\|v\| + \|v\|^2 = (\|u\| + \|v\|)^2$.

Tomando a raiz quadrada dos dois lados da igualdade temos:

$\|u+v\| = \|u\| + \|v\|$. Logo iii é verdadeira.

Exemplo: Seja < , > o produto interno usual em R^n, então $\|v\| = \sqrt{\langle v, v\rangle}$.

Se v = (x_1,...,x_n)

$$\|v\| = \sqrt{\langle (x_1,...x_n),(x_1,...,x_n)\rangle} = \sqrt{x_1^2 + x_2^2 + ... + x_n^2}.$$

1.5 – Normas em R^n

Norma 1. Denotada por $\|\,.\,\|_1$

A norma 1 soma os módulos dos componentes do vetor. Assim, se x = (x_1, . . ., x_n), então

$$\|x\|_1 = |x|_1 + \ldots + |x|_n = \sum_{i=1}^{n} |x_i|$$

Exemplo:

Seja x o vetor x = (1, -5, 3) em R^3. Como a norma 1, $\|x\|_1$ soma os valores absolutos dos componentes, temos então que $\|x\|_1 = |1| + |-5| + |3| = 9$.

Norma-infinito ou **norma uniforme.** Denotada por $\|\,.\,\|_\infty$

A norma-infinito ou norma uniforme escolhe o maior entre os módulos das componentes. Assim, se x = (x_1, . . ., x_n), então

$$\|x\|_\infty = \max_{1\leq i\leq n} |x_i|$$

Exemplo:

Seja x o vetor x = (1, 3, -8, -4) em R^4. A norma infinito escolhe o maior dentre os módulos das componentes. Portanto $\|x\|_\infty = 8$.

Norma p. Denotada por $\|\,.\,\|_p$, $p \geq 1$; $p \in$ R

A norma p é a raiz de ordem p da soma de cada uma das componentes elevada a p. Assim, se $x = (x_1, \ldots, x_n)$

$$\|x\|_p = (x_1^p + \ldots + x_n^p)^{1/p} = \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p}.$$

Observação:
Em particular se p = 2, temos

$$\|x\|_2 = \left(\sum_{i=1}^{n} |x_i|^2\right)^{1/2} = \sqrt{< x, x >}.$$

A norma $\|\,.\,\|_2$ é a norma em R^n proveniente do produto interior. Mostra-se que para $p \neq 2$, $\|\,.\,\|_p$ não corresponde a nenhum produto interno, pois o Teorema de Pitágoras não é válido.

1.6 – Caracterização de normas vetoriais induzidas por produtos internos

Teorema 2. Seja $\|\ \|$ uma norma em um espaço vetorial V tal que $\|u+v\|^2 + \|u-v\|^2 = 2\left(\|u\|^2 + \|v\|^2\right)$, para todo u, v $\in$ V. Então existe um produto interior < , > em V tal que $\|v\| = < v, v >^{\frac{1}{2}}$

Demonstração:

É suficiente mostrar que a função $f(u,v) = \frac{1}{2}\left(\|u+v\|^2 - \|u\|^2 - \|v\|^2\right)$ define um produto interior.

$$\text{i) } f(u,u) = \frac{1}{2}\left(\|u+u\|^2 - \|u\|^2 - \|u\|^2\right) = \frac{1}{2}\left(\|2u\|^2 - \|u\|^2 - \|u\|^2\right)$$
$$= \frac{1}{2}\left[(4\|u\|)^2 - \|u\|^2 - \|u\|^2\right] = \frac{1}{2}\left(4\|u\|^2 - 2\|u\|^2\right) = \frac{1}{2}\left(2\|u\|^2\right) = \|u\|^2 .$$

Logo, $f(u,u) = \|u\|^2$, e como $\|\ \|$ é uma norma, concluímos que f(u,u) ≥ 0 e f(u) = 0 se, e somente se, u = 0 (vetor nulo).

ii) f(ku,v) = k f(u,v), onde k é um escalar.

