

PROJETO SUPERVISIONADO

MÉTODOS ITERATIVOS PARA SISTEMAS LINEARES

Nelson Gomes Brasil Junior
Aluno

Profa. Dra. Márcia A. Gomes-Ruggiero
Orientadora

Departamento de Matemática Aplicada
IMECC – UNICAMP

Dezembro, 2011

1 Introdução

Os métodos iterativos para a resolução de um sistema linear da forma $A\mathbf{x} = \mathbf{b}$, sendo A uma matriz $N \times N$, e $\mathbf{x}, \mathbf{b} \in \mathbb{R}^N$ foram inicialmente propostos na década de 50, como uma alternativa aos denominados métodos diretos (Eliminação de Gauss, Cholesky). Estes métodos realizam um trabalho da ordem de N^3 operações para resolver um sistema, e quando N é grande resulta em um esforço computacional com custos proibitivos. Outro aspecto negativo dos métodos diretos é a alteração da estrutura original da matriz. Se o sistema linear a ser resolvido vier da discretização de um problema de valor de contorno, a matriz de coeficientes será uma matriz esparsa. Ao aplicar um método direto (fatoração LU, por exemplo) ocorrem preenchimentos na matriz, introduzindo valores onde originalmente eram nulos.

Os métodos iterativos que estudamos neste projeto são os chamados métodos de Krylov, que consistem em obter aproximações para a solução do sistema nos chamados subespaços de Krylov. Para este projeto, estudamos os métodos dos gradientes conjugados e GMRES.

O trabalho está organizado da seguinte forma: Na Seção [2](#) estudamos o método dos gradientes conjugados. Apresentamos uma introdução teórica ao métodos, assim como algumas propriedades dos subespaços de Krylov. Descrevemos um algoritmo para o método dos gradientes conjugados, alguns resultados de convergência, assim como testes computacionais para exemplificar a teoria. Na Seção [3](#) descrevemos o método GMRES, iniciando com as ideias básicas do método e uma breve explicação da iteração de Arnoldi. Tal como fizemos com gradientes conjugados, apresentamos um algoritmo e resultados de convergência. Por fim, realizamos testes computacionais, usando o MatLab 7.9.

2 Gradientes conjugados

Nesta seção estudaremos alguns aspectos do método dos gradientes conjugados. Tomamos como base as referências [3](#), [4](#).

O método dos gradientes conjugados começou a ser estudado na década de 50, sendo que foi proposto como um método direto, não como um método iterativo. Muitos autores citam o artigo de Herstenes e Stifel [2](#) como referência base ao que foi desenvolvido posteriormente.

2.1 Aspectos teóricos dos gradientes conjugados

O método dos gradientes conjugados é aplicado para a resolução de sistemas lineares da forma $A\mathbf{x} = \mathbf{b}$, onde a matriz A é *simétrica definida positiva*, isto é, matrizes $N \times N$, tais

que, para todo $\mathbf{x} \in \mathbb{R}^N$, $\mathbf{x} \neq \mathbf{0}$, temos

$$\mathbf{x}^T A \mathbf{x} > 0 \quad (1)$$

Note que a relação expressa em (1) induz uma norma vetorial, que chamaremos de *norma-A*. Definiremos tal norma como $\|\mathbf{x}\|_A = \sqrt{\mathbf{x}^T A \mathbf{x}}$. Temos ainda que, se a matriz A é definida positiva, sua inversa A^{-1} também é definida positiva, e portanto também induz uma norma vetorial.

Observe que um problema de minimização da forma

$$\begin{aligned} \min \quad & \phi(\mathbf{x}) = \mathbf{x}^T A \mathbf{x} - \mathbf{x}^T \mathbf{b} \\ \text{s.a} \quad & \mathbf{x} \in \mathbb{R}^n \end{aligned} \quad (2)$$

é um problema convexo se assumirmos que A é simétrica definida positiva. Desta forma (2) tem como ponto ótimo, \mathbf{x}_* tal que $\nabla\phi(\mathbf{x}_*) = 0$, ou seja, o ponto satisfaz

$$\nabla\phi(\mathbf{x}_*) = 0 \iff A\mathbf{x}_* - \mathbf{b} = 0 \iff A\mathbf{x}_* = \mathbf{b},$$

em que \mathbf{x}_* seja solução do sistema linear que queremos resolver.

Se a dimensão do problema é grande, (2) pode ser resolvido iterativamente impondo que a cada iteração k é obter o mínimo da função $\phi(\mathbf{x})$ com \mathbf{x} restrito a um subespaço $\mathcal{S} \in \mathbb{R}^k$. Desta forma, na iteração N a expectativa obter o mnimo de $\phi(\mathbf{x})$ em \mathbb{R}^n .

O resíduo da solução a cada iteração k é definido como $\mathbf{r}_k = \mathbf{b} - A\mathbf{x}_k$. Observando o problema de otimização, vemos que o resíduo é também $\mathbf{r}_k = -\nabla\phi(\mathbf{x}_k)$, o negativo do gradiente no ponto considerado.

O método dos gradientes conjugados trabalha com a resolução do problema (2) substituindo a restrição que o vetor procurado pertence a \mathbb{R}^N pela restrição de \mathbf{x} no subespaço de Krylov de dimensão k , na k -ésima iteração.

Os subespaços de Krylov são definidos como o espaço gerado por potências da matriz A multiplicadas pelo vetor \mathbf{b} , a saber

$$\mathcal{K}_n = \text{span}\{\mathbf{b}, A\mathbf{b}, \dots, A^{n-1}\mathbf{b}\}.$$

Os subespaços de Krylov satisfazem as seguintes propriedades.

- $\mathcal{K}_0 = \{\mathbf{0}\}$ e $\mathcal{K}_0 \subseteq \mathcal{K}_1 \subseteq \mathcal{K}_2 \subseteq \dots \subseteq \mathcal{K}_n$;
- Se $\mathcal{K}_k = \mathcal{K}_{k+1}$, então $\mathcal{K}_i = \mathcal{K}_k$ para todo $i \geq k$;
- $A^{-1}\mathbf{b} \in \mathcal{K}_n$.

Dado uma aproximação inicial \mathbf{x}_0 , vamos procurar, na iteração j um vetor $\mathbf{x}_j \in \mathbf{x}_0 + \mathcal{K}_j$ tal que \mathbf{x}_j seja o mínimo da função ϕ restrito a este subespaço, ou ainda, que o erro em \mathbf{x} seja menor possível, na norma- A . Podemos enunciar o seguinte lema:

Lema 1 *Seja $\mathcal{S} \subset \mathbb{R}^n$. Se \mathbf{x}_k minimiza $\phi(\mathbf{x})$ sobre \mathcal{S} , então \mathbf{x}_k também minimiza $\|\mathbf{x} - \mathbf{x}_*\|_A = \|\mathbf{r}\|_{A^{-1}}$ sobre \mathcal{S} .*

Demonstração: Note que

$$\|\mathbf{x} - \mathbf{x}_*\|_A^2 = (\mathbf{x} - \mathbf{x}_*)^T A (\mathbf{x} - \mathbf{x}_*) = \mathbf{x}^T A \mathbf{x} - \mathbf{x}^T A \mathbf{x}_* - (\mathbf{x}_*)^T A \mathbf{x} + (\mathbf{x}_*)^T A \mathbf{x}_*.$$

Como A é simétrica, e $A \mathbf{x}_* = \mathbf{b}$, temos

$$-\mathbf{x}^T A \mathbf{x}_* - (\mathbf{x}_*)^T A \mathbf{x} = -2\mathbf{x}^T A \mathbf{x}_* = -2\mathbf{x}^T \mathbf{b}.$$

Daí

$$\|\mathbf{x} - \mathbf{x}_*\|_A^2 = 2\phi(\mathbf{x}) + (\mathbf{x}_*)^T A \mathbf{x}_*.$$

Como $(\mathbf{x}_*)^T A \mathbf{x}_*$ é independente de \mathbf{x} , temos que minimizar $\phi(\mathbf{x})$ é equivalente a minimizar $\|\mathbf{x} - \mathbf{x}_*\|_A$.

