

UM MÉTODO PARA A SOLUÇÃO NUMÉRICA  
DE EQUAÇÕES DIFERENCIAIS PARCIAIS ELÍPTICAS

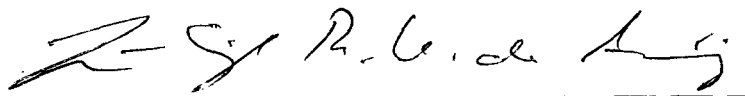
Valter de Senna

TESE SUBMETIDA AO CORPO DOCENTE DA COORDENAÇÃO DOS PROGRAMAS DE  
PÓS-GRADUAÇÃO DE ENGENHARIA DA UNIVERSIDADE FEDERAL DO RIO DE JA-  
NEIRO COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO  
GRAU DE MESTRE EM CIÊNCIA (M.Sc.)

Aprovada por:



Presidente





RIO DE JANEIRO  
ESTADO DA GUANABARA - BRASIL  
MARÇO DE 1974



AGRADECIMENTOS

Desejamos expressar nossa gratidão a todos que de uma forma ou de outra, contribuíram para a realização deste trabalho.

Em especial, queremos agradecer:

- Ao Professor Martin A. Diamond, pela sua paciente e valiosa orientação acadêmica.
- Aos nossos pais, pelo incentivo constante.
- A Miriam P. Carqueja e Conceição M. Almeida pela bem humorada dedicação ao trabalho datilográfico.

SUMÁRIO

Este trabalho é o estudo de um algoritmo, econômico, para a solução de equações diferenciais parciais elípticas. Esse tipo de equações diferenciais dá origem, quando da sua solução numérica, a sistemas de equações lineares de grandes dimensões, cuja solução só é possível ou eficiente, até o presente momento, por métodos iterativos. Entretanto, os métodos em uso exigem uma pré-via estimativa de parâmetros, estimativa essa em geral bastante trabalhosa. O método ora apresentado, no entanto, utiliza informações obtidas no decorrer das iterações para ir aproximando estes parâmetros, de modo a, num procedimento auto-corretivo, acelerar a taxa de convergência, fazendo com que a redução do erro seja aproximadamente igual à obtida quando parâmetros ótimos são empregados desde o início.

ABSTRACT

In this thesis an economic algorithm for the solution of large sparse systems of linear equations which arise in the solution of elliptic partial differential equations is studied. Consideration is focused on systems whose efficient solution is possible, at present, only by iterative methods. The methods in use, however, require a priori estimates of parameters which are difficult or time consuming to estimate. The method presented, on the other hand, uses information obtained during the normal execution of the iteration to approximate these parameters. The procedure is self-correcting in such a way that the convergence is accelerated. The error is reduced at a rate that is approximately equal to the rate obtained using optimal parameters.

ÍNDICE

Dedicatória .....	i
Agradecimentos .....	ii
Sumário .....	iii
Abstract .....	iv
Índice .....	v
 INTRODUÇÃO .....	 1
 CAPÍTULO I	
§ 1 - Apresentação .....	4
§ 2 - Determinação da sequência de parâmetros, $\tau_n$ .....	4
§ 3 - Implementação da sequência $\{\tau_n\}$ .....	7
§ 4 - Maneira usual de implementar o processo ...	9
 CAPÍTULO II	
§ 1 - Apresentação .....	11
§ 2 - Aproximação de um autovalor extremo .....	11
§ 3 - Apresentação de um vetor da forma $Y_N = P_{N,a,b} ((A+B)^{-1}A) y_0$ .....	19
 CAPÍTULO III	
§ 1 - Apresentação .....	21
§ 2 - Definição da sequência de intervalos .....	21
§ 3 - O caso $\mu_{N,i} \in [a_i, b_i]$ .....	24

§ 4 - Definição do vetor $\delta x_{N,i}$ .....	33
§ 5 - Apresentação de um algoritmo para aplicação do método .....	43
APÊNDICE 1 .....	47
APÊNDICE 2 .....	51
APÊNDICE 3 .....	55
REFERÊNCIAS BIBLIOGRÁFICAS .....	58

## INTRODUÇÃO

Uma solução aproximada de equações diferenciais parciais elípticas, por métodos numéricos recai, na maior parte das vezes, na resolução de um sistema de equações lineares, de grandes dimensões. Métodos diretos para a solução desse sistema,  $Ax = q$ , baseados em alguma variação do processo Gaussiano de eliminação são equivalentes a fatorar  $A$  em um produto  $LU$ , de uma matriz triangular inferior  $L$  por uma matriz triangular superior  $U$ . Tais métodos, entretanto, são computacionalmente ineficientes porque os fatores  $L$  e  $U$  não são matrizes esparsas. Por essa razão, um processo iterativo, que apresenta como conveniências adicionais a necessidade muito menor de armazenamento na memória do computador de dados intermediários, bem como a facilidade de programação, é a técnica que permanece como prática e útil.

Dois procedimentos iterativos padrão são "Successive-overrelaxation" (SOR) e "Alternating-direction-implicit" (ADI). Existem, no entanto, graves deficiências em ambos: i) SOR converge bastante lentamente, a menos que o número de equações seja pequeno comparado ao que surge na discretização de um problema de contorno elíptico; ii) ADI tem como principal dificuldade (embora o método seja bastante rápido para uma variada gama de problemas) que não existem princípios matemáticos explícitos que sirvam de guia para a aplicação do método. Depende fundamentalmente da experiência e julgamento do usuário.



A inadequação do SOR e as dificuldades do ADI motivaram várias outras técnicas para a solução do sistema  $Ax = q$ . Um desses métodos iterativos, devido a M. Diamond [1], baseado em técnicas de fatorização, entre as quais de particular interesse é a fatorização simétrica de Stone [2], será o objeto desse trabalho. Nesse método, a matriz  $A$  é substituída por uma matriz  $A+B$  que fature, pela decomposição LU, em matrizes triangulares esparsas  $L$  e  $U$ , isto é,  $A+B = LU$ . A eficiente solução de um sistema da forma  $(A+B)v=r$  que é deste modo equivalente a  $(LU)v=r$  é então utilizada para o cômputo de  $x$  (em  $Ax=q$ ), simplesmente truncando uma seqüência definida por um processo iterativo como

$$(A+B)x_{n+1} = (A+B)x_n - \tau_n (Ax_n - q)$$

De fato, a dificuldade essencial é tornar este procedimento iterativo rapidamente convergente. No capítulo I mostramos que, quando  $A$  e  $A+B$  são positivas definidas, o que acontece quando  $A$  é derivada de muitas das comumente utilizadas discretizações de problemas de contorno, e  $A+B$  é definida por uma das fatorizações devidas a Dupont [3], Kim e Saylor [4] ou Stone [2], a seqüência ótima de parâmetros,  $\tau_n$ , é a seqüência de Tchebychev deduzida a partir dos autovalores extremos de  $(A+B)^{-1}A$ . Parâmetros ótimos dão origem à melhor aproximação (no sentido "minimax") à solução,  $x$ , depois de  $N$  passos. Infelizmente, em geral apenas toscas estimativas,  $a$  e  $b$ , desses autovalores são acessíveis e os parâmetros de Tchebychev assim deduzidos são utilizados para calcular  $x_1, x_2, \dots, x_N$ ; entretanto,  $x_N$  é ótimo apenas quando  $a$  e  $b$  são

os autovalores extremos. Nos capítulos II e III mostraremos que o método por nós utilizado, usando uma iteração do tipo descrito anteriormente ou uma modificação devida a Stiefel [V.A.1] (\*), estima esses autovalores extremos ao mesmo tempo que aproxima a solução do sistema, de modo a ir melhorando o intervalo  $[a,b]$  e originando uma seqüência de intervalos que converge para o intervalo ótimo. Deste modo, a convergência é melhorada a ponto de resultados experimentais figurados em [1] e [4] mostrarem que a redução do erro obtida por esse processo é aproximadamente igual à obtida quando o intervalo ótimo é utilizado desde o início.

Na parte final do nosso trabalho encontram-se alguns apêndices onde são apresentados definições e teoremas que, de uma forma ou outra, são utilizados no texto.

---

(\*) V. A. 1 significa : vide apêndice 1

## CAPÍTULO I

### § 1. Apresentação

Neste capítulo mostraremos que a seqüência ótima de parâmetros,  $\tau_n$ , que faz o procedimento iterativo

$$(A+B)x_{n+1} = (A+B)x_n - \tau_n (Ax_n - q)$$

convergir o mais rapidamente possível é a seqüência de Tchebychev, deduzida a partir dos autovalores extremos [V.A.3, def. 1] da matriz  $(A+B)^{-1}A$ .

Em seguida, daremos dois processos de implementação desta seqüência bem como a maneira usual em que esta implementação é feita.

### § 2. Determinação da Seqüência de Parâmetros, $\tau_n$

Consideremos uma iteração da forma

$$(A+B)x_{n+1} = (A+B)x_n - \tau_n (Ax_n - q) \tag{1.2.1}$$

onde  $A$  e  $(A+B)$  são positivas definidas [V.A.3, def. 2].

Lema 1.1 - Se o erro na  $n$ -ésima iteração da forma (1.2.1) é dado por  $e_n = x - x_n$ , onde  $x$  é a solução exata do sistema  $Ax=q$ , então

$$e_n = \prod_{i=0}^{n-1} (I - \tau_i (A+B)^{-1} A) e_0 ,$$

sendo  $e_0$  o erro inicial.

