UNIVERSIDADE FEDERAL DO PARANÁ
Diego Dutra Zontini

MÉTODOS COMPUTACIONAIS PARA INVERSAS GENERALIZADAS

Curitiba, 2014.

UNIVERSIDADE FEDERAL DO PARANÁ
Diego Dutra Zontini

MÉTODOS COMPUTACIONAIS PARA INVERSAS GENERALIZADAS

Tese apresentada ao Programa de Pós-Graduação em Matemática e Matemática Aplicada da Universidade Federal do Paraná, como requisito parcial à obtenção do grau de Doutor em Matemática Aplicada.

Orientador: Prof. Dr. Yuan Jin Yun.

Curitiba, 2014.

Ministério da Educação
Universidade Federal do Paraná
Setor de Ciências Exatas/Departamento de Matemática
Programa de Pós-Graduação em Matemática - PPGM

ATA DA 2ª DEFESA DE TESE DE DOUTORADO

Aos vinte e oito dias do mês de março de 2014, no Anfiteatro B, Bloco das PCs, da Universidade Federal do Paraná, foi instalada pelo Professor Yuan Jin Yun, a Banca Examinadora para a Segunda Defesa de Tese de Doutorado em Matemática. Estiveram presentes ao Ato, professores, alunos e visitantes.

A banca examinadora, homologada pelo Colegiado do Programa de Pós-Graduação em Matemática, ficou constituída pelos professores: Prof. Dr. Fermín Sinforiano Viloche Bazan da Universidade Federal de Santa Catarina, Prof. Dr. Wenyu Sun, NUN-China, Prof. Dr. Cássio Machiaveli Oishi, da Universidade Estadual Paulista Júlio de Mesquita Filho, Prof. Dr. Saulo Pomponet Oliveira, do Programa de Pós-Graduação em Matemática e o Prof. Dr. Yuan Jin Yun, orientador do projeto de tese, a quem coube a presidência dos trabalhos.

Às dez horas, a banca iniciou seus trabalhos, convidando o candidato **DIEGO DUTRA ZONTINI** a apresentar seu Projeto de Tese intitulado: "MÉTODOS COMPUTACIONAIS PARA INVERSAS GENERALIZADAS". Encerrada a apresentação, iniciou-se a fase de argüição pelos membros participantes. Após a argüição, a banca com pelo menos 05 (cinco) membros, reuniu-se para apreciação do desempenho do pós-graduando.

A banca considerou que o pós-graduando fez uma apresentação com a necessária concisão. A tese apresenta contribuição à área de estudos e não foram registrados problemas fundamentais de estrutura e redação, resultando em plena e satisfatória compreensão dos objetivos pretendidos.

Tendo em vista a tese e a argüição, os membros presentes da banca decidiram pela sua aprovação.

Curitiba, 28 de março de 2014.

Prof. Dr. Yuan Jin Yun
PPGMA – UFPR

Prof. Dr. Fermín S. V. Bazán
UFSC

Wenyu Sun
Prof. Dr. Wenyu Sun
NUN – China

Prof. Dr. Saulo Pomponet Oliveira
PPGMA – UFPR

Prof. Dr. Cassio Machiaveli Oishi
UNESP

Ministério da Educação
Universidade Federal do Paraná
Setor de Ciências Exatas/Departamento de Matemática
Programa de Pós-Graduação em Matemática - PPGM

PARECER DA BANCA EXAMINADORA

Após a apresentação, a banca deliberou pela aprovação da tese do candidato **DIEGO DUTRA ZONTINI** devendo, para tanto, incorporar as sugestões feitas pelos membros da banca, no prazo estabelecido pelo regimento correspondente.

Curitiba, 28 de março de 2014.

Orientador: Prof. Dr. Yuan Jin Yun
Dep. de Matemática – UFPR

Prof. Dr. Fermín S. V. Bazán
Dep. de Matemática – UFSC

Cassio M. Oishi

Prof. Dr. Cassio Machiaveli Oishi
Fac. de Ciências e Tecnologia – UNESP

Prof. Dr. Saulo Pomponet Oliveira
Dep. de Matemática – UFPR

Wenyu Sun

Prof. Dr. Wenyu Sun
NUN – China

TERMO DE APROVAÇÃO

MÉTODOS COMPUTACIONAIS PARA INVERSAS GENERALIZADAS

por

DIEGO DUTRA ZONTINI

Tese aprovada como requisito parcial para obtenção do grau de Doutor no Programa de Pós-Graduação em Matemática Aplicada, pela Comissão Examinadora composta por:

Orientador: Prof. Dr. Yuan Jin Yun
Dep. de Matemática - UFPR

Prof. Dr. Fermín S. V. Bazán
Dep. de Matemática – UFSC

Prof. Dr. Cassio Machiaveli Oishi
Fac. de Ciências e Tecnologia – UNESP

Prof. Dr. Saulo Pomponet Oliveira
Dep. de Matemática – UFPR

Wenyu Sun

Prof. Dr. Wenyu Sun
NUN – China

Curitiba, 28 de março de 2014.

Aos meus pais e esposa.

Agradecimentos

Aos meus amigos, minha família e em especial a minha esposa agradeço pela ajuda em todos os momentos, e ao tempo e apoio dispensados ao longo da elaboração desta tese.

Ao meu orientador, professor Dr. Yuan Jin Yun, agradeço pela confiança em mim depositada, pela orientação e apoio no desenvolvimento desta tese e por todos os conhecimentos comigo compartilhados ao longo do período que trabalhamos juntos.

Aos professores Dr. Fermín S. V. Bazán, Dr. Luiz Carlos Matioli, Dr. Saulo P. Oliveira, Dr. Wenyu Sun agradeço por aceitarem participar da banca de pré-defesa e pelas sugestões muito importantes para esse trabalho, em especial gostaria de agradecer o professor Dr. Fermin S. V. Bazán pelas contribuições após a pré-defesa.

Aos professores Dr. Cassio M. Oishi, Dr. Fermín S. V. Bazán, Dr. Saulo P. Oliveira e Dr. Wenyu Sun , agradeço por aceitarem participar da banca examinadora deste trabalho.

Agradeço aos professores que ao longo de graduação, mestrado e doutorado contribuíram para a minha formação acadêmica.

Ao Programa de Pós-Graduação em Matemática e Matemática Aplicada da UFPR.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES, pelo apoio financeiro e pela formação de qualidade propiciada.

"A matemática, vista corretamente, possui não apenas verdade, mas também suprema beleza - uma beleza fria e austera, como a da escultura. "

Bertrand Russell

Resumo

Neste trabalho apresentamos aspectos teóricos e computacionais sobre inversas generalizadas de matrizes e propomos três novos métodos computacionais.

Propomos inicialmente um método direto baseado em decomposição conjugada para calcular a inversa de Moore-Penrose, no qual provamos que a inversa de Moore-Penrose de uma matriz A pode ser obtida por $A^\dagger = Z\Gamma^{-1}Q^*$ se A tem posto completo e $A^\dagger = (U_1^*\Gamma_1^{-1}S_1, 0)Q^*$ caso contrário, sendo $A = Q\Gamma Z^{-1}$ uma decomposição conjugada de A e $S_1^{-1}\Gamma_1 U_1$ uma decomposição conjugada da matriz obtida pelas r primeiras linhas de Q^*A, onde $r =$posto(A).

Em seguida, propomos um método direto baseado em decomposição conjugada para calcular a inversa de Drazin, o qual consiste em um processo de deflação ortogonal que usa k decomposições conjugadas, sendo $k =$Ind(A), para construir a inversa de Drazin da forma

$$A_d = W^* \begin{pmatrix} B_1^{-1} & 0 \\ XB_1^{-1} & 0 \end{pmatrix} W, \quad \text{onde} \quad A = W^* \begin{pmatrix} B_1 & 0 \\ B_2 & N \end{pmatrix} W,$$

sendo B_1 não singular, N estritamente triangular inferior, W unitária e X solução de $XB_1 - NX = B_2$.

E por fim, propomos um método iterativo para aproximar a inversa de Moore-Penrose baseado nas equações de Penrose $AXA = A$ e $XAX = X$, o qual considera uma aproximação inicial $X_0 = \alpha A^*$, e calcula $X_{k+1} = X_k[(1+\beta)I - \beta Y_k^2]$, onde $Y_k = AX_k$. A convergência do método é provada para $0 < \alpha < 2/\rho(A^*A)$ e $\beta \in (0, 1/3]$, além disso propriedades e análises de erros são mostradas.

Palavras-chave: *Inversas generalizadas, Inversa de Moore-Penrose, Inversa de Drazin, Método direto, Método iterativo, Decomposição conjugada.*

Abstract

In this work we present theoretical and computational aspects about generalized inverses of matrices. Three new computational methods are proposed.

We propose a direct method based on conjugate decomposition for computing the Moore-Penrose inverse, which we proved that $A^{\dagger} = Z\Gamma^{-1}Q^*$ if A is a full rank matrix and $A^{\dagger} = (U_1^*\Gamma_1^{-1}S_1, 0)Q^*$ otherwise, where $A = Q\Gamma Z^{-1}$ is a conjugate decomposition of A and $S_1^{-1}\Gamma_1 U_1$ is a conjugate decomposition of the matrix obtained by the r first rows of Q^*A. Here r =rank(A).