Agora, fazendo $\mathbf{e} = \mathbf{x} - \mathbf{x}_*$, então

$$\|\mathbf{e}\|_A^2 = \mathbf{e}^T A \mathbf{e} = (A(\mathbf{x} - \mathbf{x}_*))^T A^{-1} (A(\mathbf{x} - \mathbf{x}_*)) = \|\mathbf{b} - A\mathbf{x}\|_{A^{-1}}^2.$$

□

2.2 Resultados teóricos e convergência

Nesta seção iremos desenvolver o algoritmo utilizado para gradientes conjugados, assim como as condições para a convergência do método.

Lema 2 *Seja A uma matriz simétrica definida positiva, e $\{\mathbf{x}_k\}$ a sequência gerada pelo método dos gradientes conjugados. Dados dois resíduos \mathbf{r}_k e \mathbf{r}_l , então*

$$\mathbf{r}_k^T \mathbf{r}_l = 0, \quad 0 \leq l < k.$$

Ou seja, \mathbf{r}_k e \mathbf{r}_l são ortogonais.

Demonstração: Como \mathbf{x}_k minimiza $\phi(\mathbf{x})$ sobre $\mathbf{x}_0 + \mathcal{K}_k$, nós temos que, para todo $\mathbf{v} \in \mathcal{K}_k$

$$\frac{d\phi(\mathbf{x}_k + t\mathbf{v})}{dt} = \nabla\phi(\mathbf{x}_k + t\mathbf{v})^T \mathbf{v} = 0,$$

para $t = 0$. Assim, teremos que

$$\nabla\phi(\mathbf{x}_k)^T \mathbf{v} = 0 \iff \mathbf{r}_k^T \mathbf{v} = 0, \quad \forall \mathbf{v} \in \mathcal{K}_k.$$

O resultado segue deste fato e do fato de que $\mathbf{r}_l \in \mathcal{K}_{l+1}$ e, pela propriedade dos subespaços de Krylov, $\mathcal{K}_l \subseteq \mathcal{K}_k$ para todo $l < k$. \square

O Lema 2 nos dá uma condição sobre as iterações. Se $\mathbf{x}_k = \mathbf{x}_{k+1}$, teremos que $\mathbf{r}_k = \mathbf{r}_{k+1}$, e então $\|\mathbf{r}_k\|_2^2 = \mathbf{r}_k^T \mathbf{r}_{k+1} = 0$. Ou seja, se dois iterados consecutivos forem iguais, teremos a solução do sistema.

Note ainda que $\mathbf{r}_k = \mathbf{b} - A\mathbf{x}_k \in \mathcal{K}_{k+1}$. A ortogonalidade dos resíduos nos dá que $\mathcal{K} = \text{span}\{\mathbf{r}_0, \mathbf{r}_1, \dots, \mathbf{r}_{k-1}\}$.

O próximo lema define a direção de busca do método dos gradientes conjugados, e relaciona tal direção com vetores do subespaço de Krylov.

Lema 3 *Seja A simétrica, definida positiva e seja $\{\mathbf{x}_k\}$ iterações do método de gradientes conjugados. Se $\mathbf{x}_{k+1} \neq \mathbf{x}_*$, então*

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_{k+1} \mathbf{p}_{k+1}, \quad (3)$$

e

$$\mathbf{p}_{k+1} \in \mathcal{K}_{k+1}, \quad \mathbf{p}_{k+1}^T A \mathbf{v} = 0, \quad \forall \mathbf{v} \in \mathcal{K}_k.$$

Demonstração: Ver [3, p. 20] \square

No Lema 3, $\mathbf{p} \neq \mathbf{0}$, representa a direção de busca e α é um escalar não negativo. Pré-multiplicando a expressão em (3) pela matriz A , teremos

$$A\mathbf{x}_{k+1} = A\mathbf{x}_k + \alpha_{k+1} A\mathbf{p}_{k+1} \iff \mathbf{r}_{k+1} = \mathbf{r}_k - \alpha_{k+1} \mathbf{p}_{k+1}.$$

Podemos encontrar α de modo a manter a ortogonalidade entre tais resíduos. Obtemos assim:

$$\langle \mathbf{r}_{k+1}, \mathbf{r}_k \rangle = 0 \iff \langle \mathbf{r}_k, \mathbf{r}_k \rangle - \alpha_{k+1} \langle \mathbf{r}_k, A\mathbf{p}_{k+1} \rangle = 0 \iff \alpha_{k+1} = \frac{\langle \mathbf{r}_k, \mathbf{r}_k \rangle}{\langle \mathbf{r}_k, A\mathbf{p}_{k+1} \rangle}.$$

Para obter o vetor $\mathbf{p}_{k+1} \in \mathcal{K}_{k+1}$, usamos os vetores $\mathbf{p}_k \in \mathcal{K}_k$ e \mathbf{r}_k , que já estão calculados, e escrevemos:

$$\mathbf{p}_{k+1} = \mathbf{r}_k + \beta_{k+1} \mathbf{p}_k. \quad (4)$$

Agora, como $\langle \mathbf{p}_{k+1}, A\mathbf{r}_{k-1} \rangle = \mathbf{p}_{k+1}^T A\mathbf{r}_{k-1} = 0$, podemos obter β_{k+1} :

$$\begin{aligned} \mathbf{p}_{k+1}^T A\mathbf{r}_{k-1} &= 0 \\ (\mathbf{r}_k + \beta_{k+1} \mathbf{p}_k)^T A\mathbf{r}_{k-1} &= 0 \iff \beta_{k+1} = -\frac{\mathbf{r}_k^T A\mathbf{r}_{k-1}}{\mathbf{p}_k^T A\mathbf{p}_{k-1}}. \end{aligned}$$

Podemos representar α e β de outras formas equivalentes, como mostra o próximo lema.