Temos que

$$f(ku,v) = \frac{1}{2}\left(\|ku+v\|^2 - \|ku\|^2 - \|v\|^2\right) = \frac{1}{2}\left(\|ku\|^2 + 2\|ku\|.\|v\| + \|v\|^2 - \|ku\|^2 - \|v\|^2\right) =$$
$$= \frac{1}{2}\left(2\|ku\|.\|v\|\right) = \|ku\| . \|v\| = k\|u\| . \|v\| \qquad \text{(a)}$$

Por outro lado:

$$kf(u,\ v) = k\,\frac{1}{2}\left(\|u+v\|^2 - \|u\|^2 - \|v\|^2\right) = k\,\frac{1}{2}\left(\|u\|^2 + 2\|u\| . \|v\| + \|v\|^2 - \|u\|^2 - \|v\|^2\right)$$
$$= k\,\frac{1}{2}\left(2\|u\| . \|v\|\right) = k\|u\| . \|v\|. \qquad \text{(b)}$$

de (a) e (b) constatamos que f(ku,v) = kf(u,v).

iii) Verifiquemos a bilinearidade da função $f(u,v) = \frac{1}{2}\left(\|u+v\|^2 - \|u\|^2 - \|v\|^2\right)$, ou seja

f(u + v,w) = f(u,w) + f(v,w) para todo u, v e w em V.

Por hipótese temos: $\|u\|^2 + \|v\|^2 = \frac{1}{2}\left(\|u+v\|^2 + \|u-v\|^2\right)$

Logo,

$$\|u+w\|^2 + \|u+v\|^2 = \frac{1}{2}\left(\|2u+w+v\|^2 + \|w-v\|^2\right) \qquad (1)$$

$$\|u+v+w\|^2 + \|u\|^2 = \frac{1}{2}\left(\|2u+v+w\|^2 + \|v+w\|^2\right) \qquad (2)$$

$$\|v\|^2 + \|w\|^2 = \frac{1}{2}\left(\|v+w\|^2 + \|v-w\|^2\right) \qquad (3)$$

somando $\|v+w\|^2$ na igualdade (1), obtemos:

$$\|u+w\|^2+\|u+v\|^2+\|v+w\|^2=\frac{1}{2}\left(\|2u+w+v\|^2+\|w-v\|^2\right)+\|v+w\|^2$$
$$=\frac{1}{2}\left(\|2u+w+v\|^2+\|v+w\|^2\right)+\frac{1}{2}\left(\|v+w\|^2+\|v-w\|^2\right)$$

utilizando as igualdades (2) e (3), podemos escrever:

$$\|u+w\|^2+\|u+v\|^2+\|v+w\|^2=\|u+v+w\|^2+\|u\|^2+\|v\|^2+\|w\|^2$$

ou seja,

$$\|u+v+w\|^2-\|u+v\|^2-\|w\|^2=\|u+w\|^2+\|v+w\|^2-\|u\|^2-\|v\|^2-2\|w\|^2 \qquad (4)$$

$$f(u,w)=\frac{1}{2}\left(\|u+w\|^2-\|u\|^2-\|w\|^2\right)$$

como $f(v,w)=\frac{1}{2}\left(\|v+w\|^2-\|v\|^2-\|w\|^2\right)$

$$f(u+v,w)=\frac{1}{2}\left(\|u+v+w\|^2-\|u+v\|^2-\|w\|^2\right)$$

concluímos de (4) que: f(u + v, w) = f(u,w) + f(v,w).

Aplicações do Teorema 2:

A) Verifiquemos se a norma Euclidiana $\|\ \|_E$ é induzida por um produto interior. Para facilitar a notação consideremos a norma $\|x\|_E$ em R^2. Pelo Teorema 2 é suficiente verificar a igualdade $\|u+v\|^2+\|u-v\|^2=2\left(\|u\|^2+\|v\|^2\right)$.