Afterwards, we propose a direct method based on conjugate decomposition for computing the Drazin inverse, which consists of a orthogonal deflation method using k conjugate decompositions, where k =Ind(A), which is employed to build the Drazin inverse of A as follows

$$A_d = W^* \begin{pmatrix} B_1^{-1} & 0 \\ XB_1^{-1} & 0 \end{pmatrix} W, \quad \text{where} \quad A = W^* \begin{pmatrix} B_1 & 0 \\ B_2 & N \end{pmatrix} W,$$

with B_1 nonsingular, N strictly lower triangular, W unitary and X satisfy $XB_1 - NX = B_2$.

Finally, one iterative approach is proposed for the computation of the Moore-Penrose inverse based on Penrose equations $AXA = A$ and $XAX = X$, which considers an initial approach $X_0 = \alpha A^*$, and computes $X_{k+1} = X_k[(1+\beta)I - \beta Y_k^2]$, where $Y_k = AX_k$. The convergence of the method is proved for $0 < \alpha < 2/\rho(A^*A)$ and $\beta \in (0, 1/3]$, along with some properties and an error analysis.

Keywords: *Generalized inverses, Moore-Penrose inverse, Drazin inverse, Direct method, Iterative method, Conjugate decomposition.*

Sumário

Introdução

Seja $A \in \mathbb{C}^{n\times n}$ uma matriz quadrada de ordem n com entradas complexas, sabemos que se A é não singular então existe uma única matriz X satisfazendo

$$AX = XA = I$$

sendo I a matriz identidade de ordem n. Essa matriz X é chamada de inversa de A e é denotada por $X = A^{-1}$. A inversa de uma matriz possui diversas propriedades interessantes, por exemplo

$$\begin{aligned}
(A^{-1})^{-1} &= A, \\
(A^{T})^{-1} &= (A^{-1})^{T}, \\
(A^{*})^{-1} &= (A^{-1})^{*}, \\
(AB)^{-1} &= B^{-1}A^{-1}.
\end{aligned}$$

Uma aplicação muito importante da matriz inversa, é que se considerarmos um sistema não singular de equações lineares $Ax = b$, com $A \in \mathbb{C}^{n\times n}$ não singular e $b \in \mathbb{C}^{n}$, a única solução x do sistema é dada por

$$x = A^{-1}b.$$

De maneira mais geral, é comum em problemas das mais diversas áreas, sistemas de equações lineares da forma

$$Ax = b \text{ com } A \in \mathbb{C}^{m\times n} \text{ e } b \in \mathbb{C}^{m} \tag{1}$$

Nesse caso, nem sempre teremos um único vetor satisfazendo a equação (1), podemos ter infinitos vetores que satisfazem a equação acima, ou até mesmo nenhum vetor satisfazendo a equação, como é o caso quando b não pertence a imagem de A. Isso requer que modifiquemos o conceito de solução de um sistema linear dado por (1). Para isso, podemos definir um vetor residual por

$$r := b - Ax.$$

No caso em que b pertence a imagem de A, o residuo r jamais será nulo, porém uma boa maneira de considerar uma solução para o problema é tomar x de forma que o resíduo seja minimizado em algum sentido. Temos diversas maneiras de medir o tamanho do resíduo, em geral usamos alguma norma, e por questões geométricas a norma escolhida para medir o resíduo na maioria dos problemas é a norma euclidiana, denotada por $\| \cdot \|_2$.

No caso em que existem infinitos vetores safisfazendo a equação (1), precisamos escolher um desses vetores como solução do problema, isso pode ser feito tomando como solução do problema o vetor de norma mínima que satisfaz a equação, ou seja, elegemos como solução o vetor $\overline{x}$ que é solução do seguinte problema de otimização:

$$\overline{x} = \min_{\text{s. a. } Ax=b} \|x\|_2.$$

Com as definições de solução para o problema (1) dadas acima, surge a seguinte questão: será que podemos encontrar alguma matriz X, de tal forma que possamos escrever a solução do problema (1) como $x = Xb$? A resposta é sim, podemos definir uma matriz X nessas condições, a matriz X chamada de inversa de Moore-Penrose e denotada por $A^\dagger$ resolve essa questão. De maneira mais geral podemos falar de inversas generalizadas, que são matrizes definidas por certas condições, de tal forma que quando consideramos uma matriz não singular, sua inversa generalizada se reduz a inversa convencional, além de possuir algumas propriedades similares a inversa usual. Algumas dessas inversas generalizadas nos permite caracterizar soluções de problemas do tipo (1), assim como possuem diversas outras aplicações nas mais variadas áreas.

Neste trabalho, faremos um amplo estudo sobre inversas generalizadas e os métodos obtidos para calcular tais matrizes. Duas classes de métodos são estudadas, os métodos diretos, que encontram a matriz em uma quantidade finita de operações aritméticas, e os métodos iterativos, que a partir de uma aproximação inicial X_0 geram uma sequência de matrizes $X_1, X_2, \ldots$ de tal forma que no limite essa sequência converge para a solução X do problema, ou seja,

$$\lim_{k \to \infty} X_k = X.$$

Com base nos métodos estudados, são propostos três novos métodos para aproximar inversas generalizadas, sendo um método direto para encontrar a inversa de Drazin, o qual é baseado em deflação ortogonal usando decomposição conjugada, um método direto para encontrar a inversa de Moore-Penrose usando decomposição conjugada e um método iterativo para aproximar a inversa de Moore-Penrose, o qual utiliza dois parâmetros, um para a aproximação inicial, dependendo somente do raio espectral de A^*A, e outro para a iteração, o qual não depende da matriz e influencia diretamente na velocidade de convergência do método.

Essa tese está estruturada da seguinte forma: no primeiro capítulo abordamos o conceito de inversa generalizada, bem como os principais tipos de inversas generalizadas e suas propriedades. Além disso são apresentadas algumas aplicações de inversas generalizadas.

No segundo capítulo faremos um estudo de alguns métodos para computar inversas generalizadas existentes na literatura, veremos método diretos e métodos iterativos.

No capítulo 3 propomos dois novos métodos, os quais são baseados em decomposição conjugada de matrizes, e apresentamos os resultados e aspectos teóricos importantes de cada método, assim como alguns exemplos numéricos.

No quarto capítulo propomos um novo método iterativo para aproximar a inversa de Moore-Penrose, o qual é baseado nas duas primeiras equações de Penrose. Além disso, apresentamos uma análise do erro, mostrando que a convergência é linear, um estudo do comportamento do método diante da presença de erros de arredondamento, e exemplos numéricos que mostram o funcionamento do método.

Concluímos fazendo algumas análises a respeito dos métodos propostos, discutindo as vantagens e desvantagens de cada método nos sentidos numérico e teórico, e as possíveis áreas de trabalho futuro.

Capítulo 1

Inversas generalizadas

O conceito de inversa generalizada foi introduzido por Fredholm em 1903, que chamou de “pseudo inversa” a inversa generalizada de um operador integral. A inversa generalizada de matrizes foi introduzida por Moore em 1920, que definiu uma única inversa generalizada por meio de projeção de matrizes. Pouco se fez nos próximos 30 anos, até aproximadamente 1950, quando algumas relações entre inversa generalizada e solução de problemas de mínimos quadráticos foram descobertas. Em 1955 Penrose mostrou que a inversa de Moore é a única matriz que satisfaz quatro equações, chamadas hoje em dia de equações de Penrose. Devido a esse fato, essa pseudo inversa é comumente chamada de inversa de Moore-Penrose, em homenagem as contribuições de Moore e Penrose.

Neste capítulo, apresentamos os conceitos, propriedades e aplicações de inversas generalizadas de matrizes.

Quando falamos em generalização, algumas condições surgem naturalmente, nesse caso não é diferente, para caracterizar uma inversa generalizada de uma matriz, pedimos que sejam satisfeitas as seguintes condições:

(i) Existe a inversa generalizada para uma classe de matrizes mais ampla que a classe de matrizes não singulares;

(ii) A inversa generalizada deve satisfazer algumas propriedades da inversa usual;

(iii) Se a matriz é não singular, sua inversa generalizada deve coincidir com a inversa usual.

Por exemplo, podemos dizer que X é uma inversa generalizada de A se satisfaz

$$AXA = A.$$

É fácil verificar que uma matriz X satisfazendo a equação acima contempla as condições descritas anteriormente, porém uma propriedade muito interessante da inversa usual é a unicidade, a qual não é mantida nesse caso. Vejamos algumas inversas generalizadas de grande importância na literatura.

1.1 Inversa de Moore-Penrose

Moore, em 1920 [18], definiu uma inversa generalizada e provou sua unicidade. Alguns anos depois, em 1955, Penrose [22] mostrou que para toda matriz $A \in \mathbb{C}^{m\times n}$ existe uma única matriz $X \in \mathbb{C}^{n\times m}$ satisfazendo as equações:

$$
\begin{aligned}
AXA &= A; && (1.1)\\
XAX &= X; && (1.2)\\
(AX)^* &= AX; && (1.3)\\
(XA)^* &= XA. && (1.4)
\end{aligned}
$$

Denotamos como A^* a transposta conjugada da matriz A.

As equações (1.1)-(1.4) são conhecidas como equações de Penrose. A matriz X que satisfaz as equações de Penrose coincide com a inversa generalizada definida por Moore, e por isso inversa de Moore-Penrose. A inversa de Moore-Penrose da matriz A será denotada por $A^\dagger$.