Lema 4 *Seja A semidefinida positiva. Se $\mathbf{r}_k \neq \mathbf{0}$, então*

$$\alpha_{k+1} = \frac{\|\mathbf{r}_k\|_2^2}{\mathbf{p}_{k+1}^T A \mathbf{p}_{k+1}}$$

e

$$\beta_{k+1} = \frac{\|\mathbf{r}_k\|_2^2}{\|\mathbf{r}_{k-1}\|_2^2}.$$

Demonstração: Ver [3, p. 21]. □

Podemos então, esquematizar o seguinte algoritmo para o método dos gradientes conjugados:

Algoritmo 1: Gradientes conjugados

Entrada $A, \mathbf{b}, \mathbf{x}_0, \epsilon, kmax$
 $\mathbf{r} \leftarrow \mathbf{b} - A\mathbf{x}_0$, $\mu_0 \leftarrow \|\mathbf{r}\|_2^2$, $k = 1$
while $\sqrt{\mu_{k-1}} \geq \epsilon$ e $k < kmax$ **do**
 if $k = 1$ **then**
 $\mathbf{p} \leftarrow \mathbf{r}$
 else
 $\beta = \mu_{k-1}/\mu_{k-2}$ e $\mathbf{p} \leftarrow \mathbf{r} + \beta\mathbf{p}$
 $\mathbf{w} \leftarrow A\mathbf{p}$
 $\alpha \leftarrow \mu_{k-1}/(\mathbf{p}^T \mathbf{w})$
 $\mathbf{r} \leftarrow \mathbf{r} - \alpha\mathbf{w}$
 $\mu_k = \|\mathbf{r}\|_2^2$
 $k \leftarrow k + 1$
 end if
end while

Vamos agora tratar da convergência do método dos gradientes conjugados. Primeiramente, note que o Lema 1 implica que, se o vetor \mathbf{x}_k minimiza $\phi(\mathbf{x})$ sobre \mathcal{S}_k , então

$$\|\mathbf{x}_* - \mathbf{x}_k\|_A \leq \|\mathbf{w} - \mathbf{x}_*\|_A, \quad \forall \mathbf{w} \in \mathcal{S}_k.$$

Agora, se tomarmos $\mathcal{S}_k = \mathbf{x}_0 + \mathcal{K}_k$, teremos o resultado

$$\mathbf{w} = \mathbf{x}_0 + \sum_{j=0}^{k-1} \gamma_j A^j \mathbf{r}_0. \tag{5}$$

Sendo $\mathbf{r}_0 = \mathbf{b} - A\mathbf{x}_0 = A\mathbf{x}_* - A\mathbf{x}_0 = A(\mathbf{x}_* - \mathbf{x}_0)$. Subtraindo (5) da solução exata do sistema linear, temos:

$$\mathbf{x}_* - \mathbf{w} = \mathbf{x}_* - \mathbf{x}_0 - \sum_{j=0}^{k-1} \gamma_j A^{j+1} (\mathbf{x}_* - \mathbf{x}_0) = \left(I + \sum_{j=0}^{k-1} \gamma_j A^{j+1} \right) (\mathbf{x}_* - \mathbf{x}_0),$$

que podemos reescrever como

$$\mathbf{x}_* - \mathbf{w} = p(A)(\mathbf{x}_* - \mathbf{x}_0),$$

onde

$$p(z) = 1 + \sum_{j=0}^{k-1} \gamma_j z^{j+1}.$$

Transportamos então, nosso problema de minimização do subespaços de Krylov, para o espaços dos polinômios de grau k . Ou seja

$$\|\mathbf{x}_* - \mathbf{x}_k\|_A = \min_{p \in \mathcal{P}_k, p(0)=1} \|p(A)(\mathbf{x}_* - \mathbf{x}_0)\|_A. \quad (6)$$

Usaremos agora a decomposição espectral da matriz A (tal decomposição existe, dado que a matriz A é definida positiva). Então, existem matrizes U e Λ , tal que

$$A = U\Lambda U^T,$$

com U sendo uma matriz ortogonal contendo os autovetores de A nas colunas, e Λ uma matriz diagonal, com os autovalores de A na diagonal. Como a matriz U é ortogonal, temos que as potências de A são escritas como

$$A^m = U\Lambda^m U^T.$$

Disso temos que o polinômio $p(A)$ assume a seguinte forma:

$$p(A) = Up(\Lambda)U^T.$$

Note que podemos relacionar a norma- A com a norma-2 de um vetor, pois

$$\|\mathbf{x}\|_A^2 = \mathbf{x}^T A \mathbf{x} = \mathbf{x}^T A^{1/2} A^{1/2} \mathbf{x} = \|A^{1/2} \mathbf{x}\|_2^2.$$

Assim, temos que

$$\|p(A)\mathbf{x}\|_A = \|A^{1/2}p(A)\mathbf{x}\|_2 \leq \|A^{1/2}\mathbf{x}\|_2 \|p(A)\|_2 = \|\mathbf{x}\|_A \|p(A)\|_2.$$

Com estes resultados, reescrevemos o problema de otimização dado em (6) como

$$\|\mathbf{x}_* - \mathbf{x}_k\|_A = \min_{p \in \mathcal{P}_k, p(0)=1} \|p(A)(\mathbf{x}_* - \mathbf{x}_0)\|_A = \|\mathbf{x}_* - \mathbf{x}_0\|_A \min_{p \in \mathcal{P}_k, p(0)=1} \|p(A)\|_2.$$

Agora, observe que a norma-2 de $p(A)$ é equivalente a

$$\|p(A)\|_2 = \|UP(\Lambda)U^T\|_2 = \|p(\Lambda)\|_2 = \max_{z \in \sigma(A)} |p(z)|,$$

onde $\sigma(A)$ é o conjunto dos autovalores da matriz A . Assim, reescrevemos o problema de minimização como

$$\|\mathbf{x}_* - \mathbf{x}_k\|_A = \|\mathbf{x}_* - \mathbf{x}_0\|_A \min_{p \in \mathcal{P}_k, p(0)=1} \max_{z \in \sigma(A)} |p(z)|.$$

Corolário 5 *Seja A semidefinida positiva e $\{\mathbf{x}_k\}$ uma sequência gerada por gradientes conjugados. Dado k , seja $\{\bar{p}_k\}$ qualquer polinômio de grau k tal que $\bar{p}_k(0) = 1$. Então*

$$\frac{\|\mathbf{x}_* - \mathbf{x}_k\|_A}{\|\mathbf{x}_* - \mathbf{x}_0\|_A} \leq \max_{z \in \sigma(A)} |\bar{p}_k(z)| \quad (7)$$

O polinômio p_k é conhecido como polinômio residual. Usamos tais polinômios para encontrar um limite superior o número de iterações necessárias para que método dos gradientes conjugados encontre a solução do sistema.

Podemos usar a expressão (7) para mostrar como o método dos gradientes conjugados pode ser encarado como um método direto.

Teorema 6 *Seja A uma matriz simétrica definida positiva. Então o método dos gradientes conjugados encontra a solução exata do sistema linear $A\mathbf{x} = b$ em no máximo N iterações.*

Demonstração: Seja $\{\lambda_i\}_{i=1}^N$ os autovalores de A . Construímos o seguinte polinômio teste

$$\bar{p}(z) = \prod_{i=1}^N \frac{\lambda_i - z}{\lambda_i}.$$

$\bar{p} \in \mathcal{P}_N$, pois tem grau N e $\bar{p}(0) = 1$. Então, usando (7) e o fato de que \bar{p} se anula em $\sigma(A)$, temos

$$\|\mathbf{x}_* - \mathbf{x}_k\|_A \leq \|\mathbf{x}_* - \mathbf{x}_0\|_A \max_{z \in \sigma(A)} |\bar{p}_k(z)| = 0.$$

□

Observe que o Teorema 6 nos dá a garantia de convergência do método dos gradientes conjugados. A desvantagem é que, nosso sistema pode ser relativamente grande, e N iterações pode representar uma convergência lenta. Apresentamos outro resultado de convergência para o método no Teorema 7 abaixo.