Demonstração:  $(A+B)x_{n+1} = (A+B)x_n - \tau_n (Ax_n - q)$

$\therefore x_{n+1} = x_n - \tau_n (A+B)^{-1} (Ax_n - q)$ . Daí, temos

$$x_{n+1} - x = x_n - x - \tau_n (A+B)^{-1} (Ax_n - Ax)$$

$$e_{n+1} = e_n - \tau_n (A+B)^{-1} A e_n$$

$$e_{n+1} = (I - \tau_n (A+B)^{-1} A) e_n \text{ e portanto}$$

$$e_{n+1} = \prod_{i=0}^n (I - \tau_i (A+B)^{-1} A) e_0$$

Agora, para fazer a iteração (1.2.1) convergir tão rapidamente quanto possível, devemos minimizar o raio espectral [V.A.3, def. 3] de

$M_n = \prod_{i=0}^{n-1} (I - \tau_i (A+B)^{-1}A)$ . Os autovalores de  $M_n$  são

$$\lambda_k(M_n) = \prod_{i=0}^{n-1} [1 - \tau_i \lambda_k(A+B)^{-1}A)]$$

e como  $A$  e  $(A+B)$  são positivas definidas, os autovalores de  $(A+B)^{-1}A$  são reais, logo  $\lambda_k(M_n)$  é real. Portanto, minimizar o raio espectral de  $M_n$  reduz-se a escolher uma seqüência  $\{\tau_n\}$  que minimize o valor máximo absoluto de

$$P_n(x) = \prod_{i=0}^{n-1} (1 - \tau_i x) \quad (1.2.2)$$

no intervalo  $\lambda_K \leq x \leq \lambda_1$  onde  $\lambda_K = \lambda_{\min}((A+B)^{-1}A)$  e

$$\lambda_1 = \lambda_{\max}((A+B)^{-1}A)$$

NOTA: No restante desta tese, os autovalores de  $(A+B)^{-1}A$  serão representados por

$$\lambda_{\min} = \lambda_K \leq \dots \leq \lambda_2 \leq \lambda_1 = \lambda_{\max}$$

Ora, como qualquer que seja  $n$ ,  $P_n(0) = 1$ , não se faz necessária nenhuma normalização, e o polinômio cujo valor máximo absoluto é mínimo neste intervalo, e que satisfaz  $P_n(0) = 1$  é dado por

$$P_n(x) = \frac{T_n((\lambda_K + \lambda_1 - 2x) / (\lambda_1 - \lambda_K))}{T_n((\lambda_K + \lambda_1) / (\lambda_1 - \lambda_K))} \quad (1.2.3)$$

onde  $T_n$  é o polinômio de Tchebychev de grau  $n$ . [V.A.2, def.1 e teo.3]

### § 3. Implementação da Seqüência $\{\tau_n\}$

Existem duas maneiras de implementar a seqüência  $\{\tau_n\}$  de modo que (1.2.2) é o polinômio dado em (1.2.3). No primeiro método um  $N$  é escolhido e a seqüência  $\{\tau_n\}$ ,  $n=0,1,\dots,N-1$  é computada,

sendo  $\tau_n = \frac{1}{\sigma_n}$  onde  $\sigma_n$  é raiz de  $T_N\left(\frac{\lambda_K + \lambda_1 - 2x}{\lambda_1 - \lambda_K}\right)$ . (Note que

$\frac{1}{\tau_i}$  é raiz de  $P_N(x) = \prod_{i=0}^{N-1} (1 - \tau_i x)$ ). Dizemos, neste caso, que  $\{\tau_n\}$

é uma seqüência de Tchebychev de comprimento  $N$ . Esta seqüência é então usada repetidamente até que a norma do vetor residual,  $r_n = Ax_n - q$ , seja suficientemente pequena. Esse método apresenta inconvenientes.

Primeiro, os  $x_i$  para  $i < N$  não são o resultado de uma seqüência de Tchebychev de comprimento  $i$ . Isto só é verdade para  $i=N$ . Além disso, se decidirmos que mais iterações são necessárias, não podemos simplesmente computar iterações sucessivas, pois não serão o resultado de uma seqüência de parâmetros de Tchebychev.

O outro método, devido a Stiefel [V.A.1], e que alia a vantagem de cada  $x_i$  ser o resultado de uma seqüência de comprimento  $i$ ,

ser mais estável que o anterior [4], é definido da seguinte maneira:

$$x_{n+1} = x_n + \delta x_n \quad (1.3.1)$$

$$\delta x_n = \frac{4T_n(y)}{(\lambda_1 - \lambda_K) T_{n+1}(y)} v_n + \frac{T_{n-1}(y)}{T_{n+1}(y)} \delta x_{n-1}$$

$$\delta x_0 = \frac{2}{(\lambda_1 + \lambda_K)} v_0$$

onde

$$v_n = (A+B)^{-1} r_n \quad (1.3.2)$$

$$r_n = q - Ax_n$$

e

$$y = \frac{(\lambda_1 + \lambda_K)}{(\lambda_1 - \lambda_K)}$$

O método acima, pelas vantagens citadas, será o importante para nosso estudo.

§ 4. Maneira usual de implementar o processo

Em geral,  $\lambda_1$  e  $\lambda_K$  não são conhecidos ou mesmo acessíveis e, como se fazem necessários tanto em um método como no outro, a minimização de Tchebychev é normalmente implementada encontrando-se um intervalo  $[a,b]$  contendo  $[\lambda_K, \lambda_1]$ , e usando-se  $a$  e  $b$  ao invés de  $\lambda_K$  e  $\lambda_1$ , respectivamente. Ficamos então com: no lugar de (1.2.2),

$$P_N(x) = \prod_{i=0}^{N-1} (1 - \tau_i x) = \frac{T_N\left(\frac{a+b-2x}{b-a}\right)}{T_N\left(\frac{a+b}{b-a}\right)} \quad (1.4.1)$$

e no lugar de (1.3.1),

$$x_{n+1} = x_n + \delta x_n \quad (1.4.2)$$

$$\delta x_n = \frac{4 T_n(y)}{(b-a) T_{n+1}(y)} v_n + \frac{T_{n-1}(y)}{T_{n+1}(y)} \delta x_{n-1}$$

$$\delta x_0 = \frac{2}{(b+a)} v_0$$

onde

$$v_n = (A+B)^{-1} r_n$$

$$r_n = q - Ax_n$$



e

$$y = \frac{(b+a)}{(b-a)}$$

Crítica: além de ser este intervalo  $[a,b]$ , às vezes, bastante difícil de encontrar, a seqüência deduzida utilizando-o pode estar longe de ser ótima. Um procedimento devido a M.A. Diamond [1] e [4], no entanto, possibilita-nos aproximar estes autovalores extremos,  $\lambda_1$  e  $\lambda_K$ , à medida que fazemos a iteração (1.4.1) ou (1.4.2), e aproximamos o valor de  $x$ . A descrição deste processo e a demonstração de sua convergência é objeto desta tese.

CAPÍTULO II§ 1. Apresentação

Neste capítulo mostraremos que um vetor da forma

$$y_N = P_{N,a,b} ((A+B)^{-1}A) y_0$$

onde

$$P_{N,a,b}(x) = \frac{T_N\left(\frac{a+b-2x}{b-a}\right)}{T_N\left(\frac{a+b}{b-a}\right)},$$

sob determinadas condições, possibilita-nos obter uma aproximação para um autovalor extremo. Em seguida, provaremos que o vetor  $v_n$  definido em (1.4.2) é dessa forma. Usaremos este resultado no capítulo III para obtermos uma seqüência de intervalos que converge para o intervalo ótimo  $[\lambda_K, \lambda_1]$ .

§ 2. Aproximação de um autovalor extremo.

Lema 2.1 - Se  $[\lambda_K, \lambda_1] \not\subset [a, b]$  e  $\lambda_j$  é tal que

$$|P_{N,a,b}(\lambda_j)| > |P_{N,a,b}(\lambda_i)|; \forall i; i \neq j,$$

então:

- i)  $\lambda_j \notin [a, b]$
- ii)  $\lambda_j = \lambda_1$  ou  $\lambda_j = \lambda_K$

Demonstração: Conseqüência imediata das propriedades dos polinômios de Tchebychev [V.A.2, teo 4].

Corolário 2.1 -  $P_{N,a,b}(\lambda_j) \neq 0$

Demonstração: De fato, pelo lema anterior,  $\lambda_j \notin [a,b]$

Teorema 2.1 - Seja

$$y_N = P_{N,a,b}((A+B)^{-1}A)y_0 ; \quad a < b,$$

onde

$$y_0 = \sum_{k=1}^K \alpha_k w_k \quad (2.2.1)$$

sendo  $\{w_1, w_2, \dots, w_K\}$  um conjunto completo de autovetores de  $(A+B)^{-1}A$ , ortonormal em relação a  $(A+B)$  [V.A.3, teo 2]. Suponha que  $\alpha_j \neq 0$ ,  $[\lambda_K, \lambda_1] \not\subset [a,b]$  e

$$|P_{N,a,b}(\lambda_j)| > |P_{N,a,b}(\lambda_i)| ; \quad \forall i, i \neq j .$$

Então,

$$\mu_N = \frac{\langle Ay_N, y_N \rangle}{\langle (A+B)y_N, y_N \rangle} \quad \text{aproxima } \lambda_j .$$

Demonstração: Substituindo  $y_0$  pelo seu valor, dado em (2.2.1), vem

$$y_N = P_N((A+B)^{-1}A) \sum_{k=1}^K \alpha_k w_k =$$

$$\begin{aligned}
 &= \sum_{k=1}^K \alpha_k P_N (A+B)^{-1} A w_k = \\
 &= \sum_{k=1}^K \alpha_k P_N (\lambda_k) w_k
 \end{aligned}$$

Comentário: note que  $\alpha_i P_N (\lambda_i)$  é, portanto, a  $i$ -ésima coordenada, na base  $\{w_1, w_2, \dots, w_K\}$ , do vetor  $y_N$ .

Temos então

$$\mu_N = \frac{\left\langle A \sum_{k=1}^K \alpha_k P_N (\lambda_k) w_k ; \sum_{k=1}^K \alpha_k P_N (\lambda_k) w_k \right\rangle}{\left\langle (A+B) \sum_{k=1}^K \alpha_k P_N (\lambda_k) w_k ; \sum_{k=1}^K \alpha_k P_N (\lambda_k) w_k \right\rangle}$$

Agora, substituindo  $A$  por  $(A+B)(A+B)^{-1}A$ , vem

$$\begin{aligned}
 \mu_N &= \frac{\left\langle (A+B)(A+B)^{-1}A \sum_{k=1}^K \alpha_k P_N (\lambda_k) w_k ; \sum_{k=1}^K \alpha_k P_N (\lambda_k) w_k \right\rangle}{\left\langle (A+B) \sum_{k=1}^K \alpha_k P_N (\lambda_k) w_k ; \sum_{k=1}^K \alpha_k P_N (\lambda_k) w_k \right\rangle} = \\
 &= \frac{\left\langle \sum_{k=1}^K \alpha_k P_N (\lambda_k) (A+B)^{-1}A w_k ; \sum_{k=1}^K \alpha_k P_N (\lambda_k) (A+B) w_k \right\rangle}{\left\langle \sum_{k=1}^K \alpha_k P_N (\lambda_k) w_k ; \sum_{k=1}^K \alpha_k P_N (\lambda_k) (A+B) w_k \right\rangle} =
 \end{aligned}$$

$$\begin{aligned}
&= \frac{\left\langle \sum_{k=1}^K \alpha_k P_N(\lambda_k) \lambda_k w_k ; \sum_{k=1}^K \alpha_k P_N(\lambda_k) (A+B) w_k \right\rangle}{\left\langle \sum_{k=1}^K \alpha_k P_N(\lambda_k) w_k ; \sum_{k=1}^K \alpha_k P_N(\lambda_k) (A+B) w_k \right\rangle} = \\
&= \frac{\sum_{i=1}^K \left\langle \alpha_i P_N(\lambda_i) \lambda_i w_i ; \sum_{k=1}^K \alpha_k P_N(\lambda_k) (A+B) w_k \right\rangle}{\sum_{i=1}^K \left\langle \alpha_i P_N(\lambda_i) w_i ; \sum_{k=1}^K \alpha_k P_N(\lambda_k) (A+B) w_k \right\rangle} = \\
&= \frac{\sum_{i=1}^K \alpha_i^2 \lambda_i P_N^2(\lambda_i)}{\sum_{i=1}^K \alpha_i^2 P_N^2(\lambda_i)} = \\
&= \frac{\lambda_j + \sum_{\substack{i=1 \\ i \neq j}}^K \lambda_i \left[ \frac{\alpha_i P_N(\lambda_i)}{\alpha_j P_N(\lambda_j)} \right]^2}{1 + \sum_{\substack{i=1 \\ i \neq j}}^K \left[ \frac{\alpha_i P_N(\lambda_i)}{\alpha_j P_N(\lambda_j)} \right]^2} \quad (2.2.2)
\end{aligned}$$