No teorema seguinte, vamos mostrar que a matriz X satisfazendo as quatro equações de Penrose existe e é única.

Teorema 1.1. *[29] A matriz X satisfazendo as equações de Penrose (1.1)-(1.4) existe e é única.*

Demonstração:

Se $A \in \mathbb{C}^{m\times n}$ é uma matriz de posto r, então A pode ser decomposta como $A = Q^*RP$, onde Q é uma matriz unitária e P uma matriz de permutação, de ordens m e n respectivamente, e $R \in \mathbb{C}^{m\times n}$ satisfaz

$$R = \begin{pmatrix} R_{11} & 0 \\ 0 & 0 \end{pmatrix},$$

com $R_{11} \in \mathbb{C}^{r\times r}$ triangular superior e não singular. Sejam

$$\overline{R} = \begin{pmatrix} R_{11}^{-1} & 0 \\ 0 & 0 \end{pmatrix},$$

e $X = P^*\overline{R}Q$. Temos que

(i)
$$AXA = Q^*RPP^*\overline{R}QQ^*RP = Q^*R\overline{R}RP = Q^*RP = A$$

(ii)
$$XAX = P^*\overline{R}QQ^*RPP^*\overline{R}Q = P^*\overline{R}R\overline{R}Q = P^*\overline{R}Q = X$$

(iii)
$$\begin{aligned}(AX)^* &= (Q^*RPP^*\overline{R}Q)^* = (Q^*R\overline{R}Q)^* = \left(Q\begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix}Q^*\right)^* \\ &= Q\begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix}Q^* = AX\end{aligned}$$

(iv)
$$\begin{aligned}(XA)^* &= (P^*\overline{R}QQ^*RP)^* = (P^*\overline{R}RP)^* = \left(P^*\begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix}P\right)^* \\ &= P^*\begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix}P = XA\end{aligned}$$

Portanto, $X = A^\dagger$ de onde segue a existência.

Para mostrar a unicidade, suponha que X_1 e X_2 são inversas de Moore-Penrose de A, então X_1 e X_2 satisfazem as equações (1.1)-(1.4), e assim

$$\begin{aligned}X_1 &= X_1AX_1 = X_1AX_2AX_1 = X_1(AX_2)^*(AX_1)^* = X_1(AX_1AX_2)^* \\ &= X_1(AX_2)^* = X_1AX_2 = X_1AX_2AX_2 = (X_1A)^*(X_2A)^*X_2 \\ &= (X_2AX_1A)^*X_2 = (X_2A)^*X_2 = X_2AX_2 = X_2.\end{aligned}$$

De onde segue a unicidade. ■

O próximo teorema mostra algumas propriedades que são satisfeitas pela inversa de Moore-Penrose.

Teorema 1.2. *[29] Seja $A \in \mathbb{C}^{m\times n}$ uma matriz arbitrária. Temos que as seguintes propriedades são válidas:*

1. $(A^\dagger)^\dagger = A$;
2. $(\lambda A)^\dagger = \lambda^\dagger A^\dagger$, *onde* $\lambda \in \mathbb{C}$ *e* $\lambda^\dagger = \begin{cases} \frac{1}{\lambda} & se \quad \lambda \neq 0, \\ 0 & se \quad \lambda = 0; \end{cases}$
3. $(A^*)^\dagger = (A^\dagger)^*$;

4. $(AA^*)^\dagger = (A^*)^\dagger A^\dagger$ *e* $(A^*A)^\dagger = A^\dagger (A^*)^\dagger$*;*

5. $A^\dagger = (A^*A)^\dagger A^* = A^*(AA^*)^\dagger$*;*

6. $A^* = A^*AA^\dagger = A^\dagger AA^*$*;*

7. *Se* $posto(A) = n$*, então* $A^\dagger A = I_n$ *e se* $posto(A) = m$*, então* $AA^\dagger = I_m$*;*

8. *Se* U *e* V *são matrizes unitárias, então* $(UAV)^\dagger = V^* A^\dagger U^*$*.*

Demonstração:

1. Basta mostrar que A satisfaz as quatro equações de Penrose, de fato,

$$\begin{aligned} A^\dagger A A^\dagger &= A^\dagger \\ AA^\dagger A &= A \\ (A^\dagger A)^* &= A^\dagger A \\ (AA^\dagger)^* &= AA^\dagger. \end{aligned}$$

2. Da mesma forma, vamos mostrar que $\lambda^\dagger A^\dagger$ satisfaz as equações de Penrose. Se $\lambda = 0$, seguem trivialmente, considerando $\lambda \neq 0$ temos

$$\begin{aligned} (\lambda A)(\lambda^\dagger A^\dagger)(\lambda A) &= (\lambda A)\left(\frac{1}{\lambda}A^\dagger\right)(\lambda A) = \left(\lambda \frac{1}{\lambda}\lambda\right)(AA^\dagger A) = \lambda A \\ (\lambda^\dagger A^\dagger)(\lambda A)(\lambda^\dagger A^\dagger) &= \left(\frac{1}{\lambda}A^\dagger\right)(\lambda A)(\lambda^\dagger A^\dagger) = \left(\frac{1}{\lambda}\lambda\lambda^\dagger\right)(A^\dagger AA^\dagger) = \lambda^\dagger A^\dagger \\ ((\lambda A)(\lambda^\dagger A^\dagger))^* &= ((\lambda\lambda^\dagger)(AA^\dagger))^* = (\lambda\lambda^\dagger)^*(AA^\dagger)^* = (\lambda\lambda^\dagger)(AA^\dagger) = (\lambda A)(\lambda^\dagger A^\dagger) \\ ((\lambda^\dagger A^\dagger)(\lambda A))^* &= ((\lambda^\dagger\lambda)(A^\dagger A))^* = (\lambda^\dagger\lambda)^*(A^\dagger A)^* = (\lambda^\dagger\lambda)(A^\dagger A) = (\lambda^\dagger A^\dagger)(\lambda A) \end{aligned}$$

3. Novamente, basta mostrar que $(A^\dagger)^*$ satisfaz as quatro equações de Penrose.

$$\begin{aligned} A^*(A^\dagger)^*A^* &= (AA^\dagger A)^* = A^* \\ (A^\dagger)^*A^*(A^\dagger)^* &= (A^\dagger AA^\dagger)^* = (A^\dagger)^* \\ (A^*(A^\dagger)^*)^* &= ((A^\dagger A)^*)^* = (A^\dagger A)^* = A^*(A^\dagger)^* \\ ((A^\dagger)^*A^*)^* &= ((AA^\dagger)^*)^* = (AA^\dagger)^* = (A^\dagger)^*A^* \end{aligned}$$

4. Usando o item anterior e as equações de Penrose para A, verifiquemos que valem as

equações de Penrose para AA^*.

$$\begin{aligned}
AA^*(A^*)^\dagger A^\dagger AA^* &= AA^*(A^*)^\dagger(A^\dagger A)^*A^* = AA^*(A^*)^\dagger A^*(A^*)^\dagger A^* \\
&= AA^*(A^*)^\dagger A^* = AA^*; \\
(A^*)^\dagger A^\dagger AA^*(A^*)^\dagger A^\dagger &= (A^*)^\dagger A^\dagger A(A^\dagger A)^*A^\dagger = (A^*)^\dagger A^\dagger AA^\dagger AA^\dagger \\
&= (A^*)^\dagger A^\dagger AA^\dagger = (A^*)^\dagger A^\dagger.
\end{aligned}$$

Para a terceira e quarta equações, precisamos verificar que

$$(AA^*(A^*)^\dagger A^\dagger)^* = AA^*(A^*)^\dagger A^\dagger \quad \text{e} \quad ((A^*)^\dagger A^\dagger AA^*)^* = (A^*)^\dagger A^\dagger AA^*.$$

Para isso basta notar que

$$AA^*(A^*)^\dagger A^\dagger = A(A^\dagger A)^*A^\dagger = AA^\dagger AA^\dagger = AA^\dagger$$

e

$$(A^*)^\dagger A^\dagger AA^* = (A^*)^\dagger(A^\dagger A)^*A^* = (A^*)^\dagger A^*(A^*)^\dagger A^* = (A^*)^\dagger A^*,$$

e como $(AA^\dagger)^* = AA^\dagger$ e $((A^*)^\dagger A^*)^* = (A^*)^\dagger A^*$, segue que a terceira e quarta equações de Penrose são válidas.

O caso $(A^*A)^\dagger = A^\dagger(A^*)^\dagger$ segue usando o processo anterior em A^* e o item 3.