Teorema 7 *Seja A uma matriz simétrica, definida positiva, com autovetores $\{\mathbf{v}_i\}_{i=1}^N$. Seja*

$$\mathbf{b} = \sum_{j=1}^k \gamma_j \mathbf{v}_{i_j},$$

então o algoritmo de gradientes conjugados com $\mathbf{x}_0 = 0$ encontra a solução em no máximo k iterações.

Demonstração: A demonstração pode ser encontrada em [3, p. 15]. \square

Os teoremas e resultados vistos anteriormente nos mostram que o método dos gradientes conjugados converge para a solução exata do sistema linear em um número finito (mesmo que grande) de iterações. Entretanto, na prática não estamos interessados na solução exata do sistema, mas sim em uma solução aproximada, de modo que a norma do resíduo seja limitada superiormente por algum parâmetro. Note que esta é uma das condições de parada no Algoritmo1 dado na seção 2.2.

O critério usual de parada é do resíduo relativo:

$$\|\mathbf{r}_k\|_2 \leq \eta \|\mathbf{r}_0\|_2, \quad \eta \in (0, 1). \quad (8)$$

Observe que, se a chute inicial for $\mathbf{x}_0 = \mathbf{0}$, a expressão em (8) se reduz a

$$\|\mathbf{b} - A\mathbf{x}_k\|_2 \leq \eta \|\mathbf{b}\|_2.$$

Os Lemas 8 e 9 estabelecer uma relação entre a norma-2 do resíduo e a norma- A do erro em \mathbf{x}_k , usando os seguintes lemas:

Lema 8 *Seja A uma matriz simétrica definida positiva e $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ seus autovalores. Então, para todo vetor $\mathbf{z} \in \mathbb{R}^n$, temos*

$$\sqrt{\lambda_n} \|\mathbf{z}\|_A \leq \|A\mathbf{z}\|_2 \leq \sqrt{\lambda_1} \|\mathbf{z}\|_A. \quad (9)$$

Lema 9 *Seja A uma matriz simétrica definida positiva e o sistema $A\mathbf{x} = \mathbf{b}$. Então*

$$\frac{\|\mathbf{b} - A\mathbf{x}_0\|_2}{\|\mathbf{b} - A\mathbf{x}_k\|_2} \leq \sqrt{\frac{\lambda_1}{\lambda_N}} \frac{\|\mathbf{x}_k - \mathbf{x}_*\|_A}{\|\mathbf{x}_0 - \mathbf{x}_*\|_A}. \quad (10)$$

Observe que em (10), o número de condição de A ($cond_2(A) = \frac{\lambda_1}{\lambda_N}$) desempenha um papel importante na relação entre a norma-2 do resíduo e a norma- A do erro, pois a partir dele conseguimos um limitante superior para o número de iterações necessários para reduzir a norma do resíduo relativo. Para matrizes com o número de condição alto (i.e., matrizes mal condicionadas), o método pode fazer um número grande de iterações até reduzir a norma-2 do resíduo.

Podemos obter uma relação mais precisa do número máximo de iterações que o método poderá realizar. Se os autovalores de A estiverem agrupados em um número pequeno de intervalos, o número de condição de A pode ser grande e mesmo assim o método dos gradientes conjugados pode funcionar muito bem. Isto acontece porque quanto mais agrupados estiverem os autovalores, menor é o valor em módulo que um polinômio residual pode assumir no intervalo, garantindo a convergência mais veloz. O exemplo a seguir foi apresentado em [3, p. 18–19] e comprova este fato.

Exemplo 1 Assuma que $\mathbf{x}_0 = 0$ e que os autovalores de A estão contidos nos intervalos $(1, 1.5)$ e $(399, 400)$. Com essa informação podemos supor que o número de condição de A é $\text{cond}_2(A) \leq 400$. De [1] temos a seguinte relação sobre o erro em \mathbf{x}_k :

$$\|\mathbf{x}_k - \mathbf{x}_*\|_A \leq 2\|\mathbf{x}_0 - \mathbf{x}_*\|_A \left[\frac{\sqrt{\text{cond}_2(A)} - 1}{\sqrt{\text{cond}_2(A)} + 1} \right]^k. \quad (11)$$

Substituindo o número de condição de A em (11), temos

$$\frac{\|\mathbf{x}_k - \mathbf{x}_*\|_A}{\|\mathbf{x}_*\|_A} \leq 2(19/21)^k \approx 2(0.91)^k.$$

que é uma convergência lenta, dado a proximidade de um do termo elevado a k . Entretanto, se usarmos como polinômio residual o polinômio $\bar{p}_{3k}(z) \in \mathcal{P}_{3k}$ tal que

$$\bar{p}_{3k}(z) = \frac{(1.25 - z)^k (400 - z)^{2k}}{(1.25)^k (400)^{2k}}.$$

Podemos ver que

$$\max_{z \in \sigma(A)} |\bar{p}_{3k}(z)| \leq (0.25/1.25)^k = (0.2)^k$$

que é uma estimativa melhor para a convergência. De fato, se quisermos estimar em quantas iterações teríamos

$$\|\mathbf{x}_k - \mathbf{x}_*\|_A \leq 10^{-3} \|\mathbf{x}_*\|_A,$$

ou seja, para qual k $2(0.91)^k < 10^{-3}$, teríamos

$$k > -\log_{10}(2000)/\log_{10}(0.91) \approx 3.3/0.04 = 82.5.$$

A estimativa baseada no agrupamento dos autovalores nos dá

$$(0.2)^k \leq 10^{-3} \Leftrightarrow k > -3/\log_{10}(0.2) = 4.3.$$

Ou seja, enquanto (11) estima 83 iterações, a outra análise, se baseando no agrupamento dos autovalores nos dá 15 iterações (note que 15 é o maior inteiro múltiplo de 3 que é maior do que $3 \times 4.3 = 12.9$).

2.3 Testes computacionais

Agora vamos analisar, a partir de testes computacionais feitos com o auxílio do MatLab a convergência do método dos gradientes conjugados. Para os testes, usamos sistemas lineares $A\mathbf{x} = \mathbf{b}$, com dimensão $n = 1000$.

Primeiramente, geramos um vetor aleatório $\mathbf{u} \in \mathbb{R}^{1000}$, distribuído uniformemente entre 0 e 1. A partir deste vetor, geramos uma matriz $U = I - (2\mathbf{u}\mathbf{u}^T)/(\mathbf{u}^T\mathbf{u})$ que é ortogonal. A

partir desta, montamos para cada teste uma matriz diagonal D , com os autovalores de A na diagonal, e fizemos $A = UDU^T$, obtendo assim a matriz de coeficientes desejada. Geramos o vetor \mathbf{b} aleatoriamente no intervalo $[-25, 25]$. Usamos o método implementado pelo `MatLab`, tendo como chute inicial o vetor nulo, e precisão de 10^{-8} . A Tabela 1 mostra o resultados obtidos.