O erro absoluto  $\epsilon_N = (\lambda_j - \mu_N) \bar{\epsilon}$

$$\epsilon_N = \frac{\sum_{\substack{i=1 \\ i \neq j}}^K (\lambda_j - \lambda_i) \left[ \frac{\alpha_i P_N(\lambda_i)}{\alpha_j P_N(\lambda_j)} \right]^2}{1 + \sum_{\substack{i=1 \\ i \neq j}}^K \left[ \frac{\alpha_i P_N(\lambda_i)}{\alpha_j P_N(\lambda_j)} \right]^2} \quad (2.2.3)$$

Veremos agora que  $\varepsilon_N \rightarrow 0$  quando  $N \rightarrow +\infty$ , de modo que  $\mu_N \rightarrow \lambda_j$ .

De fato,

$$\varepsilon_N \leq \left| \sum_{i=1}^K (\lambda_j - \lambda_i) \left[ \frac{\alpha_i P_N(\lambda_i)}{\alpha_j P_N(\lambda_j)} \right]^2 \right| \leq$$

$$\leq \left| \sum_{i \in J} (\lambda_j - \lambda_i) \left[ \frac{\alpha_i P_N(\lambda_i)}{\alpha_j P_N(\lambda_j)} \right]^2 \right| + \quad (2.2.4)$$

$$+ \left| \sum_{i \in H} (\lambda_j - \lambda_i) \left[ \frac{\alpha_i P_N(\lambda_i)}{\alpha_j P_N(\lambda_j)} \right]^2 \right| \quad (2.2.5)$$

sendo

$$J = \{i / \lambda_i \in [a, b]\} \text{ e}$$

$$H = \{i / \lambda_i \notin [a, b] ; i \neq j\} .$$

Agora,

$$\left| \sum_{i \in J} (\lambda_j - \lambda_i) \left[ \frac{\alpha_i P_N(\lambda_i)}{\alpha_j P_N(\lambda_j)} \right]^2 \right| \leq$$

$$\leq \sum_{i \in J} |\lambda_j - \lambda_i| \left( \frac{\alpha_i}{\alpha_j} \right)^2 \left( \frac{P_N(\lambda_i)}{P_N(\lambda_j)} \right)^2 =$$

$$= \sum_{i \in J} \left[ c_i \frac{P_N(\lambda_i)}{P_N(\lambda_j)} \right]^2 \quad \text{onde}$$

$$c_i = \left[ \left( \frac{\alpha_i}{\alpha_j} \right)^2 |\lambda_j - \lambda_i| \right]^{1/2}$$

Vemos portanto que a expressão dada em (2.2.4) tende a zero quando  $N$  cresce, pois os termos  $c_i \frac{P_N(\lambda_i)}{P_N(\lambda_j)}$  estão decrescendo a uma taxa ótima [V.A.2, teo.1 e Cor. 1].

Em relação à expressão dada em (2.2.5) temos, analogamente,

$$\left| \sum_{i \in H} (\lambda_j - \lambda_i) \left[ \frac{\alpha_i P_N(\lambda_i)}{\alpha_j P_N(\lambda_j)} \right]^2 \right| \leq$$

$$\leq \sum_{i \in H} \left[ c_i \frac{P_N(\lambda_i)}{P_N(\lambda_j)} \right]^2 . \quad \text{Entretanto, como } \lambda_i \notin [a, b], \text{ não po}$$

demos utilizar o resultado aplicado no caso da expressão (2.2.4) . Isto contudo não se constitui em obstáculo intransponível, uma vez que

$$\begin{aligned}
\left| \frac{P_N(\lambda_i)}{P_N(\lambda_j)} \right| &= \left| \frac{T_N\left(\frac{a+b-2\lambda_i}{b-a}\right)}{T_N\left(\frac{a+b-2\lambda_j}{b-a}\right)} \right| = \\
&= \frac{\cosh N \cosh^{-1} \left| \frac{a+b-2\lambda_i}{b-a} \right|}{\cosh N \cosh^{-1} \left| \frac{a+b-2\lambda_j}{b-a} \right|} = \\
&= \frac{e^{N \cosh^{-1} \left| \frac{a+b-2\lambda_i}{b-a} \right|} + e^{-N \cosh^{-1} \left| \frac{a+b-2\lambda_i}{b-a} \right|}}{e^{N \cosh^{-1} \left| \frac{a+b-2\lambda_j}{b-a} \right|} + e^{-N \cosh^{-1} \left| \frac{a+b-2\lambda_j}{b-a} \right|}}
\end{aligned}$$

e, como

$$\cosh^{-1} \left| \frac{a+b-2\lambda_i}{b-a} \right| - \cosh^{-1} \left| \frac{a+b-2\lambda_j}{b-a} \right| < 0$$

pois

$$1 < \left| \frac{a+b-2\lambda_i}{b-a} \right| < \left| \frac{a+b-2\lambda_j}{b-a} \right| ; \text{ [V.A.2 , teo 4]}$$

segue-se que

$$\lim_{N \rightarrow +\infty} \frac{P_N(\lambda_i)}{P_N(\lambda_j)} = 0$$



Assim, obtemos que  $\epsilon_N \rightarrow 0$  quando  $N \rightarrow +\infty$  e, por conseguinte,  $\mu_N \rightarrow \lambda_j$ .

Comentários:

i) observe que a única coisa realmente necessária para

$y_N = f_N((A+B)^{-1}A)y_0$  fornecer uma aproximação a  $\lambda_j$  é que

$$\lim_{N \rightarrow +\infty} \frac{f_N(\lambda_i)}{f_N(\lambda_j)} = 0 ; \quad \forall i \neq j; \quad \text{desde que } \alpha_j \neq 0.$$

ii) quanto mais próximo de 1 estiver  $\left| \frac{a+b-2\lambda_j}{b-a} \right|$ , ou seja,

quanto melhor for a aproximação  $[a,b]$  a  $[\lambda_K, \lambda_1]$ , maior

terá que ser o  $N$  para que  $\frac{P_N(\lambda_i)}{P_N(\lambda_j)}$  seja aproximadamente

igual a zero.

Um fato importante e que transparece na expressão (2.2.3) é o seguinte:

Corolário 2.2 -  $\forall N, \mu_N \in [\lambda_K, \lambda_1]$

Demonstração: Basta notar que se  $\lambda_j = \lambda_1$ , então  $\epsilon_N > 0$ , e se  $\lambda_j = \lambda_K$ , então  $\epsilon_N < 0$ .

Esse resultado, na verdade, é mais geral como revela um exame mais detido da maneira como a expressão (2.2.3) foi deduzida. Realmente,  $\alpha_i P_N(\lambda_i)$  é simplesmente a  $i$ -ésima coordenada,

na base  $\{w_1, w_2, \dots, w_K\}$  do vetor  $y_N$ . De modo que, se tomamos um vetor qualquer  $y$  e escrevemos como combinação linear dos vetores da base  $\{w_1, w_2, \dots, w_K\}$ ,  $y = \sum_{k=1}^K \alpha_k w_k$ , calculando o quociente

$$\mu = \frac{\langle AY, Y \rangle}{\langle (A+B)Y, Y \rangle}, \text{ obtemos}$$

$$\mu \in [\lambda_K, \lambda_1], \text{ pois } \epsilon = \lambda_j - \mu =$$

$$= \frac{\sum_{\substack{i=1 \\ i \neq j}}^K (\lambda_j - \lambda_i) \left[ \frac{\alpha_i}{\alpha_j} \right]^2}{1 + \sum_{\substack{i=1 \\ i \neq j}}^K \left[ \frac{\alpha_i}{\alpha_j} \right]^2} \quad e$$

$$\epsilon \geq 0 \quad \text{se } \lambda_j = \lambda_1 \quad e$$

$$\epsilon \leq 0 \quad \text{se } \lambda_j = \lambda_K$$

§ 3. Apresentação de um vetor da forma  $y_N = P_{N,a,b} ((A+B)^{-1}A)y_0$

Teorema 2.2 - O vetor  $v_n = (A+B)^{-1}r_n$ , definido em (1.4.2),

satisfaz  $v_n = P_{n,a,b} ((A+B)^{-1}A)v_0$ .

Demonstração:

$$\begin{aligned} v_n &= (A+B)^{-1} r_n = (A+B)^{-1} (q - Ax_n) = \\ &= (A+B)^{-1} (Ax - Ax_n) = (A+B)^{-1} A e_n \quad e, \end{aligned}$$

como  $e_n = x - x_n = P_{n,a,b} ((A+B)^{-1} A) e_0$

vem

$$\begin{aligned} v_n &= (A+B)^{-1} A P_{n,a,b} ((A+B)^{-1} A) e_0 = \\ &= P_{n,a,b} ((A+B)^{-1} A) (A+B)^{-1} A e_0 = \\ &= P_{n,a,b} ((A+B)^{-1} A) (A+B)^{-1} A (x - x_0) = \\ &= P_{n,a,b} ((A+B)^{-1} A) (A+B)^{-1} (q - Ax_0) = \\ &= P_{n,a,b} ((A+B)^{-1} A) v_0 . \end{aligned}$$

CAPÍTULO III§ 1. Apresentação

Neste capítulo, utilizando os resultados do capítulo anterior, apresentaremos um método devido a M. Diamond de, utilizando a iteração (1.4.2), resolver o sistema  $Ax = q$ , e simultaneamente melhorar o intervalo  $[a, b]$ . Neste método, um número  $N$  é escolhido e  $N$  iterações do tipo (1.4.2) são realizadas com cada um dos intervalos,  $[a_i, b_i]$ , de uma seqüência de intervalos que converge para o intervalo ótimo  $[\lambda_K, \lambda_1]$ .