5. Usando as equações de Penrose e os itens 3 e 4 temos:

$$A^\dagger = A^\dagger AA^\dagger = A^\dagger(AA^\dagger)^* = A^\dagger(A^*)^\dagger A^* = (A^*A)^\dagger A^*$$

e

$$A^\dagger = A^\dagger AA^\dagger = (A^\dagger A)^*A^\dagger = A^*(A^*)^\dagger A^\dagger = A^*(AA^*)^\dagger$$

6. Segue das equações de Penrose e do item 3 que:

$$A^* = A^*(A^*)^\dagger A^* = A^*(AA^\dagger)^* = A^*AA^\dagger \text{ e } A^* = A^*(A^*)^\dagger A^* = (A^\dagger A)^*A^* = A^\dagger AA^*.$$

7. Se posto$(A) = n$, então posto$(A^*A) = n$, logo A^*A é não singular, assim pelo item 5 temos:

$$A^\dagger A = (A^*A)^\dagger A^*A = (A^*A)^{-1}A^*A = I_n.$$

Da mesma forma, se posto$(A) = m$ então posto$(AA^*) = m$, logo AA^* é não singular, e usando o item 5 segue que

$$AA^\dagger = AA^*(AA^*)^\dagger = AA^*(AA^*)^{-1} = I_m.$$

8. Basta verificar que são válidas as equações de Penrose

$$\begin{aligned}
UAVV^*A^\dagger U^*UAV &= UAA^\dagger AV = UAV \\
V^*A^\dagger U^*UAVV^*A^\dagger U^* &= V^*A^\dagger AA^\dagger U^* = V^*A^\dagger U^* \\
(UAVV^*A^\dagger U^*)^* &= (UAA^\dagger U^*)^* = U(AA^\dagger)^*U^* = UAA^\dagger U^* = UAVV^*A^\dagger U^* \\
(V^*A^\dagger U^*UAV)^* &= (V^*A^\dagger AV)^* = V^*(A^\dagger A)^*V = V^*A^\dagger AV = V^*A^\dagger U^*UAV
\end{aligned}$$

■

A propriedade 8 do Teorema anterior mostra que no caso de matrizes unitárias temos que a inversa de Moore-Penrose de um produto é o produto das inversas de Moore-Penrose, isso nos fornece uma maneira de calcular a inversa de Moore-Penrose usando a decomposição em valores singulares, ou seja, se

$$A = U\Sigma V^* = U \begin{pmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{pmatrix} V^*,$$

com U e V unitárias e Σ_1 diagonal, então temos que

$$A^\dagger = V\Sigma^\dagger U^* = V \begin{pmatrix} \Sigma_1^{-1} & 0 \\ 0 & 0 \end{pmatrix} U^*,$$

mais adiante veremos esse processo com mais detalhes.

A inversa de Moore-Penrose também possui diversas propriedades em relação ao núcleo e a imagem de A e A^*.

Definição 1.3. Seja $A \in \mathbb{C}^{m\times n}$. O conjunto imagem de A, denotado por $\mathcal{R}(A)$ é o conjunto

$$\mathcal{R}(A) = \{y \in \mathbb{C}^m : y = Ax, x \in \mathbb{C}^n\},$$

e o núcleo de A, denotado por $\mathcal{N}(A)$ é o conjunto

$$\mathcal{N}(A) = \{x \in \mathbb{C}^n : Ax = 0\}.$$

Definição 1.4. Se $A \in \mathbb{C}^{m\times n}$, então o posto de A, denotado por posto(A), é o número de colunas de A linearmente independentes, o qual coincide com o número de linhas de A linearmente independentes. Caso tenhamos posto$(A) = \min\{m, n\}$, dizemos que A é uma matriz de posto completo.

O próximo teorema mostra as propriedades básicas da inversa de Moore-Penrose a respeito dos conjuntos núcleo e imagem.

Teorema 1.5. *[29] Sejam $A \in \mathbb{C}^{m\times n}$ e $A^\dagger$ a inversa de Moore-Penrose de A. Temos que*

1. $\mathcal{R}(A) = \mathcal{R}(AA^\dagger) = \mathcal{R}(AA^*)$;

2. $\mathcal{R}(A^\dagger) = \mathcal{R}(A^*) = \mathcal{R}(A^\dagger A) = \mathcal{R}(A^*A)$;

3. $\mathcal{R}(I - A^\dagger A) = \mathcal{N}(A^\dagger A) = \mathcal{N}(A) = \mathcal{R}(A^*)^\perp$;

4. $\mathcal{R}(I - AA^\dagger) = \mathcal{N}(AA^\dagger) = \mathcal{N}(A^\dagger) = \mathcal{N}(A^*) = \mathcal{R}(A)^\perp$;

5. $\mathcal{R}(AB) = \mathcal{R}(A) \Leftrightarrow$ *posto*$(AB) =$ *posto*(A).

6. $\mathcal{N}(AB) = \mathcal{N}(B) \Leftrightarrow$ *posto*$(AB) =$ *posto*(B).

Demonstração:

1. Se $y \in \mathcal{R}(A)$, então $\exists\ x \in \mathbb{C}^n$ tal que $Ax = y$. Pela primeira equação de Penrose segue que $AA^\dagger Ax = y$, e tomando $z = Ax$ temos $AA^\dagger z = y$ portanto $y \in \mathcal{R}(AA^\dagger)$, e assim $\mathcal{R}(A) \subset \mathcal{R}(AA^\dagger)$.

   Agora, considere $y \in \mathcal{R}(AA^\dagger)$, então $\exists\ x \in \mathbb{C}^m$ tal que $AA^\dagger x = y$. Usando o item 5 do Teorema 1.2 temos que $AA^*(AA^*)^\dagger x = y$, assim se $z = (AA^*)^\dagger x$ temos que $AA^*z = y$, ou seja, $y \in \mathcal{R}(AA^*)$ de onde segue que $\mathcal{R}(AA^\dagger) \subset \mathcal{R}(AA^*)$.

   Por fim, tomando $y \in \mathcal{R}(AA^*)$, temos que $\exists\ x \in \mathbb{C}^m$ tal que $AA^*x = y$, e se considerarmos $z = A^*x$ temos $Az = y$, de onde segue que $y \in \mathcal{R}(A)$, e assim $\mathcal{R}(AA^*) \subset \mathcal{R}(A)$.

   Juntando as informações acima temos

   $$\mathcal{R}(A) \subset \mathcal{R}(AA^\dagger) \subset \mathcal{R}(AA^*) \subset \mathcal{R}(A),$$

   portanto
   $$\mathcal{R}(A) = \mathcal{R}(AA^\dagger) = \mathcal{R}(AA^*).$$

2. Se $y \in \mathcal{R}(A^\dagger)$, então $\exists\ x \in \mathbb{C}^m$ tal que $A^\dagger x = y$. Usando o item 5 do Teorema 1.2 temos $A^*(AA^*)^\dagger x = y$, e chamando $z = (AA^*)^\dagger x$ temos $A^*z = y$, portanto $y \in \mathcal{R}(A^*)$ e temos que $\mathcal{R}(A^\dagger) \subset \mathcal{R}(A^*)$.

   Agora considere $y \in \mathcal{R}(A^*)$, então existe $x \in \mathbb{C}^m$ tal que $A^*x = y$, e pelo item 6 do Teorema 1.2 temos que $A^\dagger AA^*x = y$, tomando $z = A^*x$ temos $A^\dagger Az = y$, portanto $y \in \mathcal{R}(A^\dagger A)$, e temos que $\mathcal{R}(A^*) \subset \mathcal{R}(A^\dagger A)$.

Considerando agora $y \in \mathcal{R}(A^\dagger A)$, temos que $\exists\ x \in \mathbb{C}^n$ tal que $A^\dagger Ax = y$, e usando os itens 5 e 6 do Teorema 1.2 segue que $A^*AA^\dagger(AA^*)^\dagger Ax = y$, se tomarmos $z = A^\dagger(AA^*)^\dagger Ax$ temos $A^*Az = y$, logo $y \in \mathcal{R}(A^*A)$ e portanto $\mathcal{R}(A^\dagger A) \subset \mathcal{R}(A^*A)$.

Por fim, se $y \in \mathcal{R}(A^*A)$, então $\exists\ x \in \mathbb{C}^n$ tal que $A^*Ax = y$. Pelo item 6 do Teorema 1.2 segue que $A^\dagger AA^*Ax = y$, e para $z = AA^*Ax$ temos $A^\dagger z = y$, logo $y \in \mathcal{R}(A^\dagger)$ e $\mathcal{R}(A^*A) \subset \mathcal{R}(A^\dagger)$.

Juntando as afirmações acima temos

$$\mathcal{R}(A^\dagger) \subset \mathcal{R}(A^*) \subset \mathcal{R}(A^\dagger A) \subset \mathcal{R}(A^*A) \subset \mathcal{R}(A^\dagger),$$

de onde implica que

$$\mathcal{R}(A^\dagger) = \mathcal{R}(A^*) = \mathcal{R}(A^\dagger A) = \mathcal{R}(A^*A).$$

3. Se $x \in \mathcal{R}(A^*)^\perp$, então $y^*x = 0,\ \forall\ y \in \mathcal{R}(A^*)$. Considere $w \in \mathbb{C}^m$ arbitrário, então $w^*Ax = (A^*w)^*x = 0$, pois $(A^*w) \in \mathcal{R}(A^*)$, como $w^*Ax = 0\ \forall\ w \in \mathbb{C}^m$ segue que $Ax = 0$, e portanto $x \in \mathcal{N}(A)$ e temos $\mathcal{R}(A^*)^\perp \subset \mathcal{N}(A)$.

   Agora, se $x \in \mathcal{N}(A)$, então $Ax = 0$, portanto $A^\dagger Ax = A^\dagger 0 = 0$, logo $x \in \mathcal{N}(A^\dagger A)$, e assim $\mathcal{N}(A) \subset \mathcal{N}(A^\dagger A)$.

   Considere $x \in \mathcal{N}(A^\dagger A)$, então $A^\dagger Ax = 0$ de onde temos que $x - A^\dagger Ax = x$, ou seja,$(I - A^\dagger A)x = x$ e portanto $x \in \mathcal{R}(I - A^\dagger A)$, de onde segue que $\mathcal{N}(A^\dagger A) \subset \mathcal{R}(I - A^\dagger A)$.