Seja $u=(x_1,x_2)$ e $v=(y_1,y_2)$.

$$\|u+v\|^2+\|u-v\|^2=\|(x_1+y_1),(x_2+y_2)\|^2+\|(x_1-y_1),(x_2-y_2)\|^2$$
$$=(x_1+y_1)^2+(x_2+y_2)^2+(x_1-y_1)^2+(x_2-y_2)^2$$
$$=x_1^2+y_1^2+x_2^2+y_2^2+x_1^2+y_1^2+x_2^2+y_2^2+2(x_1y_1+x_2y_2-x_1y_1-x_2y_2)$$
$$=2\left(x_1^2+x_2^2+y_1^2+y_2^2\right)$$

e $2\left(\|u\|^2+\|v\|^2\right)=2\left(\left(x_1^2+x_2^2\right)+\left(y_1^2+y_2^2\right)\right)$, assim $\|u+v\|^2+\|u-v\|^2=2\left(\|u\|^2+\|v\|^2\right)$ para qualquer u, v em R^2, do Teorema 2, concluímos que a norma $\|x\|_E$ é induzida por um produto interior.

B) Analisemos se a norma $\|x\|_{\infty} = \max\{|x_1|, |x_2|, ..., |x_n|\}$ é induzida por um produto interior, isto é, vejamos se a igualdade $\|u+v\|^2 + \|u-v\|^2 = 2(\|u\|^2 + \|v\|^2)$ se verifica para qualquer escolha de u e v.

Observemos que se u = (2, 1, 0, ..., 0) e v = (1, 2, 0, ..., 0), então:

$\|u\| = 2 = \|v\|$, $\|u+v\| = 3$ e $\|u-v\| = 1$.

Logo, $$\begin{aligned} 3^2 + 1^2 &\neq 2(2^2 + 2^2) \\ 10 &\neq 16 \end{aligned}$$

Concluímos do Teorema 2 que não existe produto interior em R^n que induz a norma infinito.

C) Vejamos mais um exemplo, a norma matricial $\|A\| = tr(A^t A)$, onde A^t denota a transposta de A e tr denota o traço de A, isto é, a soma dos elementos da diagonal principal. Novamente consideramos o caso de matriz 2x2.

Seja $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ logo

$tr(A^t A) = tr\left(\begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{bmatrix} . \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}\right) = tr\left(\begin{bmatrix} a_{11}^2 + a_{21}^2 & \\ & a_{12}^2 + a_{22}^2 \end{bmatrix}\right) = a_{11}^2 + a_{21}^2 + a_{12}^2 + a_{22}^2$, por outro lado $\|A\| = a_{11}^2 + a_{21}^2 + a_{12}^2 + a_{22}^2$ se $B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$ com

$\|B\| = b_{11}^2 + b_{12}^2 + b_{21}^2 + b_{22}^2$,

temos que: $A + B = \begin{pmatrix} a_{11}+b_{11} & a_{12}+b_{12} \\ a_{21}+b_{21} & a_{22}+b_{22} \end{pmatrix}$ e $A - B = \begin{pmatrix} a_{11}-b_{11} & a_{12}-b_{12} \\ a_{21}-b_{21} & a_{22}-b_{22} \end{pmatrix}$

$$\|A+B\| = (a_{11}+b_{11})^2 + (a_{12}+b_{12})^2 + (a_{21}+b_{21})^2 + (a_{22}+b_{22})^2$$
$$\|A-B\| = (a_{11}-b_{11})^2 + (a_{12}-b_{12})^2 + (a_{21}-b_{21})^2 + (a_{22}-b_{22})^2$$

Por último verifiquemos a igualdade $\|u+v\|^2 + \|u-v\|^2 = 2(\|u\|^2 + \|v\|^2)$, assim teremos a igualdade acima, mostrando que a norma $\|A\|$ é induzida por um produto interior.