A parte inicial deste capítulo consiste, com ligeiras modificações, no parágrafo 4 e parte do parágrafo 5 de um artigo, ainda não publicado, de Diamond.

§ 2. Definição da seqüência de intervalos

Os vetores calculados utilizando o intervalo  $[a_i, b_i]$  serão identificados por um segundo índice. Assim, os vetores calculados na  $(n+1)$  - ésima iteração da forma (1.4.2), com o intervalo  $[a_i, b_i]$ , serão representados por :

$$r_{n,i} = q - Ax_{n,i}$$

(3.2.1)

$$v_{n,i} = (A+B)^{-1} r_{n,i}$$

$$\delta x_{n,i} = \frac{4T_n(y_i)}{(b_i - a_i)T_{n+1}(y_i)} v_{n,i} + \frac{T_{n-1}(y_i)}{T_{n+1}(y_i)} \delta x_{n-1,i}$$

onde

$$y_i = \frac{(a_i + b_i)}{(b_i - a_i)}$$

e

$$x_{n+1,i} = x_{n,i} + \delta x_{n,i}$$

Ora, pelos teoremas 2.1 e 2.2 do capítulo anterior, se  $[\lambda_K, \lambda_1] \not\subset [a_i, b_i]$  e

$$|P_{N,a_i,b_i}(\lambda_j)| > |P_{N,a_i,b_i}(\lambda_k)| \quad ; \quad \forall k, k \neq j,$$

então quando  $N$  cresce

$$\mu_{N,i} = \frac{\langle Av_{N,i} ; v_{N,i} \rangle}{\langle (A+B)v_{N,i} ; v_{N,i} \rangle}$$

aproxima  $\lambda_j$ . Ora, pelo lema 2.1,  $\lambda_j = \lambda_1$  ou  $\lambda_j = \lambda_K$ , de modo que podemos afirmar que quando  $N$  cresce,  $\mu_{N,i}$  é ou uma aproximação a  $\lambda_1$  ou uma aproximação a  $\lambda_K$ , dependendo de

$$|P_{N,a_i,b_i}(\lambda_1)| > |P_{N,a_i,b_i}(\lambda_K)| \quad \text{ou} \quad |P_{N,a_i,b_i}(\lambda_K)| > |P_{N,a_i,b_i}(\lambda_1)|$$

Comentário: o caso  $|P_{N,a_i,b_i}(\lambda_1)| = |P_{N,a_i,b_i}(\lambda_K)|$ , que corresponde ao caso  $a_i - \lambda_K = \lambda_1 - b_i$ , será discutido mais adiante, no parágrafo 3.

Assim, o  $i$ -ésimo intervalo é melhorado substituindo-se  $a_i$  por  $\mu_{N,i}$  se  $\mu_{N,i}$  aproxima o menor autovalor, e substituindo  $b_i$  por  $\mu_{N,i}$  se a aproximação é ao maior autovalor. Assumindo que  $[a_i, b_i] \subset [\lambda_K, \lambda_1]$ , concluímos que  $\mu_{N,i}$  aproxima  $\lambda_K$  se  $\mu_{N,i} < a_i$  e aproxima  $\lambda_1$  se  $\mu_{N,i} > b_i$ . O caso  $\mu_{N,i} \in [a_i, b_i]$  que guarda relação com o caso  $|P_{N,a_i,b_i}(\lambda_1)| = |P_{N,a_i,b_i}(\lambda_K)|$  será objeto do parágrafo 3.

Mostraremos agora como definir o intervalo  $[a_1, b_1]$  de modo que  $[a_i, b_i] \subset [\lambda_K, \lambda_1]$ ,  $\forall i$ . Em primeiro lugar, note que se  $A$  e  $A+B$  são positivas definidas, pelo corolário 2,2,  $\mu_{N,i} \in [\lambda_K, \lambda_1]$ . Então, se  $[a_i, b_i] \subset [\lambda_K, \lambda_1]$  e  $[a_{i+1}, b_{i+1}]$  é definido substituindo-se  $a_i$  ou  $b_i$  por  $\mu_{N,i}$ , segue-se que  $[a_{i+1}, b_{i+1}] \subset [\lambda_K, \lambda_1]$ .

Portanto, os intervalos satisfazem  $[a_i, b_i] \subset [\lambda_K, \lambda_1]$  se  $[a_1, b_1] \subset [\lambda_K, \lambda_1]$ .

Tal intervalo inicial,  $[a_1, b_1]$ , pode ser determinado da seguinte maneira: seja  $K$  a ordem de  $A$  e  $e_k$ ,  $k = 1, 2, \dots, K$  a  $k$ -ésima coluna da matriz identidade  $K \times K$ .

Defina  $d_k = \frac{\langle Ae_k, e_k \rangle}{\langle (A+B)e_k, e_k \rangle}$ . Conforme comentário relativo ao corolário 2.2,  $d_k \in [\lambda_K, \lambda_1]$ . Fazendo  $a_1 = \min_k d_k$  e  $b_1 = \max_k d_k$ , obtemos um intervalo  $[a_1, b_1]$  contido em  $[\lambda_K, \lambda_1]$ .

Note que  $d_k$  é a razão entre o  $k$ -ésimo elemento da diagonal de  $A$  e o  $k$ -ésimo elemento da diagonal de  $(A+B)$ . É, portanto, facilmente computável.

Se ocorrer  $a_1 = b_1$  podemos redefinir o intervalo inicial do seguinte modo: escolha um elemento de  $A$ , digamos  $(A)_{jk}$ , tal que  $(A)_{jk} = 0$  mas  $(A+B)_{jk} \neq 0$ . Defina

$$c = \frac{\langle A(e_j + e_k); (e_j + e_k) \rangle}{\langle (A+B)(e_j + e_k); (e_j + e_k) \rangle} \quad (3.2.2)$$

e

$$a_1' = \min(a_1, c) ; b_1' = \max(b_1, c).$$

### § 3. O caso $\mu_{N,i} \in [a_i, b_i]$

Vamos estudar agora o caso  $a_i \leq \mu_{N,i} \leq b_i$ . Se esta situação ocorre, então ou:

- 1) o número de iterações,  $N$ , não é grande o bastante para que a convergência tenha ocorrido; ou
- 2)  $P_{N, a_i, b_i}(\lambda_K) \cong |P_{N, a_i, b_i}(\lambda_1)|$

Em geral, podemos distinguir estes dois casos da seguinte maneira: se

$$\mu_{n,i} = \frac{\langle Av_{n,i}; v_{n,i} \rangle}{\langle (A+B)v_{n,i}; v_{n,i} \rangle},$$

calculado para cada  $n$ , formar uma seqüência crescente (ou decrescente) concluímos que  $N$  é pequeno e que fazendo mais iterações faremos  $\mu_{n,i}$  crescer ultrapassando  $b_i$  (ou decrescer ultrapassando  $a_i$ ). Se, por outro lado, os  $\mu_{n,i}$  estão oscilando, concluímos que  $P_{N,a_i,b_i}(\lambda_K) \cong P_{N,a_i,b_i}(\lambda_1)$ .

No caso 1, simplesmente continuamos iterando, portanto aumentando  $N$ .

Comentário: note aqui a vantagem da iteração (1.4.2) sobre (1.4.1) ao permitir aumentar  $N$  sem precisar redefinir a seqüência  $\{\tau_n\}$ .

Mostraremos agora que o caso 2 eventualmente ocorre em toda aplicação do método.

Teorema 3.1 - Seja  $y_0 = \sum_{k=1}^K \alpha_k w_k$  onde  $\{w_k\}$  é um conjunto completo de autovetores de  $(A+B)^{-1}A$  e  $\alpha_k \alpha_1 \neq 0$ .

Suponha  $[a_1, b_1] \subset [\lambda_K, \lambda_1]$ , e



$$[a_{i+1}, b_{i+1}] = \begin{cases} [\mu_{N,i}, b_i] & \text{se } \mu_{N,i} < a_i \\ [a_i, b_i] & \text{se } a_i \leq \mu_{N,i} \leq b_i \\ [a_i, \mu_{N,i}] & \text{se } \mu_{N,i} > b_i \end{cases}$$

onde

$$\mu_{N,i} = \frac{\langle Ay_{N,i}, y_{N,i} \rangle}{\langle (A+B)y_{N,i}, y_{N,i} \rangle}$$

e

$$y_{N,i} = \prod_{j=1}^i P_{N,a_j,b_j} ((A+B)^{-1}A)y_0$$

Então, quando  $i \rightarrow +\infty$ ,

$$[a_i, b_i] \rightarrow [a, b] \quad \text{tal que}$$

$$a - \lambda_K = \lambda_1 - b$$

Demonstração: Como foi dito no parágrafo anterior,

$$[a_i, b_i] \subset [a_{i+1}, b_{i+1}] \subset [\lambda_K, \lambda_1].$$

Portanto,  $[a_i, b_i]$  tende para algum limite, digamos  $[a, b]$ . Suponha que  $a - \lambda_K < \lambda_1 - b$ . Daí,  $[P_{N,a,b}(\lambda_1)] > [P_{N,a,b}(\lambda_k)]$ ;  $k=2, \dots, K$ .

$$\text{Agora, } \lim_{i \rightarrow \infty} \frac{P_{N, a_i, b_i}(\lambda_k)}{P_{N, a_i, b_i}(\lambda_1)} = \frac{P_{N, a, b}(\lambda_k)}{P_{N, a, b}(\lambda_1)} ; \forall k$$

$$\text{e como } \frac{P_{N, a, b}(\lambda_k)}{P_{N, a, b}(\lambda_1)} < 1 ; k=2, \dots, K$$

segue-se que existe  $I_0$ , tal que  $\forall i > I_0$  e  $k=2, \dots, K$ ,

$$\frac{P_{N, a_i, b_i}(\lambda_k)}{P_{N, a_i, b_i}(\lambda_1)} < C < 1$$

$$\lim_{i \rightarrow \infty} \frac{\prod_{j=1}^i P_{N, a_j, b_j}(\lambda_k)}{\prod_{j=1}^i P_{N, a_j, b_j}(\lambda_1)} = 0 ; k=2, \dots, K$$

e portanto  $\lim_{i \rightarrow +\infty} \mu_{N, i} = \lambda_1$ . Concluimos então que  $b = \lambda_1$  e a suposição inicial  $a - \lambda_K < \lambda_1 - b$  é absurda, uma vez que  $\lambda_1 - b = 0$ . Analogamente, obtemos uma contradição ao supormos  $a - \lambda_K > \lambda_1 - b$ .