   E por fim, se $x \in \mathcal{R}(I - A^\dagger A)$, então existe $z \in \mathbb{C}^n$ tal que $z - A^\dagger Az = x$, tomando $w \in \mathcal{R}(A^*)$, temos que $\exists\ t \in \mathbb{C}^m$ tal que $A^*t = w$ e assim temos

$$w^*x = (A^*t)(z - A^\dagger Az) = t^*Az - t^*AA^\dagger Az = t^*Az - t^*Az = 0,$$

   de onde temos que $x \in \mathcal{R}(A^*)^\perp$ e portanto $\mathcal{R}(I - A^\dagger A) \subset \mathcal{R}(A^*)^\perp$.

   Juntando as afirmações acima temos

$$\mathcal{R}(I - A^\dagger A) \subset \mathcal{R}(A^*)^\perp \subset \mathcal{N}(A) \subset \mathcal{R}(A^\dagger A) \subset \mathcal{R}(I - A^\dagger A)$$

   e assim

$$\mathcal{R}(I - A^\dagger A) = \mathcal{R}(A^*)^\perp = \mathcal{N}(A) = \mathcal{R}(A^\dagger A).$$

4. Se $x \in \mathcal{R}(I - AA^\dagger)$ e $y \in \mathcal{R}(A)$, então $\exists\ z \in \mathbb{C}^m$ e $w \in \mathbb{C}^n$ tal que $z - AA^\dagger z = x$ e $Aw = y$, assim, usando o item 5 do Teorema 1.2 temos que

$$y^*x = (Aw)^*(z - AA^\dagger z) = w^*A^*z - w^*A^*AA^\dagger z = w^*A^*z - w^*A^*z = 0,$$

de onde temos que $x \in \mathcal{R}(A)^{\perp}$ e portanto $\mathcal{R}(I - AA^{\dagger}) \subset \mathcal{R}(A)^{\perp}$.

Agora considere $x \in \mathcal{R}(A)^{\perp}$ e $w \in \mathbb{C}^n$ arbitrário, então $w^*A^*x = (Aw)^*x = 0$, pois $Aw \in \mathcal{R}(A)$, e como $w^*A^*x = 0$ para todo $w \in \mathbb{C}^n$ segue que $A^*x = 0$, logo $x \in \mathcal{N}(A^*)$, e assim $\mathcal{R}(A)^{\perp} \subset \mathcal{N}(A^*)$.

Agora, se $x \in \mathcal{N}(A^*)$, então $A^*x = 0$, e pelo item 5 do teorema 1.2 segue que $A^{\dagger}x = (A^*A)^{\dagger}A^*x = (A^*A)^{\dagger}0 = 0$, logo $x \in \mathcal{N}(A^{\dagger})$ e temos que $\mathcal{N}(A^*) \subset \mathcal{N}(A^{\dagger})$.

Considere agora $x \in \mathcal{N}(A^{\dagger})$, então $A^{\dagger}x = 0$ e temos que $AA^{\dagger}x = A0 = 0$, e portanto $x \in \mathcal{N}(AA^{\dagger})$ e $\mathcal{N}(A^{\dagger}) \subset \mathcal{N}(AA^{\dagger})$.

Por fim, para $x \in \mathcal{N}(AA^{\dagger})$ temos que $AA^{\dagger}x = 0$, portanto $x - AA^{\dagger}x = x$, ou seja, $(I - AA^{\dagger})x = x$, de onde segue que $x \in \mathcal{R}(I - AA^{\dagger})$ e portanto $\mathcal{N}(AA^{\dagger}) \subset \mathcal{R}(I - AA^{\dagger})$.

Segue das afirmações acima que

$$\mathcal{R}(I - AA^{\dagger}) \subset \mathcal{R}(A)^{\perp} \subset \mathcal{N}(A^*) \subset \mathcal{N}(A^{\dagger}) \subset \mathcal{N}(AA^{\dagger}) \subset \mathcal{R}(I - AA^{\dagger}),$$

e assim

$$\mathcal{R}(I - AA^{\dagger}) = \mathcal{R}(A)^{\perp} = \mathcal{N}(A^*) = \mathcal{N}(A^{\dagger}) = \mathcal{N}(AA^{\dagger}).$$

5. $(\Rightarrow)$:

   Como a imagem de uma matriz é o conjunto gerado pelas colunas da matriz, então a dimensão da imagem coincide com o posto, portanto $\mathcal{R}(AB) = \mathcal{R}(A)$ implica em

   $$\text{posto}(AB) = \text{posto}(A).$$

   $(\Leftarrow)$:

   Considere $y \in \mathcal{R}(AB)$. Temos que existe x tal que $ABx = y$, de onde temos que $y \in \mathcal{R}(A)$, assim $\mathcal{R}(AB) \subset \mathcal{R}(A)$, e como $\text{posto}(AB) = \text{posto}(A)$ então $\dim(\mathcal{R}(AB)) = \dim(\mathcal{R}(A))$, de onde segue que

   $$\mathcal{R}(AB) = \mathcal{R}(A).$$

6. $(\Rightarrow)$:

   Como $\mathcal{N}(AB) = \mathcal{N}(B)$, segue do Teorema do núcleo e imagem que $\dim(\mathcal{R}(AB)) = \dim(\mathcal{R}(B))$, e portanto

   $$\text{posto}(AB) = \text{posto}(B).$$

$(\Leftarrow)$:

Seja $x \in \mathcal{N}(B)$. Temos que $Bx = 0$, e assim $ABx = 0$, ou seja, $x \in \mathcal{N}(AB)$, então temos que $\mathcal{N}(B) \subset \mathcal{N}(AB)$, e como posto$(AB)$ =posto(B) então $\dim(\mathcal{R}(AB)) = \dim(\mathcal{R}(B))$ e do Teorema do núcleo e imagem $\dim(\mathcal{N}(AB)) = \dim(\mathcal{N}(B))$, de onde segue que

$$\mathcal{N}(AB) = \mathcal{N}(B).$$

■

1.1.1 Algumas aplicações da inversa de Moore-Penrose

A aplicação mais natural da inversa de Moore-Penrose é em sistemas de equações lineares $Ax = b$ com $A \in \mathbb{C}^{m \times n}$. Quando pensamos em um sistema $Ax = b$ a solução x depende diretamente do conjunto imagem de A, ou seja, se $b \in \mathcal{R}(A)$ então sabemos que existe algum vetor x tal que $Ax = b$ e o sistema possui solução, porém ainda assim é possível que essa solução não seja única, nesse caso a solução do sistema é definida por

$$x = \arg \min_{Az=b} \|z\|$$

para alguma norma $\|\cdot\|$. Por questões geométricas a norma utilizada com maior frêquencia é a norma euclidiana $\|\cdot\|_2$, a qual será denotada daqui em diante apenas por $\|\cdot\|$.

No caso em que $b \notin \mathcal{R}(A)$, não há solução do sistema, então definimos como solução o vetor x que minimiza o erro em algum sentido. Para tanto, definimos o resíduo do sistema por $r = Ax - b$, e a solução do sistema é definida por

$$x = \arg \min_{z \in \mathbb{C}^n} \|Az - b\|.$$

Sabemos que quando A é quadrada e não singular, a solução do sistema é dada por $A^{-1}b$, e no caso da inversa de Moore-Penrose, será que podemos caracterizar a solução de maneira análoga? A é resposta é sim, como mostra o próximo teorema.

Teorema 1.6. *[29] Dado um sistema de equações lineares $Ax = b$, com $A \in \mathbb{C}^{m \times n}$, então $x = A^\dagger b$ é um minizador de $\|Ax - b\|$, e além disso é o minimizador de norma-2 mínima, ou seja, $A^\dagger b$ é a solução do sistema $Ax = b$.*

Demonstração:

Se $A = U\Sigma V^*$ é a decomposição SVD de A, então $A^\dagger = V\Sigma^\dagger U^*$. Seja r =posto(A). Temos que $\sigma_k = 0$ para $k > r$, assim se $\alpha = V^* x$ então

$$\begin{aligned}
\|Ax - b\|^2 &= \|U^*(Ax - b)\|^2 = \|U^*AVV^*x - U^*b\|^2 = \|(U^*AV)V^*x - U^*b\|^2 \\
&= \|\Sigma\alpha - U^*b\|^2 = \sum_{i=1}^{r} (\sigma_i \alpha_i - u_i^* b)^2 + \sum_{i=r+1}^{m} (u_i^* b)^2
\end{aligned}$$

como apenas a primeira parcela depende de x, temos que se x é um minimizador de $\|Ax - b\|$ então

$$\sigma_i \alpha_i - u_i^* b = 0, \text{ ou seja, } \alpha_i = \frac{u_i^* b}{\sigma_i}$$

e se tomarmos $\alpha_{r+1} = \ldots = \alpha_n = 0$ teremos o minimizador de norma-2 mínima dado por

$$x = \sum_{i=1}^{r} \frac{u_i^* b}{\sigma_i} v_i = V \Sigma^\dagger U^* b = A^\dagger b.$$

∎

Atualmente a solução de sistemas dessa forma é encontrada sem o cálculo explícito da inversa de Moore-Penrose, porém a inversa de Moore-Penrose aparece em diversos problemas de variadas áreas, por exemplo na engenharia elétrica o cálculo da matriz de impedância [1] Z de duas redes de resistores conectadas em paralelo, com matrizes de impedância Z_1 e Z_2, é dada por $Z = Z_1(Z_1 + Z_2)^\dagger Z_2$, mais detalhes podem ser encontrados em [5].