Podemos notar que nos problemas 1 e 2 o número de iterações foi elevado, devido ao espalhamento dos autovalores. Os problemas 3 e 4 convergiram em uma iteração, conforme o esperado, dado que possuem apenas um autovalor. Note ainda que o resíduo relativo foi tão pequeno quanto a precisão da máquina (10^{-16}). O problema 5 teve também uma convergência rápida, e essa convergência também deve-se ao fato de possuir apenas dois autovalores distintos. O problema 6 não convergiu, devido a singularidade da matriz que estamos tratando.

Os problemas 7, 8 e 9 apresentam também autovalores espalhados, porém estão mais agrupados do que os autovalores dos problemas 1 e 2. Observe que nos problemas 7, 8 e 9 temos 2, 3 e 10 clusters, respectivamente. Portanto os três problemas convergiram em menos iterações. Note ainda que, dentre os três problemas, os problema 7 foi o que apresentou convergência mais rápida, devido ao menor número de clusters.

Já no problema 10, podemos perceber que a matriz A é mal-condicionada, pois a condição de A é da ordem de 10^9 . Ou seja, como esperado, o método dos gradientes conjugados demorou a convergir para este problema.

Geramos também uma solução conhecida para o sistema linear, um vetor aleatório \mathbf{x}_* de 0's e 1's, e fizemos $\mathbf{b} = A\mathbf{x}_*$, para a matriz A do problema 10. Neste caso, o método convergiu em 71 iterações, com um resíduo relativo de 9.6×10^{-9} . Com a solução obtida, calculamos a norma do erro em \mathbf{x} , ou seja fizemos $\|\mathbf{x} - \mathbf{x}_*\| = 5.8 \times 10^{-3}$. Ou seja, apesar de o resíduo relativo ser menor do que a tolerância pedida, o vetor \mathbf{x} não está igualmente próximo da solução do sistema linear.

Tabela 1: Resultados obtidos via gradientes conjugados

Problema	Intervalo	Distribuição autovalores	Iterações	Resíduo relativo	Convergiu?
1	$\lambda_i \in [1, 1000]$	$\lambda_i = i$ $i = 1, \dots, 1000$	117	9.30×10^{-9}	sim
2		λ_i aleatório $i = 1, \dots, 1000$	161	6.50×10^{-9}	sim
3		$\lambda_i = 1$ $i = 1, \dots, 1000$	1	0	sim
4		$\lambda_i = 500$ $i = 1, \dots, 1000$	1	1.00×10^{-16}	sim
5		$\lambda_1 = 1$ $\lambda_i = 500$ $i = 2, \dots, 1000$	2	3.80×10^{-10}	sim
6		$\lambda_1 = 0$ $\lambda_i = 500$ $i = 2, \dots, 1000$	1	2.30×10^{-2}	não
7	$\lambda_i \in [400, 500]$ $\lambda_i \in [100, 200]$	λ_i aleatório $i = 1, \dots, 500$ $i = 501, \dots, 1000$	10	1.20×10^{-9}	sim
8	$\lambda_i \in [100, 200]$ $\lambda_i \in [1600, 1700]$ $\lambda_i \in [2800, 2900]$	λ_i aleatório $i = 1, \dots, 300$ $i = 301, \dots, 600$ $i = 601, \dots, 1000$	15	6.30×10^{-9}	sim
9	$\lambda_i \in [100, 200]$ $\lambda_i \in [400, 500]$ $\lambda_i \in [700, 800]$ $\lambda_i \in [1000, 1100]$ $\lambda_i \in [1300, 1400]$ $\lambda_i \in [1600, 1700]$ $\lambda_i \in [1900, 2000]$ $\lambda_i \in [2200, 2300]$ $\lambda_i \in [2500, 2600]$ $\lambda_i \in [2800, 2900]$	λ_i aleatório $i = 1, \dots, 100$ $i = 101, \dots, 200$ $i = 201, \dots, 300$ $i = 301, \dots, 400$ $i = 401, \dots, 500$ $i = 501, \dots, 600$ $i = 601, \dots, 700$ $i = 701, \dots, 800$ $i = 801, \dots, 900$ $i = 901, \dots, 1000$	24	5.10×10^{-9}	sim
10		$\lambda_i = 10^{-6}$ $\lambda_i = i$ $i = 2, \dots, 1000$	204	9.50×10^{-9}	sim

3 GMRES

Passaremos agora ao estudo do GMRES (Generalized Minimal Residual), outro método iterativo para resolução de sistemas lineares. Esta seção será baseada em [4, 3].

O método foi proposto por Saad e Schultz, em 1986 [5]. Ele se diferencia do método dos gradientes conjugados por não ser restrito a matrizes simétricas definidas positivas. Porém, com isso temos que o uso dos polinômios residuais torna mais difícil, assim como não temos garantia da aplicação do teorema espectral.

3.1 Aspectos teóricos

vamos supor nesta seção que o chute inicial dado para o método do GMRES é $\mathbf{x}_0 = \mathbf{0}$.

O GMRES gera uma sequência $\{\mathbf{x}_i\}$ de tal forma que, na iteração k , o vetor obtido é o que minimiza a norma-2 do resíduo $\|\mathbf{b} - A\mathbf{x}\|_2$ no subespaço de Krylov \mathcal{K}_k .

Se denotarmos por K_k a matriz cujas colunas são os vetores geradores do subespaço de Krylov na k -ésima iteração: $\mathbf{b}, A\mathbf{b}, \dots, A^{k-1}\mathbf{b}$, teremos que um vetor $\mathbf{v} \in \mathcal{K}_k$ pode ser escrito como $\mathbf{v} = K_k\mathbf{c}$, $\mathbf{c} \in \mathbb{R}^k$. Então, o problema de minimização (Quadrados Mínimos) equivale a:

$$\min_{\mathbf{x}_k \in \mathcal{K}_k} \|\mathbf{b} - A\mathbf{x}_k\|_2 \iff \min_{\mathbf{c} \in \mathbb{R}^k} \|\mathbf{b} - (AK_k)\mathbf{c}\|_2.$$

No segundo problema, obtemos \mathbf{c} que minimiza a norma-2 do resíduo. Para o obter o correspondente no subespaço de Krylov, fazemos $\mathbf{x}_k = K_k\mathbf{c}$.

Um modo de resolver o problema de Quadrados Mínimos é através da fatoração QR da matriz AK_k . Entretanto, este processo é instável e custoso. Um outro jeito de se resolver isso é usando o processo de Arnoldi, que descreveremos a seguir.

3.2 Iteração de Arnoldi

A iteração de Arnoldi é um processo semelhante ao processo de Gram-Schmidt, no sentido de gerar uma base ortonormal. No caso de Arnoldi, gera-se uma base ortonormal para o subespaço de Krylov.