Uma vez tendo ocorrido que  $a_i - \lambda_K \cong \lambda_1 - b_i$  podemos melhorar ambos os extremos do intervalo sempre que um deles é melhorado.

De fato, se  $a_{i+1} - \lambda_K < a_i - \lambda_K$  então  $\lambda_1 - (b_i + a_i - a_{i+1}) < \lambda_1 - b_i$ .

Comentário: uma vez tendo acontecido o caso 2, usando o artifício acima, este caso fatalmente continuará ocorrendo.

Estudaremos agora que medidas devem ser tomadas após a ocorrência do caso 2. Consideremos dois sub-casos:

$$2a) \quad a_i \cong \lambda_K \quad e \quad b_i \cong \lambda_1$$

$$2b) \quad a_i \not\cong \lambda_K \quad e \quad b_i \not\cong \lambda_1$$

Considere 2a. Se  $a_i = \lambda_K$  e  $b_i = \lambda_1$  então

$$P_{N, a_i, b_i}(\lambda_K) = |P_{N, a_i, b_i}(\lambda_1)|.$$

Portanto, se o intervalo é melhorado a ponto de  $[a_i, b_i] \cong [\lambda_K, \lambda_1]$ ,

$P_{N, a_i, b_i}(\lambda_K) \cong |P_{N, a_i, b_i}(\lambda_1)|$ . Neste caso, o intervalo é quase

ótimo e pode ser usado nas iterações subseqüentes sem prejuízo.

No caso 2b a situação é esquematizada na figura 1.

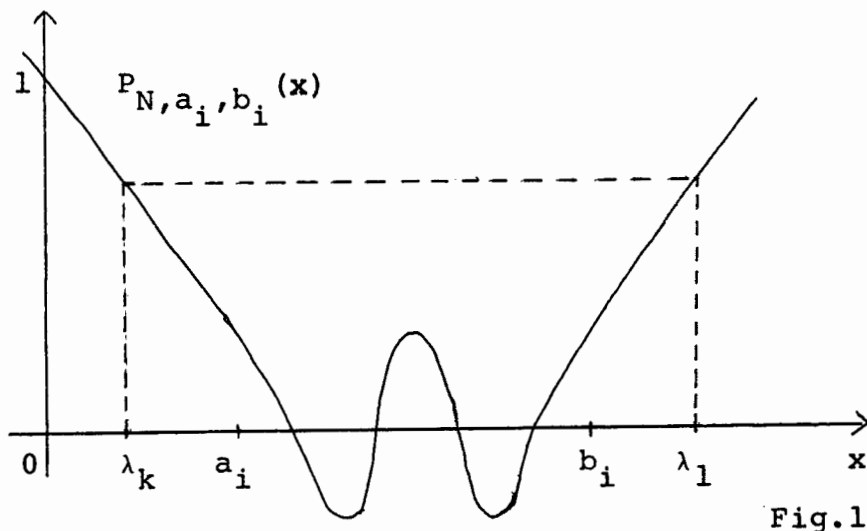


Fig.1

A convergência de  $\{x_{n,i}\}$  poderá ser lenta se o intervalo  $[a_i, b_i]$  for apenas uma estimativa grosseira de  $[\lambda_K, \lambda_1]$ . Além disso, note que simplesmente aumentar o grau de  $P_{N,a_i,b_i}$ , isto é, fazer mais iterações, não garantirá a convergência da seqüência  $\{\mu_{N,i}\}$  para  $\lambda_1$  ou  $\lambda_K$ .

Concluimos então que o intervalo não poderá mais ser melhorado utilizando-se o método descrito anteriormente.

Vamos agora descrever uma técnica que forçará a convergência de  $\{\mu_{N,i}\}$  quando  $P_{N,a_i,b_i}(\lambda_K) \approx |P_{N,a_i,b_i}(\lambda_1)|$ . A idéia é conseguir um polinômio  $P_{N,a_i,b_i}^*(x)$  tal que

$$P_{N,a_i,b_i}^*(\lambda_K) \approx |P_{N,a_i,b_i}^*(\lambda_1)|$$

e que obviamente continuemos a ter um vetor da forma

$$y_N = P_{N,a_i,b_i}^* ((A+B)^{-1}A) y_0$$

com  $\frac{\langle Ay_N, y_N \rangle}{\langle (A+B)y_N, y_N \rangle}$  aproximando um autovalor extremo. Isto não

é difícil. Mostraremos a seguir que, fazendo uma iteração da forma

$$(A+B)x_{n+1} = (A+B)x_n - (1/b_i)(Ax_n - q), \quad (3.3.1)$$

obtemos um polinômio  $P_{N,a_i,b_i}^*$  com as características desejadas.

Suponhamos que  $N$  iterações foram feitas com o intervalo  $[a_i, b_i]$  e chegamos a

$$e_{N,i} = P_{N,a_i,b_i} ((A+B)^{-1}A) e_{N,i-1} .$$

Então, uma iteração da forma

$$(A+B)x_{N+1,i} = (A+B)x_{N,i} - \frac{1}{b_i} (Ax_{N,i} - q)$$

fornece

$$\begin{aligned} e_{N+1,i} &= (I - \frac{1}{b_i} (A+B)^{-1}A) e_{N,i} = \\ &= (I - \frac{1}{b_i} (A+B)^{-1}A) P_{N,a_i,b_i} ((A+B)^{-1}A) e_{N,i-1} = \\ &= P_{N+1,a_i,b_i}^* ((A+B)^{-1}A) e_{N,i-1} \end{aligned} \quad (3.3.2)$$

com  $P_{N+1,a_i,b_i}^*(x) = (1 - \frac{x}{b_i}) P_{N,a_i,b_i}(x)$ .

É fácil ver que

$$P_{N+1,a_i,b_i}^*(\lambda_K) \neq |P_{N+1,a_i,b_i}^*(\lambda_1)|$$

O vetor  $v_{N+1,i}$  passa a ser

$$v_{N+1,i} = \left( I - \frac{1}{b_i} (A+B)^{-1} A \right) P_{N,a_i,b_i} \left( (A+B)^{-1} A \right) v_{N,i-1} =$$

$$= P_{N+1,a_i,b_i}^* \left( (A+B)^{-1} A \right) v_{N,i-1}.$$

Calcula-se a seguir

$$\mu_{N+1,i} = \frac{\langle Av_{N+1,i} ; v_{N+1,i} \rangle}{\langle (A+B)v_{N+1,i} ; v_{N+1,i} \rangle}$$

e melhora-se o intervalo, se

$$\mu_{N+1,i} \notin [a_i, b_i]$$

Suponha agora que

$$\mu_{N+1,i} \in [a_i, b_i]$$

Duas alternativas se nos colocam, uma vez que continuaremos a usar o intervalo  $[a_i, b_i]$ . A primeira delas é agir como se tivéssemos mudado o intervalo, ou seja, definir

$$[a_{i+1}, b_{i+1}] = [a_i, b_i] ;$$

$$x_{0,i+1} = x_{N+1,i} ;$$

$$\delta x_{0,i+1} = \frac{2}{(a_{i+1} + b_{i+1})} v_{0,i+1} \quad \text{etc.}$$

A segunda alternativa é continuar calculando  $x_{N+2,i}$ ,  $x_{N+3,i}, \dots, x_{2N+1,i}$ . Usaremos esta segunda alternativa e além disso o cálculo de  $\mu_{N+1,i}$  será omitido.  $x_{N+1,i}$  é definido usando (3.3.1) e  $x_{N+2,i}, \dots, x_{2N+1,i}$  calculados em seguida.

Isto é feito pelas seguintes razões:

- i) passaremos a ter polinômios de Tchebychev de graus maiores.
- ii) um autovalor aproximado,  $\mu$ , é calculado apenas uma vez a cada  $N$  iterações, e não duas.
- iii) é de se esperar a ocorrência do caso 2a que não é em nada beneficiado pelo cálculo suplementar de  $\mu_{N+1,i}$ .

Um problema entretanto surge agora. Os vetores  $x_{N+2,i}, \dots, x_{2N+1,i}$  não podem ser computados simplesmente reiniciando a iteração (3.2.1) pois  $\delta x_{N,i}$  não está definido. De fato, calcular  $x_{N+2,i}$  exige  $x_{N+1,i}$  e  $\delta x_{N,i}$ . Primeiro  $\delta x_{N+1,i}$  é calculado e então fazemos  $x_{N+2,i} = x_{N+1,i} + \delta x_{N+1,i}$ . Se a iteração (3.2.1) é usada para calcular  $x_{1,i}, \dots, x_{N,i}$ , o último vetor a ser calculado é  $\delta x_{N-1,i}$ .

Se  $x_{N+1,i}$  é calculado usando (3.3.1), então é necessário definir  $\delta x_{N,i}$  antes de reiniciar (3.2.1). É claro que o vetor  $\delta x_{N,i}$  deve

ser definido de modo que a iteração  $x_{N+2,i}, x_{N+3,i}, \dots$  convirja a  $x$  e que  $v_{N+2,i}, v_{N+3,i}, \dots$  forneça a aproximação a um autovalor extremo. Mostraremos a seguir como isto é feito.

§ 4. Definição do vetor  $\delta x_{N,i}$

Teorema 3.2 - Seja  $x_0$  qualquer vetor inicial e defina  $x_n, n=1, \dots, N$ ; e  $\delta x_n, n=0, \dots, N-1$  através de (1.4.2). Seja  $x_{N+1} = x_N + \delta_{N+1}$  onde

$$\delta_{N+1} = \tau(A+B)^{-1}(q - Ax_N)$$

e seja

$$\delta x_N = (I - \tau(A+B)^{-1}A) \delta x_{N-1}$$

Se  $x_n, n \geq N+1$ , é gerado por

$$x_{n+1} = x_n + \delta x_n$$

com

$$\delta x_n = \frac{4T_{n-1}(y)}{(b-a)T_n(y)} (A+B)^{-1}(q - Ax_n) + \frac{T_{n-2}(y)}{T_n(y)} \delta x_{n-1}$$



onde

$$y = \frac{(a+b)}{(b-a)}$$

Então

$$e_n = (I - \tau(A+B)^{-1}A)P_{n-1}((A+B)^{-1}A)e_0,$$

onde

$$P_n(x) = \frac{T_n\left(\frac{a+b-2x}{b-a}\right)}{T_n(y)}; \text{ para } n \geq N+1.$$

Demonstração: Provaremos por indução em  $e_{N+j}$ ,  $j=1,2,\dots$

i) Por analogia com (3.3.2)

$$\begin{aligned} e_{N+1} &= (I - \tau(A+B)^{-1}A)P_N((A+B)^{-1}A)e_0 = \\ &= P_{N+1}^*((A+B)^{-1}A)e_0. \end{aligned}$$

A proposição é portanto válida para  $j=1$ .

ii) Suponhamos válida  $\forall m \leq j$ ,  $j > 1$ , ou seja

$$\begin{aligned} e_{N+m} &= P_{N+m}^*((A+B)^{-1}A)e_0 = \\ &= (I - \tau(A+B)^{-1}A)P_{N+m-1}((A+B)^{-1}A)e_0 \end{aligned}$$

Mostraremos que é válida  $\forall m \leq j+1$ .