Observação: Na demonstração do Teorema 2 encontra-se uma fórmula para achar o produto interior que induz a norma (quando se tenha a igualdade). Por exemplo a norma $\|x\|_1 = |x_1| + |x_2| + ... + |x_n|$ é induzida pelo produto interior:

$\langle u, v\rangle = \frac{1}{2}\left(\|u+v\|^2 - \|u\|^2 - \|v\|^2\right) = \|u\|^2 + \|v\|^2$. Tomemos u = ($x_1$,$x_2$) e v = ($y_1$,$y_2$)

$$= \frac{1}{2}\left[(x_1 + y_1)^2 + (x_2 + y_2)^2 + (x_1 - y_1)^2 + (x_2 - y_2)^2\right]$$

$$= \frac{1}{2}\left[x_1^2 + y_1^2 + x_2^2 + y_2^2 + x_1^2 + y_1^2 + x_2^2 + y_2^2 + 2(x_1y_1 + x_2y_2 - x_1y_1 - x_2y_2)\right]$$

$$= \frac{1}{2}\left[2\left(x_1^2 + x_2^2 + y_1^2 + y_2^2\right)\right]$$

$$= x_1^2 + x_2^2 + y_1^2 + y_2^2 = \left(x_1^2 + x_2^2\right) + \left(y_1^2 + y_2^2\right) = \langle u, u\rangle + \langle v, v\rangle = \|u\|^2 + \|v\|^2 = \langle u, v\rangle.$$

Capítulo 2

NORMAS MATRICIAIS

2.1 – Norma de Frobenius

Definição: Para o espaço vetorial R^{mxn}, a norma definida pelo produto interno é chamada de norma de Frobenius e é denotada por $\|\cdot\|_F$. Logo, se A $\in \mathrm{R}^{mxn}$, temos

$$\|A\|_F = (\langle A , A\rangle)^{1/2} = \left(\sum_{j=1}^{n}\sum_{i=1}^{m} a^2_{i,j}\right)^{1/2}$$

Exemplo:

Sejam as matrizes $A = \begin{pmatrix} 3 & 3 \\ 1 & 2 \\ 1 & 1 \end{pmatrix}$ e $B = \begin{pmatrix} -3 & 4 \\ 3 & 0 \\ -1 & 1 \end{pmatrix}$ então a norma de Frobenius dessas duas matrizes são dadas por:

$$\|A\|_F = (9+9+1+4+1+1)^{1/2} = 5$$

$$\|B\|_F = (9+16+9+0+1+1)^{1/2} = 6$$

A norma de Frobenius tem propriedades importantes que listamos a seguir.

I. Se $\mathbf{a}_j$ representa a j-ésima coluna de A, então

$$\|A\|_F = (\langle A , A\rangle)^{1/2} = \left(\sum_{j=1}^{n}\sum_{i=1}^{m} a^2_{i,j}\right)^{1/2} = \left(\sum_{j=1}^{n}\|a_j\|_2^2\right)^{1/2}$$

II. Se **a**(i,:) representa a i-ésima linha de A, então

$$\|A\|_F = \left(\sum_{i=1}^{m}\sum_{j=1}^{n} a^2_{i,j}\right)^{1/2} = \left(\sum_{i=1}^{m}\|a(i,:)^T\|_2^2\right)^{1/2}$$

III. Se x$\in \mathrm{R}^n$, então

$$\|Ax\|_2 = \left[\sum_{i=1}^{m}\left(\sum_{j=1}^{n} a_{i,j}x_j\right)^2\right]^{1/2} = \left[\sum_{i=1}^{m}(a(i,:)x)^2\right]^{1/2} \le \left[\sum_{i=1}^{m}\|x\|_2^2\|a(i,:)^T\|_2^2\right]^{1/2} (Cauchy - Schwarz)$$

IV. Se B = (b_1, ..., b_r) é uma matriz n x r, então, pelas propriedades **I** e **III**,

$$
\begin{aligned}
\|AB\|_F &= \|(Ab_1, Ab_2, \ldots, Ab_r)\|_F \\
&= \left(\sum_{i=1}^{r} \|Ab_i\|_2^2 \right)^{\frac{1}{2}} \\
&\le \|A\|_F \left(\sum_{i=1}^{r} \|b_i\|_2^2 \right)^{\frac{1}{2}} \\
&= \|A\|_F \|B\|_F
\end{aligned}
$$

2.2 – Uma leve generalização da norma de Frobenius

Definamos $\|A\|_{FG} = l\,\|A\|_F$ onde $l \ge 1$

a) Verifiquemos inicialmente que $\|A\|_{FG}$ é uma norma .