A inversa de Moore-Penrose é sem dúvida a principal inversa generalizada de matrizes, porém as equações de Penrose podem ser usadas para definir outras inversas generalizadas, as quais não têm garantia de unicidade, mas podem ser úteis em diversas aplicações.

1.2 Inversas-$\{i, j, k\}$

Dada uma matriz $A \in \mathbb{C}^{m \times n}$, dizemos que a matriz $X \in \mathbb{C}^{n \times m}$ é uma inversa-$\{i, j, k\}$ de A se X satisfaz a i-ésima, a j-ésima e a k-ésima equação de Penrose, por exemplo X é uma inversa-$\{1\}$ se satisfaz $AXA = A$, X é uma inversa-$\{1, 2\}$ se satisfaz $AXA = A$ e $XAX = X$, entre outros. Vejamos alguns exemplos interessantes de inversas-$\{i, j, k\}$.

1.2.1 Inversa-$\{1\}$ e suas aplicações

Dada uma matriz $A \in \mathbb{C}^{m \times n}$, $X \in \mathbb{C}^{n \times m}$ satisfazendo $AXA = A$ é uma inversa-$\{1\}$ de A. A existência de tal inversa segue da existência da inversa de Moore-Penrose, a qual também é uma inversa-$\{1\}$, porém não temos unicidade nesse caso, mas ainda assim a inversa-$\{1\}$ possui diversas aplicações interessantes.

[1] A impedância é a carga resistiva total de um circuito de corrente alternada, ou seja quando um determinado componente cria uma resistência e gasta energia em forma de calor, tem se o Efeito Joule, isso chamamos de resistência, e se o componente não gasta energia em forma de calor temos a reatância, então quando estão presentes a resistência e reatância chamamos de impedância. A impedância é expressa como um número complexo, possuindo uma parte real equivalente a resistência R, e uma parte imaginária dada pela reatância X. A impedância é designada pelo símbolo Z, e indica a oposição total que um circuito oferece ao fluxo de uma corrente elétrica variável no tempo.

Inversa-$\{1\}$ e o conjunto solução de $Ax = b$ com $b \in \mathcal{R}(A)$: Vimos que a inversa de Moore-Penrose pode ser usada para obter a solução de um sistema $Ax = b$ para $A \in \mathbb{C}^{m\times n}$ qualquer, porém no caso particular em que $A \in \mathcal{R}(A)$ sabemos que existe x satisfazendo a equação, porém nem sempre temos um único vetor x que satisfaz tal equação, mas podemos caracterizar o conjunto dos elementos que satizfazem $Ax = b$ usando inversas-$\{1\}$ de A.

Teorema 1.7. *[29] Para $A \in \mathbb{C}^{m\times n}$, $X \in \mathbb{C}^{m\times n}$ tem a propriedade de que Xb é uma solução de $Ax = b$ para todo $b \in \mathcal{R}(A)$ se e somente se $AXA = A$.*

Demonstração:

$(\Rightarrow)$: Se Xb é solução de $Ax = b$ para todo $b \in \mathcal{R}(A)$ então temos que $AXb = b$. Em particular, se $A = [a_1, a_2, \ldots, a_n]$, temos que $Ae_i = a_i$, e portanto $a_i \in \mathcal{R}(A)$ e então $AXA = AX[a_1, a_2, \ldots, a_n] = [AXa_1, AXa_2, \ldots, AXa_n] = [a_1, a_2, \ldots, a_n] = A$.

$(\Leftarrow)$: Suponha que $AXA = A$, então dado $b \in \mathcal{R}(A)$, $\exists\ x_b \in \mathbb{C}^n$ tal que $Ax_b = b$, sendo assim

$$AXb = AXAx_b = Ax_b = b.$$

■

Logo, o conjunto solução do sistema $Ax = b$, com $A \in \mathbb{C}^{m\times n}$ e $b \in \mathcal{R}(A)$ é dado por

$$\{Xb : X \in \mathbb{C}^{n\times m} \text{ e } AXA = A\}.$$

Inversa-$\{1\}$ e a solução da equação matricial $AXB = D$:

Teorema 1.8. *[29] Se $A \in \mathbb{C}^{m\times n}$, $B \in \mathbb{C}^{p\times q}$ e $D \in \mathbb{C}^{m\times q}$, então a equação matricial*

$$AXB = D \tag{1.5}$$

é consistente se e somente se para $A^{(1)}$ e $B^{(1)}$ inversas-$\{1\}$ de A e B respectivamente vale

$$AA^{(1)}DB^{(1)}B = D.$$

Nesse caso, a solução geral da equação matricial (1.5) é dada por

$$X = A^{(1)}DB^{(1)} + Y - A^{(1)}AYBB^{(1)} \tag{1.6}$$

para $Y \in \mathbb{C}^{n\times p}$ arbitrário.

Demonstração:

Supondo que vale $AA^{(1)}DB^{(1)}B = D$, claramente temos que $X = A^{(1)}DB^{(1)}$ é uma solução da equação (1.5), ou seja, a equação é consistente. Agora supondo que a equação

(1.5) seja consistente, então existe X tal que $AXB = D$, e usando a definição de inversas-$\{1\}$ para A e B temos

$$D = AXB = AA^{(1)}AXBB^{(1)}B = AA^{(1)}DB^{(1)}B.$$

Além disso, para uma equação consistente temos que se X é da forma (1.6) então

$$\begin{aligned} AXB &= A(A^{(1)}DB^{(1)} + Y - A^{(1)}AYBB^{(1)})B \\ &= AA^{(1)}DB^{(1)}B + AYB - AA^{(1)}AYBB^{(1)}B \\ &= AA^{(1)}DB^{(1)}B + AYB - AYB = AA^{(1)}DB^{(1)}B = D \end{aligned}$$

assim, X é solução da equação (1.5).

Por outro lado, se X é uma solução de (1.5), então

$$X = A^{(1)}DB^{(1)} + X - A^{(1)}DB^{(1)} = A^{(1)}DB^{(1)} + X - A^{(1)}AXBB^{(1)}$$

que é da forma (1.6). ■

Um caso particular do teorema acima nos dá uma caracterização para consistência de um sistema linear, para tanto basta tomar $B = I$ e $D = b$, de onde segue o seguinte corolário.

Corolário 1.9. *Se $A \in \mathbb{C}^{m\times n}$ e $b \in \mathbb{C}^m$, então o sistema de equações lineares $Ax = b$ é consistente se, e somente se para alguma $A^{(1)}$ inversa-$\{1\}$ de A, tenhamos*

$$AA^{(1)}b = b,$$

nesse caso, a solução geral do sistema é dada por

$$x = A^{(1)}b + (I - A^{(1)}A)y$$

para $y \in \mathbb{C}^n$ arbitrário.

Inversa-$\{1\}$ e a solução comum aos sistemas $Ax = a$ e $Bx = b$: Se $A \in \mathbb{C}^{m\times n}$, $B \in \mathbb{C}^{p\times n}$, $a \in \mathbb{C}^m$ e $b \in \mathbb{C}^p$, então podemos considerar o seguinte problema: encontrar $x \in \mathbb{C}^n$ tal que sejam satisfeitas as equações

$$Ax = a \quad \text{e} \quad Bx = b.$$

Tal problema é equivalente a resolver o seguinte sistema particionado

$$\begin{pmatrix} A \\ B \end{pmatrix} x = \begin{pmatrix} a \\ b \end{pmatrix}.$$

Os dois teoremas a seguir nos dão uma solução do problema acima em termos de inversas-$\{1\}$.

Por $(A \mid B)$ denotamos a matriz cujas primeiras colunas correspondem as colunas de A e as últimas colunas correspondem as colunas de B.

Teorema 1.10. *[29] Sejam $A \in \mathbb{C}^{m\times n}$, $B \in \mathbb{C}^{p\times n}$ e $C \in \mathbb{C}^{m\times r}$. Se $Y^{(1)}$ denota uma inversa-$\{1\}$ de Y, então*

(i) Uma inversa-$\{1\}$ da matriz particionada $M = \begin{pmatrix} A \\ B \end{pmatrix}$ é dada por

$$X_M = \left((I - D(BD)^{(1)}B)A^{(1)} \mid D(BD)^{(1)}\right),$$

onde $D = I - A^{(1)}A$.

(ii) Uma inversa-$\{1\}$ da matriz particionada $N = \begin{pmatrix} A \mid C \end{pmatrix}$ é dada por

$$X_N = \begin{pmatrix} A^{(1)}(I - C(EC)^{(1)}E) \\ (EC)^{(1)}E \end{pmatrix},$$

onde $E = I - AA^{(1)}$.