De fato, como

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x); \forall n \quad [\text{V.A.2, teo 2}] \quad e$$

$$P_n(x) = \frac{T_n\left(\frac{a+b-2x}{b-a}\right)}{T_n\left(\frac{a+b}{b-a}\right)}, \quad \text{sendo } \frac{a+b}{b-a} = y, \text{ vem}$$

$$P_{N+j}(x) = - \frac{T_{N+j-1}(y) 4x}{T_{N+j}(y) (b-a)} P_{N+j-1}(x) +$$

$$+ P_{N+j-1}(x) + \frac{T_{N+j-2}(y)}{T_{N+j}(y)} (P_{N+j-1}(x) - P_{N+j-2}(x)).$$

[V.A.2, cor.2]

Daí,  $P_{N+j+1}^*((A+B)^{-1}A)$  satisfaz

$$P_{N+j+1}^*((A+B)^{-1}A) = - \frac{T_{N+j-1}(y) 4((A+B)^{-1}A)}{T_{N+j}(y) (b-a)} P_{N+j}^*((A+B)^{-1}A) +$$

$$+ P_{N+j}^*((A+B)^{-1}A) +$$

$$+ \frac{T_{N+j-2}(y)}{T_{N+j}(y)} (P_{N+j}^*((A+B)^{-1}A) - P_{N+j-1}^*((A+B)^{-1}A)). \quad (3.4.1)$$

$$\text{Agora, } e_{N+j+1} = x - x_{N+j+1} =$$

$$= x - x_{N+j} - \delta x_{N+j} =$$

$$e_{N+j} - \frac{4T_{N+j-1}(y)}{(b-a)T_{N+j}(y)} (A+B)^{-1} (q - Ax_{N+j}) - \frac{T_{N+j-2}(y)}{T_{N+j}(y)} \delta x_{N+j-1}$$

$$\text{mas, } q - Ax_{N+j} = A(x - x_{N+j}) = Ae_{N+j} \quad e$$

$$\delta x_{N+j-1} = x_{N+j} - x_{N+j-1} = e_{N+j-1} - e_{N+j} \quad (3.4.2)$$

Logo

$$e_{N+j+1} = e_{N+j} - \frac{4T_{N+j-1}(y)}{(b-a)T_{N+j}(y)} (A+B)^{-1} Ae_{N+j} -$$

$$- \frac{T_{N+j-2}(y)}{T_{N+j}(y)} (e_{N+j-1} - e_{N+j}).$$

Portanto, substituindo  $e_{N+j}$  e  $e_{N+j-1}$  pelos seus valores,

$P_{N+j}^* ((A+B)^{-1} A) e_0$  e  $P_{N+j-1}^* ((A+B)^{-1} A) e_0$  respectivamente obtemos ,  
de acordo com (3.4.1)  $e_{N+j+1} = P_{N+j+1}^* ((A+B)^{-1} A) e_0$  .

O caso  $j=1$ , isto é, a passagem de  $e_{N+1}$  para  $e_{N+2}$  que não pode ser mostrada utilizando (3.4.2) exige apenas que se observe que

$$\begin{aligned}
\delta x_{N+j-1} &= \delta x_N = (I - \tau(A+B)^{-1}A) \delta x_{N-1} = \\
&= (I - \tau(A+B)^{-1}A) (e_{N-1} - e_N) = \\
&= \left[ P_N^* ((A+B)^{-1}A) - P_{N+1}^* ((A+B)^{-1}A) \right] e_0.
\end{aligned}$$

Usando o fato de que o mesmo polinômio que aparece em  $e_n$  aparece em  $v_n$  (vide demonstração do teorema 2.2) concluímos:

$$v_{N+j} = P_{N+j}^* ((A+B)^{-1}A) v_0$$

A convergência de  $x_{N+2,i}$ ,  $x_{N+3,i}$ , ... a  $x$  é consequência de  $e_{N+j} = P_{N+j}^* ((A+B)^{-1}A) e_0$ .

De fato, pelo teorema 3.1, após certo número de iterações o caso  $P_N(\lambda_K) \cong |P_N(\lambda_1)|$  fatalmente ocorre e, como  $P_N(\lambda_K) < 1$  devemos ter  $|P_N(\lambda_1)| < 1$ . Ora, pelo lema 2.1, o  $\max_{i=1, \dots, K} |P_N(\lambda_i)|$  ocorre

em  $\lambda_1$  ou  $\lambda_K$ , portanto, após certo número de iterações o

$\max_{i=1, \dots, K} |P_N(\lambda_i)| < 1$  e o procedimento (que inicialmente pode divergir) torna-se convergente. Cumpre notar que, na pior das hipóteses, quando  $\lambda_K \cong 0$  acarretando  $P_N(\lambda_K) \cong 1$ , a multiplicação de

$P_N(x)$  por  $(1 - \frac{x}{b_i})$  força que  $|P_N^*(\lambda_1)| < 1$ .

Comentário: a suposição  $(1 - \frac{x}{b_i}) < 1$ ,  $\lambda_K \leq x \leq \lambda_1$  decorre

de estarmos considerando o caso  $|P_N(\lambda_1)| \cong P_N(\lambda_K)$  que implica em  $b_i > \frac{\lambda_K + \lambda_1}{2}$ .

Vamos estudar agora o comportamento da seqüência  $\mu_{N+2,i}, \mu_{N+3,i}, \dots$

Suponhamos então que  $N$  iterações do tipo (3.2.1) foram feitas.

Calculamos  $v_{N,i}$  e  $\mu_{N,i} = \frac{\langle Av_{N,i}, v_{N,i} \rangle}{\langle (A+B)v_{N,i}, v_{N,i} \rangle}$ .

Se  $\mu_{N,i} \in [a_i, b_i]$  fazemos uma iteração da forma (3.3.1) e a seguir mais  $N$  iterações do tipo (3.2.1) obtendo  $v_{2N+1,i}$  e

$$\mu_{2N+1,i} = \frac{\langle Av_{2N+1,i}, v_{2N+1,i} \rangle}{\langle (A+B)v_{2N+1,i}, v_{2N+1,i} \rangle}$$

Se  $\mu_{2N+1,i}$  ainda pertence a  $[a_i, b_i]$ , fazemos outra iteração da forma (3.3.1) e, novamente, mais  $N$  iterações do tipo (3.2.1). Calculamos  $\mu_{3N+2,i}$ . Se  $\mu_{3N+2,i} \notin [a_i, b_i]$  melhoramos o intervalo; caso contrário seguimos com o processo, obtendo

$$\mu_{4N+3,i}, \mu_{5N+4,i}, \dots$$

correspondentes a

$$v_{4N+3,i}, v_{5N+4,i}, \dots$$

Ora,  $v_{N+1,i} = P_{N+1, a_i, b_i}^* ((A+B)^{-1}A)v_{0,i}$  e, pela descrição do procedimento, é fácil ver que

$$v_{2N+1,i} = \left( I - \frac{(A+B)^{-1}A}{b_i} \right) P_{2N,a_i,b_i} \left( (A+B)^{-1}A \right) v_{0,i}$$

$$v_{2N+2,i} = \left( I - \frac{(A+B)^{-1}A}{b_i} \right)^2 P_{2N,a_i,b_i} \left( (A+B)^{-1}A \right) v_{0,i}$$

e, de modo geral

$$v_{jN+(j-1),i} = \left( I - \frac{(A+B)^{-1}A}{b_i} \right)^{(j-1)} P_{jN,a_i,b_i} \left( (A+B)^{-1}A \right) v_{0,i}$$

Comentário:  $\left( I - \frac{(A+B)^{-1}A}{b_i} \right)$  comuta com  $P_{jN,a_i,b_i} \left( (A+B)^{-1}A \right)$ .

Agora, por um desenvolvimento análogo ao feito para o teorema 2.1 (vide comentário feito em seguida a esse teorema), concluímos que o nosso problema reduz-se a estudar o comportamento do

$$\begin{aligned} & \lim_{j \rightarrow \infty} \left| \frac{P_{jN+(j-1),a_i,b_i}^*(\lambda_1)}{P_{jN+(j-1),a_i,b_i}^*(\lambda_K)} \right| = \\ & = \lim_{j \rightarrow \infty} \left| \frac{\left( 1 - \frac{\lambda_1}{b_i} \right)^{(j-1)} P_{jN,a_i,b_i}(\lambda_1)}{\left( 1 - \frac{\lambda_K}{b_i} \right)^{(j-1)} P_{jN,a_i,b_i}(\lambda_K)} \right| = \end{aligned}$$

$$= \lim_{j \rightarrow \infty} \left| \frac{\left( \frac{b_i - \lambda_1}{b_i - \lambda_K} \right)^{(j-1)} P_{jN, a_i, b_i}(\lambda_1)}{P_{jN, a_i, b_i}(\lambda_K)} \right|$$

Ora, como estamos no caso  $\left| P_{N, a_i, b_i}(\lambda_1) \right| \cong P_{N, a_i, b_i}(\lambda_K)$ , é razoável supormos  $b_i - \lambda_K > \lambda_1 - b_i$ . Portanto, quando  $j$  cresce,

$$\left| \frac{b_i - \lambda_1}{b_i - \lambda_K} \right|^{(j-1)}$$

vai a zero e daí, se o

$$\lim_{j \rightarrow \infty} \frac{P_{jN, a_i, b_i}(\lambda_1)}{P_{jN, a_i, b_i}(\lambda_K)}$$

foi finito (e para isso basta e é necessário que tenhamos

$$P_{N, a_i, b_i}(\lambda_1) \leq P_{N, a_i, b_i}(\lambda_K),$$

concluimos que  $\mu_{jN+(j-1), i}$  aproxima  $\lambda_K$ .