i) $\|A\|_{FG} = l\,\|A\|_F \ge 0$, pois λ > 0 e $\|A\|_F \ge 0$.

ii) $\|kA\|_{FG} = l\,\|kA\|_{FG} = l\,k\|A\|_F = k\,l\,\|A\|_F = k\|A\|_{FG}$.

iii) $\|A+B\|_{FG} = l\,\|A+B\|_F \le l\,(\|A\|_F + \|B\|_F)$

$$
\begin{aligned}
&= l\,\|A\|_F + l\,\|B\|_F \\
&= \|A\|_{FG} + \|B\|_{FG.}
\end{aligned}
$$

b) Vejamos que a norma $\|A\|_{FG}$ herda a propriedade $\|A.B\|_F \le \|A\|_F.\|B\|_F$.

De fato, $\|A.B\|_{FG} = l\,\|A.B\|_F \le l\,\|A\|_F.\|B\|_F$

$$
\begin{aligned}
&\le l^2\|A\|_F.\|B\|_F \;\; (\text{pois } l \ge 1) \\
&= l\,\|A\|_F.\,l\,\|B\|_F \\
&= \|A\|_{FG}.\|B\|_{FG}
\end{aligned}
$$

Assim, $\|A.B\|_{FG} \le \|A\|_{FG}.\|B\|_{FG}$.

Capítulo 3

NORMAS MATRICIAIS SUBORDINADAS

Consideremos que cada matriz m x n é um operador linear de R^n em R^m, para qualquer conjunto de normas vetoriais, podemos definir uma norma do operador comparando $\|A\mathbf{x}\|$ e $\|\mathbf{x}\|$ para cada x não nulo, definindo

(1) $\|A\| = \max\limits_{x \neq 0} \dfrac{\|A\mathbf{x}\|}{\|\mathbf{x}\|}$

É possível mostrar que existe um $\mathbf{x}_0$ particular em R^n que maximiza $\dfrac{\|A\mathbf{x}\|}{\|\mathbf{x}\|}$. Supondo que $\dfrac{\|A\mathbf{x}\|}{\|\mathbf{x}\|}$ sempre pode ser maximizada, vamos mostrar que (1) define, de fato, uma norma em R^{mxn}. Para isso, precisamos verificar que cada uma das três condições abaixo se verifica:

i) Para cada $x \neq 0, \dfrac{\|A\mathbf{x}\|}{\|\mathbf{x}\|} \geq 0$ e $\|A\mathbf{x}\| = \max\limits_{x \neq 0} \dfrac{\|A\mathbf{x}\|}{\|\mathbf{x}\|} \geq 0.$

De fato, se $\|A = 0\|$, então A$\mathbf{x}$ = 0 para todo $\mathbf{x} \in R^n$.

Isso implica que $a_j = Ae_j = 0$ para j=1, . . ., n e , portanto, A tem que ser matriz nula.

ii) $\|\alpha A\| = \max\limits_{x \neq 0} \dfrac{\|\alpha A\mathbf{x}\|}{\|\mathbf{x}\|} = |\alpha| \max\limits_{x \neq 0} \dfrac{\|A\mathbf{x}\|}{\|\mathbf{x}\|} = |\alpha|.\|A\|;$

iii) Se $\mathbf{x} \neq 0,$ então $\|A + B\| = \max \dfrac{\|(A+B)\mathbf{x}\|}{\|\mathbf{x}\|}.$

De fato, $\|A + B\| \leq \max\limits_{x \neq 0} \dfrac{\|A\mathbf{x}\| + \|B\mathbf{x}\|}{\|\mathbf{x}\|}$

$$\leq \max_{x \neq 0} \frac{\|A\mathbf{x}\| + \|B\mathbf{x}\|}{\|\mathbf{x}\|}$$

$$\leq \max_{x \neq 0} \frac{\|A\mathbf{x}\|}{\|\mathbf{x}\|} + \max_{x \neq 0} \frac{\|B\mathbf{x}\|}{\|\mathbf{x}\|}$$

$$= \|A\| + \|B\|$$

Logo, (1) define uma norma em R^{mxn}. Para cada família de normas vetorias $\| \ \|_V$, poderíamos, então, definir uma família de normas matriciais por (1). As normas matriciais definidas por (1) são ditas **subordinadas** às normas vetorias $\| \ \|_V$.