Definição 10 Uma matriz H é denominada *Hessenberg superior* se tem a forma:

$$H = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1(n-1)} & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2(n-1)} & a_{2n} \\ 0 & a_{32} & a_{33} & \cdots & a_{3(n-1)} & a_{3n} \\ 0 & 0 & a_{43} & \cdots & a_{4(n-1)} & a_{4n} \\ \vdots & \vdots & \vdots & \ddots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & a_{n(n-1)} & a_{nn} \end{pmatrix}$$

Dada uma matriz A , a iteração de Arnoldi gera uma matriz *Hessenberg superior*, que seja unitariamente semelhante à matriz A : $A = QHQ^*$.

O processo de Arnoldi baseia-se no fato que, cada subespaço de Krylov de dimensão $j + 1$ pode ser obtido a partir da base ortogonal do subespaço de Krylov de dimensão j , usando o conjunto gerador $\{\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_j, A\mathbf{q}_j\}$.

Realizamos o processo da seguinte forma:

- O processo tem início com o vetor \mathbf{b} . A partir dele, obtemos $\mathbf{q}_1 = \mathbf{b}/\|\mathbf{b}\|_2$.
- O próximo vetor da sequência de Krylov é $A\mathbf{b}$, e como $A\mathbf{b} = \|\mathbf{b}\|_2(A\mathbf{q}_1)$ podemos considerar apenas $A\mathbf{q}_1$, pois $\text{span}\{\mathbf{b}, A\mathbf{b}\} = \text{span}\{\mathbf{q}_1, A\mathbf{q}_1\}$.
- O vetor \mathbf{q}_2 deve ser obtido de modo que $\text{span}\{\mathbf{q}_1, \mathbf{q}_2\} = \text{span}\{\mathbf{q}_1, A\mathbf{q}_1\}$, e ainda ser ortogonal a \mathbf{q}_1 e ter norma-2 igual a 1.
- Continuaremos este processo e, na etapa j teremos $\{\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_j\}$ é uma base ortogonal para o subespaço de Krylov \mathcal{K}_j .
- Temos então que $\mathcal{K}_{j+1} = \{\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_j, A\mathbf{q}_j\}$ e daí ortogonalizamos $A\mathbf{q}_j$ com relação aos outros vetores.

Matricialmente, teremos na etapa k : $AQ_k = Q_{k+1}\tilde{H}_k$, onde Q_k é ortogonal, e \tilde{H}_k é uma matriz $(k + 1) \times k$ e tem estrutura de uma matriz de Hessenberg de ordem k , com mais uma linha com entradas iguais a zero para $j = 1, \dots, k$ e na entrada $(k + 1, k)$, temos a componente $h_{k+1,k}$. Observe, então que podemos escrever a relação matricial como:

$$AQ_k = Q_k H_k + h_{k+1,k} \mathbf{q}_{k+1} \mathbf{e}_k^T. \quad (12)$$

Note ainda que, se $h_{k+1,k} = 0$ em (12), teremos que $A\mathcal{K}_k \subseteq \mathcal{K}_k$, ou seja, \mathcal{K}_k será um espaço invariante sobre A . Assim, teremos $\mathcal{K}_j = \mathcal{K}_r$, $r > j > k$ e a solução do sistema linear $A\mathbf{x} = \mathbf{b}$ pertence a \mathcal{K}_k , se A não singular.

A seguir apresentamos o algoritmo de Arnoldi, que é usado para construir as matrizes Q_k, Q_{k+1} e \tilde{H}_s .

Algoritmo 2: Arnoldi

Entrada: \mathbf{b}
 $\mathbf{q}_1 = \mathbf{b}/\|\mathbf{b}\|_2$
for $s = 1, 2, \dots$ **do**
 $\mathbf{v} = A\mathbf{q}_s$
 for $i = 1, \dots, s$ **do**
 $h_{is} = \mathbf{q}_i^T \mathbf{v}$
 $\mathbf{v} = \mathbf{v} - h_{is}\mathbf{q}_i$
 $h_{s+1,s} = \|\mathbf{v}\|_2$
 $\mathbf{q}_{s+1} = \mathbf{v}/h_{s+1,s}$
 end for
end for

3.3 Resultados teóricos

Vamos considerar novamente o problema de Quadrados Mínimos que o GMRES resolve a cada iteração:

$$\min_{\mathbf{c} \in \mathbb{R}^k} \|\mathbf{b} - AK_k \mathbf{c}\|_2 \quad (13)$$

Agora, seja Q_k a matriz formada pela base ortogonal do k -ésimo subespaço de Krylov. Então, um vetor neste subespaço é escrito como $K_k \mathbf{c} = Q_k \mathbf{a}$. Então usando isso em (13):

$$\min_{\mathbf{c} \in \mathbb{R}^k} \|\mathbf{b} - AK_k \mathbf{c}\|_2 \iff \min_{\mathbf{a} \in \mathbb{R}^k} \|\mathbf{b} - AQ_k \mathbf{a}\|_2.$$

Usando a relação da iteração de Arnoldi, temos

$$\min_{\mathbf{a} \in \mathbb{R}^k} \|\mathbf{b} - AQ_k \mathbf{a}\|_2 \iff \min_{\mathbf{a} \in \mathbb{R}^k} \|\mathbf{b} - Q_{k+1} \tilde{H}_k \mathbf{a}\|_2.$$

Note então, que as colunas de Q_{k+1} são ortogonais, e por definição, $\mathbf{q}_1 = \frac{\mathbf{b}}{\|\mathbf{b}\|_2}$. Ou seja:

$$\mathbf{q}_j^T \mathbf{b} = \begin{cases} 0, & \text{se } j > 1; \\ \|\mathbf{b}\|_2, & \text{se } j = 1. \end{cases} \quad (14)$$

Usando o fato que uma matriz ortogonal preserva a norma-2 de um vetor, temos:

$$\min_{\mathbf{a} \in \mathbb{R}^k} \|\mathbf{b} - Q_{k+1} \tilde{H}_k \mathbf{a}\|_2 \iff \min_{\mathbf{a} \in \mathbb{R}^k} \|Q_{k+1}^T \mathbf{b} - \tilde{H}_k \mathbf{a}\|_2.$$

Agora, usando a expressão em (14), obtemos:

$$\min_{\mathbf{a} \in \mathbb{R}^k} \|Q_{k+1}^T \mathbf{b} - \tilde{H}_k \mathbf{a}\|_2 \iff \min_{\mathbf{a} \in \mathbb{R}^k} \|\|\mathbf{b}\|_2 \mathbf{e}_1 - \tilde{H}_k \mathbf{a}\|_2.$$

Sendo assim, vamos elaborar um algoritmo geral para o GMRES:

Algoritmo 3: GMRES

Entrada: $A, \mathbf{b}, \mathbf{x}_0 = \mathbf{0}$

$\mathbf{r}_0 \leftarrow \mathbf{b}$

$\mathbf{q}_1 = \mathbf{b}/\|\mathbf{b}\|_2$

for $k = 1, 2, 3, \dots$ **do**

 Use o processo de Arnoldi descrito no Algoritmo 2, com $s = k$

 Encontre (via fatoração QR de \tilde{H}_k) o vetor \mathbf{a} que minimiza $\|\|\mathbf{b}\|_2 \mathbf{e}_1 - \tilde{H}_k \mathbf{a}\|_2$

 Faça $\mathbf{x}_k = Q_k \mathbf{a}$

end for

No Algoritmo 3 devemos fazer um passo do método de Arnoldi a cada iteração do GMRES. Em vez disso podemos, em uma iteração m pré-estabelecida, reiniciar o processo usando como aproximação inicial o resíduo da iteração m . Esta alternativa faz o cálculo da iteração de Arnoldi ser mais rápida, e com isso o GMRES ser mais rápido por consequência. O processo também requer menos memória, já que armazena m vetores a cada ciclo. Tal processo é conhecido como GMRES com restart.