Caso  $\left| \frac{P_{jN, a_i, b_i}(\lambda_1)}{P_{jN, a_i, b_i}(\lambda_K)} \right| \rightarrow +\infty$  quando  $j \rightarrow +\infty$  (consequência de

$$P_{N, a_i, b_i}(\lambda_1) > P_{N, a_i, b_i}(\lambda_K)) \text{ temos:}$$

$$\lim_{j \rightarrow \infty} \left| \frac{b_i^{-\lambda_1}}{b_i^{-\lambda_K}} \right|^{(j-1)} \frac{P_{jN, a_i, b_i}(\lambda_1)}{P_{jN, a_i, b_i}(\lambda_K)} =$$

$$= \lim_{j \rightarrow \infty} \left| \frac{b_i^{-\lambda_1}}{b_i^{-\lambda_K}} \right|^{(j-1)} \frac{\cosh(jN \cosh^{-1} \left| \frac{a_i + b_i - 2\lambda_1}{b_i - a_i} \right|)}{\cosh(jN \cosh^{-1} \left| \frac{a_i + b_i - 2\lambda_K}{b_i - a_i} \right|)}$$

Seja  $r = \left| \frac{b_i^{-\lambda_1}}{b_i^{-\lambda_K}} \right|$ ,

$$s_1 = \cosh^{-1} \left| \frac{a_i + b_i - 2\lambda_1}{b_i - a_i} \right|$$

$$s_2 = \cosh^{-1} \left| \frac{a_i + b_i - 2\lambda_K}{b_i - a_i} \right|$$

Então,  $r < 1$ ;  $s_1 > s_2 > 0$ . Daí, vem

$$\lim_{j \rightarrow \infty} \left| \frac{b_i^{-\lambda_1}}{b_i^{-\lambda_K}} \right|^{(j-1)} \frac{\cosh(jN \cosh^{-1} \left| \frac{a_i + b_i - 2\lambda_1}{b_i - a_i} \right|)}{\cosh(jN \cosh^{-1} \left| \frac{a_i + b_i - 2\lambda_K}{b_i - a_i} \right|)} =$$



$$\begin{aligned}
&= \lim_{j \rightarrow \infty} r^{(j-1)} \frac{e^{jNs_1} + e^{-jNs_1}}{e^{jNs_2} + e^{-jNs_2}} \leq \\
&\leq \lim_{j \rightarrow \infty} r^{(j-1)} e^{jN(s_1-s_2)} = \frac{1}{r} \lim_{j \rightarrow \infty} (r e^{N(s_1-s_2)})^j =
\end{aligned}$$

$$= \begin{cases} 0 & \text{se } r e^{N(s_1-s_2)} < 1 \\ 1 & \text{se } r e^{N(s_1-s_2)} = 1/r \\ +\infty & \text{se } r e^{N(s_1-s_2)} > 1 \end{cases}$$

Ora, como  $s_1 - s_2 > 0$ , se quisermos que a seqüência dos  $\mu_{jN+(j-1),i}$  aproxime  $\lambda_K$ , temos que limitar o valor de  $N$ . Cumpre notar que, se

$$\lim_{j \rightarrow \infty} r^{(j-1)} \frac{e^{js_1} + e^{-js_1}}{e^{js_2} + e^{-js_2}} = +\infty$$

não há com que nos preocuparmos pois, simplesmente, ao invés de aproximarmos  $\lambda_K$ , aproximamos  $\lambda_1$  embora, é verdade, a convergência da seqüência dos  $\mu_{jN+(j-1),i}$  a  $\lambda_1$  seja infelizmente retardada pelas iterações do tipo (3.3.1). É conveniente portanto que as aproximações  $\mu_{jN+(j-1),i}$ , computadas após a ocorrência do caso

$\mu_{N,i} \in [a_i, b_i]$ , se o intervalo inicial  $[a_1, b_1]$  ainda não foi razoavelmente melhorado ou seja enquanto o valor de  $r$  não foi bastante reduzido, sejam calculadas com pequenos valores de  $N$ .

Comentário: Mesmo no caso altamente improvável, como a própria forma da expressão evidencia, do

$$\lim_{j \rightarrow \infty} r = \frac{(j-1) \frac{e^{js_1} + e^{-js_1}}{e^{js_2} + e^{-js_2}}}{e^{js_2} + e^{-js_2}}$$

ser finito e diferente de zero, embora não consigamos mais melhorar o intervalo, o algoritmo é convergente.

#### § 5. Apresentação de um algoritmo para aplicação do método.

1 - Defina um intervalo inicial  $[a_1, b_1]$  onde

$$a_1 = \min_k \frac{(A)_{kk}}{(A+B)_{kk}}$$

e

$$b_1 = \max_k \frac{(A)_{kk}}{(A+B)_{kk}}$$

Se  $a_1 = b_1$ , redefina o intervalo usando (3.2.2)

2 - Escolha um vetor inicial  $x_{0,1}$  e, um  $\epsilon > 0$  que será usado como "end test". A iteração para quando a norma do vetor residual,  $r_n$ , é menor que  $\epsilon$ .

3 - Escolha um inteiro  $N$  que será o número de iterações a serem executadas antes de tentar melhorar o intervalo.

Comentário: Quando  $N$  é aumentado, a convergência de  $a_i$  e  $b_i$  para  $\lambda_K$  e  $\lambda_1$  é melhorada, mas a convergência de  $x_n$  a  $x$  é retardada se os valores de  $a_i$  e  $b_i$  em uso são aproximações grosseiras de  $\lambda_K$  e  $\lambda_1$ .

Segundo Diamond [1] valores de  $N$  iguais a 5,6,7,8,9,10 são igualmente eficientes. Depois do intervalo ter sido melhorado algumas vezes, pode ser conveniente aumentar  $N$ .

4 - Execute  $N$  passos da iteração (3.2.1), gerando  $x_{N,i}$ ;  $v_{N-1,i}$  e  $\delta x_{N-1,i}$ .

5 - Calcule  $r_{N,i} = q - Ax_{N,i}$ . Se  $\|r_{N,i}\| < \epsilon$  pare; caso contrário calcule

$$v_{N,i} = (A+B)^{-1} r_{N,i} \quad e$$

$$\mu_{N,i} = \frac{\langle Av_{N,i}, v_{N,i} \rangle}{\langle (A+B)v_{N,i}, v_{N,i} \rangle} = \frac{\langle Av_{N,i}, v_{N,i} \rangle}{\langle r_{N,i}, v_{N,i} \rangle}.$$

6 - i) Se  $\mu_{N,i} < a_i$ , defina

$$a_{i+1} = \min(\min_{j \leq i} a_j, \mu_{N,i}) \quad e \quad b_{i+1} = b_i.$$

Some 1 a  $i$  e volte para 4. (Use  $v_{N,i}$  calculado em 5, na primeira iteração de 4).

ii) Se  $\mu_{N,i} > b_i$ , defina  $a_{i+1} = b_i$  e  $b_{i+1} = \mu_{N,i}$ .  
 Some 1 a  $i$  e volte para 4. (Use  $v_{N,i}$  calculado em 5 na primeira iteração de 4).

Comentário: A definição do intervalo seguinte quando  $\mu_{N,i} > b_i$  é diferente da definição quando  $\mu_{N,i} < a_i$  pela seguinte razão: componentes do vetor erro correspondentes a autovalores  $\lambda_k > b_i$  serão aumentadas se, por acaso,  $\left| P_{N,a_i,b_i}(\lambda_k) \right| > 1$ . Por esta razão, quando  $\mu_{N,i} > b_i$ , o intervalo seguinte é definido de modo a forçar uma redução nas componentes dominantes do vetor erro. As componentes dominantes quando  $\mu_{N,i} < a_i$ , por outro lado, nunca são aumentadas uma vez que  $\left| P_{N,a_i,b_i}(\lambda_k) \right| < 1$  para todo  $\lambda_k \in (0, a_i)$ . Portanto, quando  $\mu_{N,i} < a_i$ , não é necessário preocuparmo-nos com a redução das componentes dominantes.

iii) Se  $\mu_{N,i} \in [a_i, b_i]$ , faça uma iteração da forma (3.3.1) e  $N$  iterações do tipo (3.2.1). Volte para 5.

Comentários: i) O objetivo desse passo é contornar o problema  $P_{N,a_i,b_i}(\lambda_k) \cong \left| P_{N,a_i,b_i}(\lambda_1) \right|$  como foi discutido nos parágrafos 3 e 4.

ii) O algoritmo é sempre convergente uma vez que obrigamos

$$\max_k \left| P_{N,a_i,b_i}(\lambda_k) \right| < 1$$

iii) Resultados experimentados e uma estimativa do número de operações aritméticas requeridas pelo método podem ser encontrados em [1] e [4].

APÊNDICE 1Método de Stiefel para Implementação da Sequência  $\{\tau_n\}$ 

Todos os trabalhos por nós consultados sobre o método de Stiefel omitem a maneira como ele foi deduzido. Dão, para isso, como referência, um trabalho do próprio Stiefel, datado de 1958, intitulado "Kernel polynomials in linear algebra and their numerical applications", publicado no Nat. Bur. Standards Appl. Math. Ser. , vol. 49, que não nos foi possível obter. Em vista disso, damos a seguir o que julgamos ser uma maneira de deduzi-lo.

Em primeiro lugar, o erro  $e_n = x - x_n$  tem que satisfazer  $e_n = P_n((A+B)^{-1}A)e_0$  onde

$$P_n(x) = \frac{T_n\left(\frac{a+b-2x}{b-a}\right)}{T_n\left(\frac{a+b}{b-a}\right)}. \text{ Daí, temos}$$

$$\begin{aligned} x - x_{n+1} &= e_{n+1} = P_{n+1}((A+B)^{-1}A)e_0 = \\ &= \left[ -\frac{4T_n(y)(A+B)^{-1}A}{T_{n+1}(y)(b-a)} P_n((A+B)^{-1}A) + P_n((A+B)^{-1}A) + \right. \\ &\left. + \frac{T_{n-1}(y)}{T_{n+1}(y)} (P_n((A+B)^{-1}A) - P_{n-1}((A+B)^{-1}A)) \right] e_0 \text{ onde } y = \frac{a+b}{b-a}. \end{aligned}$$