Teorema 3. Se a família de normas matricias $\| \ \|_M$ é subordinada à família de normas vetorias $\| \ \|_V$, então $\| \ \|_M$ satisfaz a propriedade $\|AB\|_M \leq \|A\|_M \|B\|_M$.

Demonstração:

Se **x** é um vetor qualquer não-nulo em R^n, então $\dfrac{\|A\mathbf{x}\|_V}{\|\mathbf{x}\|_V} \leq \max\limits_{y \neq 0} \dfrac{\|A\mathbf{y}\|_V}{\|\mathbf{y}\|_V} = \|A\|_M$.

Logo,

(2) $\|A\mathbf{x}\|_V \leq \|A\|_M . \|\mathbf{x}\|_V$

Se B é uma matriz n x r, segue-se de (2)

$\|AB\mathbf{x}\|_V \leq \|A\|_M \|B\mathbf{x}\|_V \leq \|A\|_M \|B\|_M \|\mathbf{x}\|_V$.

Logo, para todo $\mathbf{x} \neq 0$,

$\dfrac{\|AB\mathbf{x}\|_V}{\|\mathbf{x}\|_V} \leq \|A\|_M \|B\|_M$ e portanto,

$\|AB\|_M = \max\limits_{x \neq 0} \dfrac{\|ABx\|_V}{\|x\|_V} \leq \|A\|_M \|B\|_M$.

Observações:

$\|A\|_2 = \max\limits_{x \neq 0} \dfrac{\|A\mathbf{x}\|_2}{\|\mathbf{x}\|_2}$ é difícil de ser calculada, porém as normas $\|A\|_1 = \max\limits_{x \neq 0} \dfrac{\|A\mathbf{x}\|_1}{\|\mathbf{x}\|_1}$ e $\|A\|_\infty = \max\limits_{x \neq 0} \dfrac{\|A\mathbf{x}\|_\infty}{\|\mathbf{x}\|_\infty}$ são fáceis de se calcular.

Teorema 4. Se A é uma matriz m x n, então

$$\|A\|_1 = \max_{1\le j\le n}\left(\sum_{i=1}^{m}|a_{ij}|\right) \text{ e } \|A\|_\infty = \max_{1\le i\le m}\left(\sum_{j=1}^{n}|a_{ij}|\right).$$

Demonstração:

Vamos provar que $\|A\|_1 = \max_{i\le j\le n}\left(\sum_{i=1}^{m}|a_{ij}|\right)$. Seja $a = \max_{i\le j\le n}\sum_{i=1}^{m}|a_{ij}| = \sum_{i=1}^{m}|a_{ik}|$

Note que k é o índice da coluna onde ocorre o máximo. Seja x um vetor arbitrário em R^n. Então

$$Ax = \left(\sum_{j=1}^{n}a_{1j}x_j, \sum_{j=1}^{n}a_{2j}x_j, \ldots, \sum_{j=1}^{n}a_{mj}x_j\right)^T$$

e temos que

$$\begin{aligned}
\|Ax\|_1 &= \sum_{i=1}^{m}\left|\sum_{j=1}^{n}a_{ij}x_j\right| \\
&\le \sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}x_j| \\
&= \sum_{j=1}^{n}\left(|x_j|\sum_{i=1}^{m}|a_{ij}|\right) \\
&\le a\sum_{j=1}^{n}|x_j| \\
&= a\|x\|_1
\end{aligned}$$

Logo, qualquer que seja o vetor não-nulo x em R^n,

$$\frac{\|Ax\|_1}{\|x\|_1} \le a$$

e portanto,

$$\|A\|_1 = \max_{x\ne 0}\frac{\|Ax\|_1}{\|x\|_1} \le a \qquad (1)$$

Por outro lado,

$$\|Ae_k\|_1 = \|a_k\|_1 = a$$

Como $\|e_k\|_1 = 1$, temos que

$$\|A\|_1 = \max_{x \neq 0} \frac{\|Ax\|_1}{\|x\|_1} \geq \frac{\|Ae_k\|_1}{\|e_k\|_1} = a \qquad (2)$$

As equações (1) e (2) juntas implicam que $\|A\|_1 = a$.