Demonstração:

(i) Note que $AD = A(I - A^{(1)}A) = A - AA^{(1)}A = A - A = 0$, assim temos que

$$\begin{aligned}
MX_MM &= \begin{pmatrix} A \\ B \end{pmatrix}\left((I - D(BD)^{(1)}B)A^{(1)} \mid D(BD)^{(1)}\right)\begin{pmatrix} A \\ B \end{pmatrix} \\
&= \begin{pmatrix} AA^{(1)}A - AD(BD)^{(1)}BA^{(1)}A + AD(BD)^{(1)}B \\ BA^{(1)}A - BD(BD)^{(1)}BA^{(1)}A + BD(BD)^{(1)}B \end{pmatrix} \\
&= \begin{pmatrix} A \\ BA^{(1)}A + BD(BD)^{(1)}BD \end{pmatrix} = \begin{pmatrix} A \\ BA^{(1)}A + BD \end{pmatrix} \\
&= \begin{pmatrix} A \\ B(A^{(1)}A + D) \end{pmatrix} = \begin{pmatrix} A \\ B(A^{(1)}A + I - A^{(1)}A) \end{pmatrix} = \begin{pmatrix} A \\ B \end{pmatrix} = M
\end{aligned}$$

(ii) Note que $EA = (I - AA^{(1)})A = A - AA^{(1)}A = A - A = 0$, de onde segue que

$$\begin{aligned}
NX_NN &= (A \mid C)\begin{pmatrix} A^{(1)}(I - C(EC)^{(1)}E) \\ (EC)^{(1)}E \end{pmatrix}(A \mid C) \\
&= (AA^{(1)}(I - C(EC)^{(1)}E + C(EC)^{(1)}E)(A \mid C) \\
&= ((AA^{(1)}(I - C(EC)^{(1)}E)A + C(EC)^{(1)}EA \mid AA^{(1)}(I - C(EC)^{(1)}E)C \\
&\quad +C(EC)^{(1)}EC) \\
&= (AA^{(1)}A - AA^{(1)}C(EC)^{(1)}EA + C(EC)^{(1)}EA \mid AA^{(1)}C \\
&\quad -AA^{(1)}C(EC)^{(1)}EC + C(EC)^{(1)}EC) \\
&= (A \mid AA^{(1)}C + EC(EC)^{(1)}EC) \\
&= (A \mid AA^{(1)}C + EC) = (A \mid (AA^{(1)} + E)C) \\
&= (A \mid (AA^{(1)} + I - AA^{(1)})C) = (A \mid C) = N
\end{aligned}$$

∎

O próximo teorema nos fornece condições para a existência de uma solução comum aos sistemas $Ax = a$ e $Bx = b$, e nesse caso descreve o conjunto de todas as soluções comuns.

Teorema 1.11. *[29] Sejam $A \in \mathbb{C}^{m\times n}$, $B \in \mathbb{C}^{p\times n}$, $a \in \mathcal{R}(A)$ e $b \in \mathcal{R}(B)$. Considere x_a e x_b soluções particulares dos sistemas $Ax = a$ e $Bx = b$ respectivamente. Sejam $D = I - A^{(1)}A$ e $F = BD$. Temos que as seguintes afirmações são equivalentes:*

(i) Os sistemas $Ax = a$ e $Bx = b$ possuem uma solução comum.

(ii) $Bx_a - b \in \mathcal{R}(F) = B\mathcal{N}(A)$.

(iii) $x_a - x_b \in \mathcal{N}(A) + \mathcal{N}(B)$.

Além disso, caso exista uma solução comum aos dois sistemas, então uma solução é dada por

$$x_c = (I - DF^{(1)}B)x_a + DF^{(1)}b, \tag{1.7}$$

e o conjunto de todas as soluções comuns dos dois sistemas é dado por

$$\{x_c + D(I - F^{(1)}F)h : h \in \mathbb{C}^n\}. \tag{1.8}$$

Demonstração:

$(iii) \Rightarrow (ii)$

Supondo que $x_a - x_b \in \mathcal{N}(A) + \mathcal{N}(B)$ então podemos escrever $x_a - x_b = n_a + n_b$ tal que $An_a = 0$ e $Bn_b = 0$, de onde segue que

$$\begin{aligned} Bx_a - b &= Bx_a - Bx_b = B(x_a - x_b) = B(n_a + n_b) = Bn_a \\ &= Bn_a + A^{(1)}An_a = B(I - A^{(1)}A)n_a = Fn_a, \end{aligned}$$

portanto $Bx_a - b \in \mathcal{R}(F)$, e como

$$\mathcal{R}(F) = \{y \in \mathbb{C}^n; Fx = y\} = \{y \in \mathbb{C}^n; BDx = y\} = B\mathcal{R}(D),$$

e pelo item 3 do Teorema 1.5 temos que $\mathcal{R}(D) = \mathcal{N}(A)$, de onde segue que $Bx_a - b \in \mathcal{R}(F) = B\mathcal{N}(A)$.

$(ii) \Rightarrow (i)$

Supondo que $Bx_a - b \in \mathcal{R}(F)$, então existe $y \in \mathbb{C}^n$ tal que $Bx_a - b = Fy$, e como $AD = 0$, temos que o vetor x_c definido por (1.7) satisfaz

$$Ax_c = A((I - DF^{(1)}B)x_a + DF^{(1)}b) = Ax_a - ADF^{(1)}Bx_a + ADF^{(1)}b = Ax_a = a$$

e

$$\begin{aligned} Bx_c &= B((I - DF^{(1)}B)x_a + DF^{(1)}b) = Bx_a - BDF^{(1)}Bx_a + BDF^{(1)}b \\ &= Bx_a - FF^{(1)}Bx_a + FF^{(1)}b = Bx_a - FF^{(1)}(Bx_a - b) \\ &= Bx_a - FF^{(1)}Fy = Bx_a - Fy = Bx_a - (Bx_a - b) = b, \end{aligned}$$

ou seja, x_c é uma solução comum a $Ax = a$ e $Bx = b$.

$(i) \Rightarrow (iii)$

Supondo que existe uma solução comum a $Ax = a$ e $Bx = b$, então os conjuntos solução se interceptam. Note que dado um sistema $Yx = y$ e uma solução particular desse sitema x_y, então o conjunto solução do sistema é dado por $x_y + \mathcal{N}(Y)$. De fato, se x é uma solução do sistema $Yx = y$ então podemos escrever $x = x_y + (x - x_y)$ e $Y(x - x_y) = Yx - Yx_y = y - y = 0$. Reciprocamente, se $x \in x_y + \mathcal{N}(Y)$ então $x = x_y + z$ com $Yz = 0$, assim $Yx = Yx_y + Yz = Yx_y = y$.

Sendo assim, temos que

$$\{x_a + \mathcal{N}(A)\} \cap \{x_b + \mathcal{N}(B) \neq \emptyset\},$$

de onde implica que $\exists\ n_a \in \mathcal{N}(A)$ e $n_b \in \mathcal{N}(B)$ tal que $x_a + n_a = x_b + n_b$, ou seja $x_a - x_b = -n_a + n_b$ e portanto

$$x_a - x_b \in \mathcal{N}(A) + \mathcal{N}(B).$$

Como

$$(iii) \Rightarrow (ii) \Rightarrow (i) \Rightarrow (iii)$$

então

$$(i) \equiv (ii) \equiv (iii).$$

Agora, para ver que o conjunto de soluções em comum é dado por (1.8), considere o problema equivalente

$$\begin{pmatrix} A \\ B \end{pmatrix} x = \begin{pmatrix} a \\ b \end{pmatrix},$$

cujo conjunto solução é dado por

$$x_c + \mathcal{N} \begin{pmatrix} A \\ B \end{pmatrix},$$

e usando o item 3 do Teorema 1.5 e a inversa-$\{1\}$ de $\begin{pmatrix} A \\ B \end{pmatrix}$ dada por X_m no Teorema 1.10, obtemos

$$\begin{aligned} x_c + \mathcal{N} \begin{pmatrix} A \\ B \end{pmatrix} &= \left\{ x_c + \left(I - \begin{pmatrix} A \\ B \end{pmatrix}^{(1)} \begin{pmatrix} A \\ B \end{pmatrix} \right) h : h \in \mathbb{C}^n \right\} \\ &= \left\{ x_c + \left(I - X_m \begin{pmatrix} A \\ B \end{pmatrix} \right) h : h \in \mathbb{C}^n \right\} \end{aligned}$$

e como

$$\begin{aligned} I - X_M \begin{pmatrix} A \\ B \end{pmatrix} &= I - (I - DF^{(1)}B)A^{(1)} \mid DF^{(1)}) \begin{pmatrix} A \\ B \end{pmatrix} \\ &= I - (I - DF^{(1)}B)A^{(1)}A - DF^{(1)}B \\ &= I - A^{(1)}A + DF^{(1)}BA^{(1)}A - DF^{(1)}B \\ &= D + DF^{(1)}BA^{(1)}A - DF^{(1)}B \\ &= D(I + F^{(1)}BA^{(1)}A - F^{(1)}B) \\ &= D(I - F^{(1)}B(I - A^{(1)}A)) \\ &= D(I - F^{(1)}F), \end{aligned}$$

temos que

$$x_c + \mathcal{N} \begin{pmatrix} A \\ B \end{pmatrix} = \{x_c + D(I - F^{(1)}F)h : h \in \mathbb{C}^n\}.$$

$\blacksquare$

Inversa-$\{1\}$ e a solução comum as equações matriciais $AX = B$ e $XD = E$: Assim como podemos obter uma solução geral em comum para dois sistemas de equações lineares usando inversas-$\{1\}$, também podemos caracterizar uma solução em comum das equações matriciais $AX = B$ e $XD = E$, como mostra o teorema seguinte.