De fato, como  $T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x)$  [V.A.2, teo 2] vem:

$$\begin{aligned}
 P_{n+1}(x) &= \frac{T_{n+1}\left(\frac{a+b-2x}{b-a}\right)}{T_{n+1}(y)} = 2\left(\frac{a+b-2x}{b-a}\right) \frac{T_n\left(\frac{a+b-2x}{b-a}\right)}{T_{n+1}(y)} - \\
 & - \frac{T_{n-1}\left(\frac{a+b-2x}{b-a}\right)}{T_{n+1}(y)} = 2\left(\frac{a+b-2x}{b-a}\right) \frac{T_n(y)}{T_{n+1}(y)} P_n(x) - \frac{T_{n-1}(y)}{T_{n+1}(y)} P_{n-1}(x) = \\
 & = \left[ 2\left(\frac{a+b}{b-a}\right) \frac{T_n(y)}{T_{n+1}(y)} - \frac{4x}{b-a} \frac{T_n(y)}{T_{n+1}(y)} \right] P_n(x) - \\
 & - \frac{T_{n-1}(y)}{T_{n+1}(y)} P_{n-1}(x) = \left[ 1 + \frac{T_{n-1}(y)}{T_{n+1}(y)} - \frac{4x}{b-a} \frac{T_n(y)}{T_{n+1}(y)} \right] P_n(x) - \\
 & - \frac{T_{n-1}(y)}{T_{n+1}(y)} P_{n-1}(x) = - \frac{4T_n(y)x}{T_{n+1}(y)(b-a)} P_n(x) + P_n(x) + \\
 & + \frac{T_{n-1}(y)}{T_{n+1}(y)} (P_n(x) - P_{n-1}(x)).
 \end{aligned}$$

Chamando  $x_{n+1} - x_n = \delta x_n$  vem:

$$e_{n+1} - e_n = x - x_{n+1} - x + x_n = -\delta x_n.$$

Dai,  $e_{n+1} - e_n =$

$$= - \frac{4T_n(y)}{T_{n+1}(y)(b-a)} (A+B)^{-1}_A e_n + \frac{T_{n-1}(y)}{T_{n+1}(y)} (e_n - e_{n-1})$$

implica em

$$\delta x_n = \frac{4T_n(y)}{T_{n+1}(y)(b-a)} (A+B)^{-1}_A e_n + \frac{T_{n-1}(y)}{T_{n+1}(y)} \delta x_{n-1}.$$

Agora, sendo  $r_n = q - Ax_n$  e  $v_n = (A+B)^{-1}_A r_n$  obtemos:

$$\begin{aligned} (A+B)^{-1}_A e_n &= (A+B)^{-1}_A (x - x_n) = \\ &= (A+B)^{-1}_A (q - Ax_n) = (A+B)^{-1}_A r_n = v_n \end{aligned}$$

Resumindo, obtivemos:

$$x_{n+1} = x_n + \delta x_n$$



$$\delta x_n = \frac{4T_n(y)}{T_{n+1}(y)(b-a)} v_n + \frac{T_{n-1}(y)}{T_{n+1}(y)} \delta x_{n-1} \quad \text{e} \quad \delta x_0 = \frac{2}{a+b},$$

pois  $\delta x_0 = e_0 - e_1 = e_0 - P_1((A+B)^{-1}A)e_0 =$

$$= e_0 - \frac{\frac{(a+b-2(A+B)^{-1}A)}{b-a}}{\left(\frac{-a+b}{b-a}\right)} e_0 = \frac{2}{a+b} (A+B)^{-1}A e_0 = \frac{2}{a+b} v_0.$$

APÊNDICE 2Definição e algumas propriedades dos polinômios de Tchebchev

Definição 1 - O polinômio de Tchebychev, de grau  $N$ , representado por  $T_N$ , é definido:

$$T_N(x) = \begin{cases} \cos(n \cos^{-1} x); & |x| \leq 1 \\ \cosh(n \cosh^{-1} x); & x \geq 1 \\ (-1)^n \cosh(n \cosh^{-1} |x|); & x \leq -1 \end{cases}$$

Teorema 1 - O polinômio de Tchebychev de grau  $N$ ,  $T_N$ , tem a propriedade

$$\frac{\max_{x \in [-1, 1]} |T_N(x)|}{|T_N(y)|} \leq \frac{\max_{x \in [-1, 1]} |q(x)|}{|q(y)|},$$

qualquer que seja o polinômio  $q$ , não identicamente nulo, de grau menor ou igual a  $N$ , e qualquer que seja  $y \notin [-1, 1]$ .

Demonstração: Seja  $Q(x) = \frac{q(x)}{\max_{x \in [-1, 1]} |q(x)|}$ . Como

$$\frac{1}{|T_N(y)|} \leq \frac{1}{|Q(y)|}, \forall y \notin [-1, 1], \text{ pois, se } Q \text{ é qualquer polinômio}$$

de grau menor ou igual a  $N$ , tal que  $|Q(x)| \leq 1$  quando  $x \in [-1, 1]$ , então  $\forall y \notin [-1, 1], |Q(y)| \leq |T_N(y)|$  (caso contrário, seja

$\lambda = \frac{Q(y)}{T_N(y)}$  e considere o polinômio  $P = \lambda T_N - Q$ . Ele muda de si-

nal  $N$  vezes e se anula em  $y$ ), temos que

$$\frac{\max_{x \in [-1, 1]} |T_N(x)|}{|T_N(y)|} \leq \frac{\max_{x \in [-1, 1]} |q(x)|}{|q(y)|}$$

$\forall y \notin [-1, 1]$ .

Corolário 1 - O polinômio  $P_{N,a,b}$  satisfaz

$$\frac{\max_{x \in [a, b]} |P_{N,a,b}(x)|}{|P_{N,a,b}(y)|} \leq \frac{\max_{a \leq x \leq b} |q(x)|}{|q(y)|}$$

qualquer que seja o polinômio  $q$ , não identicamente nulo, de grau menor ou igual a  $N$  e qualquer que seja  $y \in [a, b]$ .

Teorema 2 - Os polinômios de Tchebychev,  $T_N$ , satisfazem a relação de recorrência:

$$T_{N+1}(x) = 2xT_N(x) - T_{N-1}(x)$$

Demonstração - Basta observar a identidade trigonométrica:

$$\cos(N+1)\theta + \cos(N-1)\theta = 2\cos\theta \cos N\theta$$

ou sua similar hiperbólica

$$\cosh(N+1)\theta + \cosh(N-1)\theta = 2\cosh\theta \cosh N\theta$$

Corolário 2 - O polinômio

$$P_N(x) = \frac{T_N\left(\frac{a+b-2x}{b-a}\right)}{T_N\left(\frac{a+b}{b-a}\right)} \quad \text{satisfaz}$$

$$P_{N+1}(x) = -\frac{4xT_N(y)}{(b-a)T_{N+1}(y)} P_N(x) + P_N(x) + \frac{T_{N-1}(y)}{T_{N+1}(y)} (P_N(x) - P_{N-1}(x)) \quad \text{onde}$$

$$y = \frac{a+b}{b-a}$$

Demonstração: vide apêndice 1.

Teorema 3 - O polinômio  $P_n$ , de grau  $n$ , tal que  $\max_{a \leq x \leq b} |P_n(x)|$  é mínimo e que satisfaz  $P_n(0)=1$  é dado por

$$P_n(x) = \frac{T_n\left(\frac{a+b-2x}{b-a}\right)}{T_n\left(\frac{a+b}{b-a}\right)}$$

Demonstração - Vide Forsythe e Wasow [5] pag. 227, § 21.5  
(na realidade esse resultado é consequência imediata do teorema 1)

Teorema 4 - O polinômio de Tchebychev de grau  $N, T_N$ , satisfaz

$$|T_N(y)| > \max_{x \in [-1,1]} |T_N(x)|, \quad \forall y \notin [-1,1]$$

e, além disso, se  $|y_1| > |y_2| \geq 1$  então  $|T_N(y_1)| > |T_N(y_2)|$ .

Demonstração - Segue-se trivialmente da definição.

APÊNDICE 3Definições e propriedades gerais

Definição 1. Os autovalores extremos de uma matriz são os autovalores de maior valor algébrico de menor valor algébrico.

Definição 2. Uma matriz  $A$ ,  $n \times n$ , real e simétrica é dita positiva definida se

$$x^T A x > 0, \quad \forall x \in \mathbb{R}^n, \quad x \neq 0,$$

i.é, se  $A$  é positiva definida vista como uma forma bilinear.

Teorema 1 - Os autovalores de uma matriz  $A$ , positiva definida, são positivos.

Demonstração: Conseqüência imediata da definição

Teorema 2 - Se  $A$  e  $B$  são positivas definidas, então os autovetores  $x_k$ , e os autovalores correspondentes,  $\lambda_k$ , do problema do autovalor generalizado,  $Ax_k = \lambda_k Bx_k$ , satisfazem

$$i) \quad \lambda_k > 0$$

$$ii) \quad x_k^T Bx_j = 0, \quad j \neq k$$

Demonstração:

$$i) \quad Ax_k = \lambda_k B x_k$$

$$x_k^T A x_k = \lambda_k x_k^T B x_k \longrightarrow \lambda_k > 0$$

ii) sejam  $\lambda_k$  e  $\lambda_j$  distintos e tais que

$$Ax_k = \lambda_k B x_k \quad e \quad Ax_j = \lambda_j B x_j$$

Temos  $(A x_k)^T x_j = \lambda_k (B x_k)^T x_j$ ,  $x_k^T A x_j = \lambda_j x_k^T B x_j$  ou, usando a simetria de A e B,

$$x_k^T A x_j = \lambda_k x_k^T B x_j,$$

$$x_k^T A x_j = \lambda_j x_k^T B x_j. \quad \text{Subtraindo as duas últimas duas expressões, obtemos}$$

mas duas expressões, obtemos

$$(\lambda_j - \lambda_k) x_k^T B x_j = 0 \quad e \quad \text{como, por hipótese}$$

$$\lambda_j \neq \lambda_k, \quad \text{segue-se} \quad x_k^T B x_j = 0.$$

Observação: Para um estudo mais detalhado do problema do autovalor generalizado, naquilo que interessa a este trabalho, a referência é Hildebrand [6], parágrafo 1.25.

Definição 3. O raio espectral de uma matriz  $A$  é o módulo do autovalor, da matriz  $A$ , de maior valor absoluto.



REFERÊNCIAS BIBLIOGRÁFICAS

1. M.A. DIAMOND, "An economical algorithm for the solution of finite difference equations", Department of Computer Science, University of Illinois, Urbana, Illinois, Report Nº R-71-492, December 1971.
2. AMNON BRACHA - BARAK e PAUL E. SAYLOR, "A symmetric factorization procedure for the solution of elliptic boundary value problems", SIAM Journal on Numerical Analysis, March 1973.
3. T. DUPONT, "A factorization procedure for the solution of elliptic difference equation", SIAM Journal on Numerical Analysis, December, 1968.
4. M.A. DIAMOND, "An adaptive algorithm for the application of factorization techniques in the solution of elliptic difference equations", a ser publicado.
5. FORSYTHE, G.E. e WASOW, W.R., "Finite - difference methods for partial differential equations", John Wiley e Sons, Inc. ; New York 1960.
6. HILDEBRAND, F.B., "Methods of applied mathematics", Prentice-Hall of India Private Limited, New Delhi 1968.