Exemplo:

Seja $A = \begin{bmatrix} -3 & 2 & 4 & -3 \\ 5 & -2 & -3 & 5 \\ 2 & 1 & -6 & 4 \\ 1 & 1 & 1 & 1 \end{bmatrix}$

A maior entre as somas dos valores absolutos nas colunas, ocorre na 3ª coluna,

então $\|A\|_1 = |4| + |-3| + |-6| + |1| = 14$

A maior entre as somas dos valores absolutos nas linhas, ocorre na 2ª linha,

então $\|A\|_\infty = |5| + |-2| + |-3| + |5| = 15$.

RESULTADO DA ANÁLISE E INTERPRETAÇÃO DOS DADOS E CONCLUSÕES

O método estudado, para gerar normas matriciais por meio de normas vetoriais, é de grande relevância teórica, uma vez que uma norma vetorial qualquer gera uma norma matricial, satisfazendo a desigualdade. Assim tem-se a falsa impressão de que teríamos um grande método para gerar muitas normas matriciais, porém este método fornece somente uma expressão analítica para a norma matricial e implica na maximização de funções, o que nem sempre é fácil.

Do estudo realizado destacamos a norma de Frobenius por ela satisfazer a desigualdade $\|AB\| \leq \|A\|.\|B\|$, porém, não é claro que esta norma possa ser obtida pelo método geral. Surgi assim o seguinte questionamento: a norma de Frobenius pode ser obtida pelo método geral? Esse questionamento guiou o nosso caminho.

Um outro questionamento formulado e que também merece ser destacado foi: uma norma matricial obtida por meio de um produto interior satisfaz necessariamente a desigualdade $\|AB\| \leq \|A\|.\|B\|$?. O fundamento para tal questionamento está no fato de que a norma matricial de Frobenius é obtida por um produto interior e também satisfaz a desigualdade $\|AB\| \leq \|A\|.\|B\|$.

Neste trabalho observou-se que se uma norma matricial $\|\cdot\|_F$ satisfaz a desigualdade desejada, então $\|A\|_{FG} = l\|A\|_F$ também satisfaz a desigualdade se $l \geq 1$.

Por último fazemos notar que a demonstração do teorema 2 (caraterização de normas induzidas por produtos interiores) apresentada neste trabalho, foi obtida sem auxilio bibliográfico, o que gerou muita satisfação.

Acreditamos que o mérito deste trabalho de graduação está na formulação de questionamentos iniciais importantes que certamente poderão servir de novos pontos de partida para estudos posteriores mais aprofundados.

BIBLIOGRAFIA

Cruz, Juan Alberto Rojas. *Convergência de cadeias de Markov Não-Homogêneas: Ergodicidade Fraca e Forte*, Tese de Doutorado, departamento de Matemática, Universidade de Brasília, Brasil. (1998)

Cruz, Juan Alberto Rojas; Dorea, Chang Chung Yu. *Simple Conditions for the Convergence of Simulated Annealing Type Algorithms*. Journal of Apllied Probability, v. 35, n. 4, p. 885-892, (1998).

Cruz, Juan Alberto Rojas; Dorea, Chang Chung; **R, Baigorri A.** *Assymptotic Behavior of Markovian Algorithms with General State Space*. Stochastic Modelling And Applications, v.2, n. 2, p. 1-12, (1999).

Cruz, Juan Alberto Rojas; Dorea, Chang Chung; A.G.C. Pereira. *Metropolis-Hastings Algorithm: the Non-Homogeneous Case.* Stochastic Modelling And Applications, v. 5, n. 2, p. 10-20, (2002)

Cruz, Juan Alberto Rojas; Dorea, Chang Chung. *Approximation Results for Non-Homogeneous Markov Chains and Some Applications*. SANKHYÃ, The Indian Journal of Statistics, Índia, v. 66.serie A, p. 1-10, (2004).