Porém, não podemos garantir que com esse reinício teremos a convergência do método, ou seja, existem casos em que o resíduo gerado, embora seja decrescente, não converge para zero. Em [5] é proposto um exemplo, de modo a verificar isso, usando $m = 1$ e a matriz A e o vetor \mathbf{b} dados por:

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Se escolhermos $\mathbf{x}_0 = \mathbf{0}$ e fizermos uma iteração do GMRES, vamos obter $\mathbf{x}_1 = \mathbf{0}$. Como escolhermos $m = 1$, teremos que os resíduos formam uma sequência estacionária.

3.4 Convergência

Vamos estudar agora a convergência do GMRES. Da mesma forma que fizemos no método dos gradientes conjugados, vamos escrever $\mathbf{x} \in \mathbf{x}_0 + \mathcal{K}_k$ como

$$\mathbf{x} = \mathbf{x}_0 + \sum_{j=0}^{k-1} \gamma_j A^j \mathbf{r}_0,$$

e daí

$$\mathbf{b} - A\mathbf{x} = \mathbf{b} - A\mathbf{x}_0 - \sum_{j=1}^{k-1} \gamma_{j-1} A^j \mathbf{r}_0.$$

Novamente chamamos $p(z) = 1 + \sum_{j=1}^{k-1} \gamma_{j-1} z^j$, e temos o seguinte resultado.

Teorema 11 *Seja A uma matriz não singular e seja \mathbf{x}_k a k -ésima iteração do GMRES. Então, para todo $\bar{p}_k \in \mathcal{P}_k$*

$$\|\mathbf{r}_k\|_2 = \min_{p \in \mathcal{P}_k} \|\bar{p}(A)\mathbf{r}_0\|_2 \leq \|\bar{p}_k(A)\mathbf{r}_0\|_2. \quad (15)$$

Segue ainda um corolário imediato deste teorema, gerando uma majoração para o resíduo relativo.

Corolário 12 *Sejam as mesmas condições do Teorema 11. Uma majoração para o resíduo relativo é dada por*

$$\frac{\|\mathbf{r}_k\|_2}{\|\mathbf{r}_0\|_2} \leq \|\bar{p}_k(A)\|_2.$$

Com este corolário, conseguimos provar ainda que o método GMRES possui um número finito de iterações. Ou seja, o método encontra a solução exata do sistema linear a menos de erros de arredondamento.

Teorema 13 *Seja A uma matriz não singular. Então o GMRES encontra a solução em no máximo N iterações.*

Conforme mencionado anteriormente, o GMRES é aplicado em geral para matrizes que não possuem simetria. Observe que com isso temos a desvantagem de não podermos usar a decomposição espectral para decompor a matriz A em qualquer caso. Quando A é diagonalizável, podemos escrever $A = V\Lambda V^{-1}$, com V uma matriz que pode ter componentes complexos. Novamente o polinômio $p(A) = Vp(\Lambda)V^{-1}$. Tal estrutura diagonal pode ser usado para provar o seguinte teorema:

Teorema 14 *Seja $A = V\Lambda V^{-1}$, não singular, diagonalizável. Seja \mathbf{x}_k a k -ésima iteração do GMRES. Então, para todo $\bar{p}_k \in \mathcal{P}_k$*

$$\frac{\|\mathbf{r}_k\|_2}{\|\mathbf{r}_0\|_2} \leq \text{cond}_2(V) \min_{p_k \in \mathcal{P}_k} \max_{z \in \sigma(A)} |p_k(z)|. \quad (16)$$

O resultado do Teorema 14 fornece um limitante superior em função do número de condição da matriz V . No entanto, o número de iterações está relacionado à distribuição dos autovalores. Podemos montar exemplos onde GMRES converge com um número baixo de iterações mesmo para V mal condicionada. Comprovamos este fato nos exemplos a seguir, retirados do livro [6]:

Exemplo 2 *Seja A uma matriz de dimensão 1000, cujas entradas são números reais com distribuição normal com média 2 e desvio padrão $0.5/\sqrt{1000}$. A Figura 1 mostra os autovalores da matriz A , distribuídos em um círculo de raio 0.5 e centrado em $z = 2$. O número de condição da matriz V neste caso foi de 1.1967×10^3 . A convergência neste caso foi em 10 iterações, uma convergência rápida, com o resíduo igual a 9.4×10^{-7} .*

Exemplo 3 *Modificamos agora a diagonal da matriz A , somando ao elemento a_{kk} o seguinte valor*

$$d_k = -2 + 2 \sin \theta_k + i \cos \theta_k, \quad \theta_k = \frac{k\pi}{999}, 0 \leq k \leq 999.$$

Assim, geramos uma matriz B , com autovalores diferentes dos autovalores de A . A Figura 2 mostra os autovalores da nova matriz. Aplicamos novamente o GMRES, o método levou 57 iterações para que o resíduo ficasse menor do que 10^{-6} . Observe o agrupamento dos autovalores de B está próximo à origem, o que fez a convergência ser mais lenta. O resíduo relativo na iteração 851 foi de 9.8×10^{-7} .

Este resultado formalizado no Teorema 15 apresentado em [5].

Teorema 15 *Assuma que existem ν autovalores de A com parte real não positiva e que os outros autovalores estão agrupados em um círculo centrado em $C > 0$ com o raio $R < C$. Então*

$$\min_{p \in \mathcal{P}_N, p(0)=1} \max_{z \in \sigma(A)} |\bar{p}_k(z)| \leq \left(\frac{R}{C}\right)^{N-\nu} \left(\frac{D}{d}\right)^\nu \quad (17)$$

onde $D = \max_{i=1, \dots, \nu, j=\nu+1, \dots, N} |\lambda_i - \lambda_j|$ e $d = \min_{i=1, \dots, \nu} |\lambda_i|$

A demonstração do Teorema 15 é feito em [5]. Podemos então dizer que o raio de convergência do GMRES depende da distribuição dos autovalores de A no plano complexo. Para uma convergência rápida, devemos ter tais autovalores agrupados distantes da origem. Tal distribuição é mais importante do que o número de condição de A .

3.5 Testes Computacionais

Realizamos alguns testes computacionais para comprovar o que estudamos a respeito do GMRES. Geramos matrizes de modo que os autovalores estivessem distribuídos de alguma forma no plano complexo. Para isto, criamos blocos de matrizes usando o mesmo processo usado no Exemplo 2, e depois pré e pós multiplicamos a matriz obtida por uma matriz ortogonal, obtendo uma matriz densa. Para matrizes com autovalores reais usamos o mesmo processo usado para gradientes conjugados. Usamos ainda que $N = 1000$, e testamos cada problema usando o GMRES sem restart e com restart, para 30, 50 e 100 iterações. Os resultados obtidos estão na Tabela 2.