Doob, J. L., *Stochastic Processes*, Wiley, New York, (1953).

Dorea, Chang Chung Yu. *Stationary distribution of Markov chains in R with applications to global random aptimization*, Bernoulli, v. 3, n. 4, p. 415-427, (1997)

Dorea, Chang Chung Yu. *On the efficiencyof a continuous version of the simulated annealing algorithm, Stat. And Probability Letters*, v. 31 (1997)

Gidas, B., *Nonstationary Markov chains and convergence f the annealing algoritm*, Journal of Statistical Physicis, v. 39, p. 73-131 (1985).

Hajek, B., *Cooling shedules for optimal Annealing*, Mathematics of Operations research, v. 13, n. 2, p. 311-329, (1988).

Hajnal, J., *The ergodic properties of non-homogeneous finite markov chains, Proceedings of the Cambrige Philosophical Society*, v. 52, p. 67-77, (1956)

Hajnal, J., *Weak ergodicity in non-mm, Proceedings of the Cambridge Philosophical Society*, v. 54, p. 233-246, (1958).

Isaacson, D; Madsen, D. *Markov chains,* Wiley, New York, (1976).

Isaacson, D; Madsen, D. *Strongly ergodic behavior for non-stationary Markov processes, The Annals of Probability*, v. 1, p. 329-335, (1973).

Leon, Steven J. *Álgebra Linear com Aplicações*, LTC, Rio de Janeiro, (1999).

Lipschutz, Seymour; Lipson, Marc. Álgebra Linear. 3 ed. São Paulo: Bokman, 2004.

Pereira, A.G.C. *Análise da Ergodicidade das Cadeias de Markov não-homogêneas Via condições do Tipo Doeblin*. Tese de doutorado, Departamento de Matemática, Universidade de Brasília, Brasil, (2001).

ANEXO SOBRE CADEIA DE MARKOV

Uma cadeia de Markov é um processo estocástico caracterizado pela perda de memória, a previsão feita do processo conhecendo toda a sua história é tão boa quanto a previsão feita conhecendo-se apenas o processo no presente. De uma maneira precisa, uma cadeia de Markov é uma seqüência de variáveis, X_1, X_2, X_3, ...X_n,......, tal que a distribuição (comportamento) de X_n é determinada pela distribuição de X_{n-1}.

As variáveis X_1, X_2, X_3, ..., tomam valores num conjunto S, chamado espaço de estados. Quando S é finito ou enumerável associa-se à cadeia uma seqüência de matrizes quadradas $P_{(1)}$, $P_{(2)}$, $P_{(3)}$,......, chamadas matrizes de transições da cadeia (matrizes estocásticas). Caso estas matrizes sejam todas iguais, ou seja, as transições constantes, falamos de cadeia homogênea, caso contrário a cadeia é chamada de não-homogênea. A teoria das cadeias homogêneas com espaço de estados finitos ou enumeráveis tem sido largamente estudada, tendo-se uma teoria bem desenvolvida cujas aplicações surgem nas mais variadas áreas tais como na engenharia e economia.

No estudo assintótico de cadeias de Markov, dois tipos de convergência são analisadas: a ergodicidade fraca e a ergodicidade forte. Intuitivamente a cadeia é ergótica fraca se ela tem perda de memória em relação à distribuição de probabilidade inicial e é ergótica forte se ela converge a uma matriz estocástica ou função de transição de probabilidade constante. Naturalmente a ergodicidade forte implica na ergodicidade fraca.

O interesse por Cadeias de Markov não-homogêneas tem sido revigorado nas últimas décadas devido principalmente às suas aplicações na otimização aleatória, especificamente em algoritmos com estrutura Markoviana, como o algoritmo Metropolis Hasting (para geração de amostras de uma dada densidade de difícil tratamento analítico) e o algoritmo simulated annealing (para determinar mínimos globais de uma dada função). Entretanto, resultados teóricos gerais em cadeias de Markov não-homogêneas ainda são escassos e uma teoria mais desenvolvida facilitará futuras aplicações.