A matriz do problema 1 é simétrica definida positiva, e seus autovalores pertencem ao intervalo $[1, 1000]$. Observe que a convergência não foi tão rápida neste caso, devido a este espalhamento dos autovalores, e note ainda que $\text{cond}_2(V) = 1$, pois V é uma matriz ortogonal. Agora, comparando o problema 1 com o problema 2 vemos a vantagem do primeiro. Isto é explicado se aplicarmos o Teorema 17 aos dois problemas observando que a majoração para o polinômio no problema 1 é melhor (ou seja, o máximo atingido é menor) do que o aplicado ao problema 2.

No problema 3 os autovalores também estão espalhados, porém separados em quatro grupos de autovalores. Observe que esta distribuição fez o GMRES realizar quase o mesmo número de iterações que foi realizado no problema 1.

O problema 4 funcionou muito bem. Os autovalores estão separados em dois grupos, e em cada grupo eles são iguais entre si. Ou seja, a convergência deu-se de maneira quase imediata, conseguindo um resíduo relativo que está muito próximo à precisão da máquina.

Nos problemas 5 e 6 usamos duas matrizes mal-condicionadas. A diferença é que no problema 5 geramos um autovalor próximo de zero e os outros restritos a um círculo centrado em 10 com raio 9, enquanto no problema 6 o círculo também foi centrado em 10, porém usamos um raio igual a 5. Percebemos que no primeiro problema houve convergência para o GMRES sem restart e para o com restart na iteração 100. Os outros dois restart usados não chegaram a uma solução em 1000 iterações. Já o problema 6 apresentou convergência em todos os modelos do GMRES usados. Houve também uma diferença entre o número de iterações que o GMRES sem restart fez para cada um destes problemas. Isto é explicado pela maior concentração de autovalores no problema 6.

Tabela 2: Resultados obtidos via GMRES

Problema	Distribuição autovalores	Restart	Iterações	Convergiu?	Resíduo relativo
1	$\lambda_i = i, i = 1, \dots, 1000$	0	170	Sim	6.8×10^{-9}
		30	355	Sim	9.7×10^{-9}
		50	290	Sim	8.6×10^{-9}
		100	222	Sim	6.8×10^{-9}
2	$\lambda_i = i - 501, i = 1, \dots, 500$ $\lambda_i = i - 500, i = 501, \dots, 1000$	0	999	Não	6×10^{-4}
		30	1000	Não	5.3×10^{-2}
		50	1000	Não	4.5×10^{-2}
		100	1000	Não	3.4×10^{-2}
3	$ \lambda_i - 100 \leq 50, i = 1, \dots, 250$ $ \lambda_i - 500 \leq 200, i = 251, \dots, 500$ $ \lambda_i + 150 \leq 100, i = 501, \dots, 750$ $ \lambda_i - 10 \leq 8, i = 751, \dots, 1000$	0	170	Sim	9.9×10^{-9}
		30	355	Sim	8.8×10^{-9}
		50	290	Sim	9.8×10^{-9}
		100	222	Sim	9.7×10^{-9}
4	$\lambda_i = -20, i = 1, \dots, 500$ $\lambda_i = 30, i = 501, \dots, 1000$	0	2	Sim	3×10^{-15}
		30	2	Sim	3×10^{-15}
		50	2	Sim	3×10^{-15}
		100	2	Sim	3×10^{-15}
5	$\lambda_1 = 10^{-6}$ $ \lambda_i - 10 \leq 9, i = 2, \dots, 1000$	0	227	Sim	9.8×10^{-9}
		30	1000	Não	1.5×10^{-2}
		50	1000	Não	1.1×10^{-2}
		100	398	Sim	8.8×10^{-9}
6	$\lambda_1 = 10^{-6},$ $ \lambda_i - 10 \leq 5, i = 2, \dots, 1000$	0	50	Sim	8.3×10^{-9}
		30	119	Sim	9.3×10^{-9}
		50	50	Sim	8.3×10^{-9}
		100	50	Sim	8.3×10^{-9}

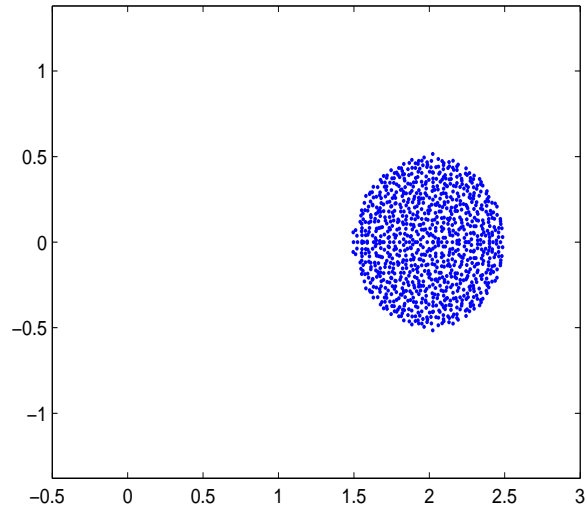


Figura 1: Distribuição dos autovalores da matriz A

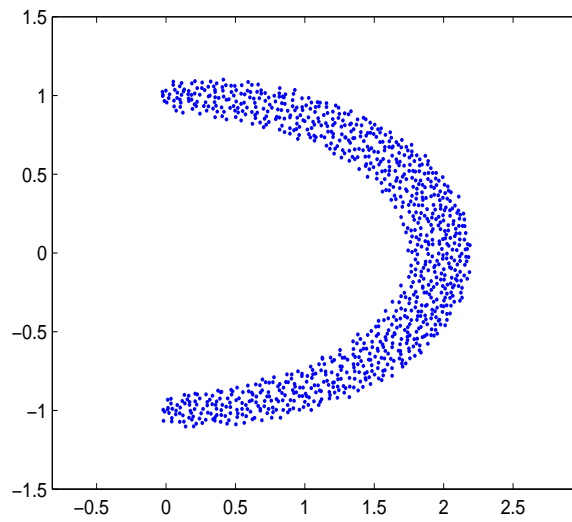


Figura 2: Distribuição dos autovalores da matriz B

Referências

- [1] J.W. Daniel. The conjugate gradient method for linear and nonlinear operator equations. *SIAM Journal on Numerical Analysis*, 4(1):10–26, 1967.

- [2] M.R. Hestenes and E. Stiefel. *Methods of conjugate gradients for solving linear systems*, 1952.
- [3] C.T. Kelley. *Iterative methods for linear and nonlinear equations*. Society for Industrial Mathematics, 1995.
- [4] Y. Saad. *Iterative methods for sparse linear systems*. PWS Pub. Co., 1996.
- [5] Y. Saad and M.H. Schultz. Gmres: A generalized minimal residual algorithm for solving nonsymmetric linear systems. *SIAM Journal on scientific and statistical computing*, 7(3):856–869, 1986.
- [6] L.N. Trefethen and D. Bau. *Numerical Linear Algebra*. Society for Industrial Mathematics, 1997.