Teorema 1.12. *[29] Sejam $A \in \mathbb{C}^{m\times n}$, $B \in \mathbb{C}^{m\times p}$, $D \in \mathbb{C}^{p\times q}$ e $E \in \mathbb{C}^{n\times q}$. Temos que as equações matriciais*

$$AX = B \quad e \quad XD = E$$

possuem uma solução em comum se e somente se cada equação possui solução e $AE = BD$.

Neste caso, uma solução em comum é dada por

$$X_c = A^{(1)}B + ED^{(1)} - A^{(1)}AED^{(1)},$$

e o conjunto de todas as soluções é dado por

$$\{X_c + (I - A^{(1)}A)Y(I - DD^{(1)}) : Y \in \mathbb{C}^{n\times p}\}. \tag{1.9}$$

Demonstração:

$(\Rightarrow)$:

Suponha que exista uma solução em comum X, então $AX = B$ e $XD = E$. Multiplicando a primeira equação por D à direita e a segunda por A à esquerda temos

$$AXD = BD \quad \text{e} \quad AXD = AE,$$

de onde segue que $AE = BD$.

$(\Leftarrow)$:

Como cada equação é consistente, ou seja possui solução, então pelo Teorema 1.8, podemos escrever a equação $AX = B$ como $AXI = B$ e $XD = E$ como $IXD = E$, onde I denota a matriz identidade. Assim temos que

$$AA^{(1)}BII^{(1)} = B \quad \text{e} \quad II^{(1)}ED^{(1)}D = E$$

portanto $AA^{(1)} = B$ e $D^{(1)}D = E$, e usando que $AE = BD$ temos

$$\begin{aligned} AX_c &= A(A^{(1)}B + ED^{(1)} - A^{(1)}AED^{(1)}) = AA^{(1)}B + AED^{(1)} - AA^{(1)}AED^{(1)} \\ &= B + AED^{(1)} - AED^{(1)} = B \end{aligned}$$

e

$$\begin{aligned} X_cD &= (A^{(1)}B + ED^{(1)} - A^{(1)}AED^{(1)})D = A^{(1)}BD + ED^{(1)}D - A^{(1)}AED^{(1)}D \\ &= A^{(1)}BD + E - A^{(1)}AE = A^{(1)}BD + E - A^{(1)}BD = E, \end{aligned}$$

logo X_c é uma solução comum a $AX = B$ e $XD = E$.

Para ver que o conjunto de soluções em comum é dado por (1.9), note que dado X da forma $X_c + (I - A^{(1)}A)Y(I - DD^{(1)})$ temos

$$\begin{aligned} AX &= A(X_c + (I - A^{(1)}A)Y(I - DD^{(1)})) = AX_c + A(I - A^{(1)}A)Y(I - DD^{(1)}) \\ &= B + (A - AA^{(1)}A)Y(I - DD^{(1)}) = B \end{aligned}$$

e

$$\begin{aligned} XD &= (X_c + (I - A^{(1)}A)Y(I - DD^{(1)}))D = X_cD + (I - A^{(1)}A)Y(I - DD^{(1)})D \\ &= E + (I - A^{(1)}A)Y(D - DD^{(1)}D) = E, \end{aligned}$$

portanto X é uma solução comum as duas equações. Reciprocamente, se X satisfaz as duas equações, então

- $A^{(1)}A(X - X_c) = A^{(1)}AX - A^{(1)}AX_c = A^{(1)}B - A^{(1)}B = 0$;
- $(X - X_c)DD^{(1)} = XDD^{(1)} - X_cDD^{(1)} = ED^{(1)} - ED^{(1)} = 0$;
- $A^{(1)}A(X - X_c)DD^{(1)} = [A^{(1)}A(X - X_c)]DD^{(1)} = 0$;

de onde segue que

$$\begin{aligned} X &= X_c + X - X_c \\ &= X_c + (X - X_c) - (X - X_c)DD^{(1)} - A^{(1)}A(X - X_c) + A^{(1)}A(X - X_c)DD^{(1)} \\ &= X_c + (X - X_c)(I - DD^{(1)}) - A^{(1)}A(X - X_c)(I - DD^{(1)}) \\ &= X_c + (I - A^{(1)}A)(X - X_c)(I - DD^{(1)}) = X_c + (I - A^{(1)}A)Y(I - DD^{(1)}), \end{aligned}$$

com $Y = X - X_c$. ■

1.2.2 Inversa-$\{1, 4\}$ e a solução de norma mínima de sistemas de equações lineares consistentes

Vimos anteriomente que para um sistema $Ax = b$, com $A \in \mathbb{C}^{m \times n}$ e $b \in \mathcal{R}(A)$, a solução pode não ser única, e neste caso definimos como solução o vetor x de norma mínima que satisfaz $Ax = b$. Também vimos que $A^{\dagger}b$ é o vetor de norma-2 mínima que satisfaz o sistema, porém a maneira de obter essa solução não é única. O próximo teorema mostra que podemos obter essa solução por Xb, onde X é uma matriz que satisfaz as equações (1.1) e (1.4).

Teorema 1.13. *[29] Se $A \in \mathbb{C}^{m\times n}$ e $b \in \mathbb{C}^m$, tal que $b \in \mathcal{R}(A)$, então para $X \in \mathbb{C}^{n\times m}$ temos que $AXb = b$ e $\|Xb\| < \|z\|$ para todo $z \neq Xb$, $z \in \{x : Ax = b\}$ se e somente se X satisfaz*

$$AXA = A \quad e \quad (XA)^* = XA.$$

Além disso, se X satisfaz as duas equações acima, então o conjunto de soluções de $Ax = b$ é dado por

$$\{Xb + (I - XA)h : h \in \mathbb{C}^n\}. \tag{1.10}$$

Demonstração:

$(\Leftarrow)$:

Do fato de X satisfazer $AXA = A$ e $b \in \mathcal{R}(A)$ temos que

$$AXb = AXAx = Ax = b,$$

portanto Xb satisfaz o sistema. Para mostrar que Xb é a solução de norma mínima, vejamos primeiro que o conjunto solução de $Ax = b$ é dado por (1.10).

Seja x solução de $Ax = b$. Temos que

$$x = Xb + x - Xb = Xb + x - XAx = Xb + (I - XA)x,$$

ou seja, x pertence a (1.10). Reciprocamente, tomando um vetor da forma $Xb+(I-XA)h$ para $h \in \mathbb{C}^n$ arbitrário temos

$$A(Xb + (I - XA)h) = AXb + A(I - XA)h = b + (A - AXA)h = b + (A - A)h = b,$$

o que prova que o conjunto solução de $Ax = b$ é dado por (1.10).

Vejamos agora que Xb tem norma-2 mínima. Segue de $(XA)^* = XA$ que

$$Xb = XAx = (XA)^*x = A^*X^*x,$$

portanto $Xb \in \mathcal{R}(A^*)$ e pelo item 3 do Teorema 1.5 temos que

$$\langle Xb, (I + XA)h\rangle = 0,$$

seguindo assim do Teorema de Pitágoras que

$$\|Xb\|^2 \leq \|Xb\|^2 + \|(I - XA)h\|^2 = \|Xb + (I - XA)h\|^2,$$

portanto $\|Xb\| < \|z\|$ para todo $z \neq Xb$, $z \in \{Xb + (I - XA)h : h \in \mathbb{C}^n\}$.

($\Rightarrow$) :

Seja Y uma matriz que satisfaz $AYA = A$ e $(YA)^* = YA$. Vimos acima que Yb é a solução de norma mínima do sistema $Ax = b$, e como por hipótese Xb também é solução de norma mínima de $Ax = b$ segue que

$$Xb = Yb, \ \forall \ b \in \mathcal{R}(A),$$

se $A = [a_1, a_2, \ldots, a_n]$, claramente $a_i \in \mathcal{R}(A)$ para $i = 1, 2, \ldots, n$, de onde temos que $X[a_1, a_2, \ldots, a_n] = Y[a_1, a_2, \ldots, a_n]$, ou seja $XA = YA$, assim

$$AXA = AYA = A \quad \text{ e } \quad (XA)^* = (YA)^* = YA = XA.$$

■

1.2.3 Inversa-$\{1, 3\}$ e a solução de sistemas de equações lineares inconsistentes

Como vimos anteriormente, no caso de um sistema $Ax = b$ com $A \in \mathbb{C}^{m \times n}$ em que $b \notin \mathcal{R}(A)$, não existe nenhum vetor x que satisfaça $Ax = b$, então a solução do problema é tomada como sendo o vetor x que minimiza a norma do resíduo $Ax - b$.

Vimos também que a inversa de Moore-Penrose pode ser usada para obter tal solução, ou seja, o vetor que minimiza a norma do resíduo é dado por $A^\dagger b$. Porém, assim como no caso em que b pertence a imagem de A, podemos caracterizar a solução desse problema em termos de uma inversa mais fraca que a inversa de Moore-Penrose. O teorema a seguir mostra que a solução do problema pode ser dada por uma inversa-$\{1, 3\}$.

Teorema 1.14. *[29] Se $A \in \mathbb{C}^{m \times n}$ e $b \in \mathbb{C}^m$ tal que $b \notin \mathcal{R}(A)$, então Xb é o minimizador de $\|Ax - b\|$ se e somente se X satisfaz*

$$AXA = A \quad e \quad (AX)^* = AX.$$

Demonstração:

($\Leftarrow$) :

Seja $b = b_1 + b_2$ de tal forma que $b_1 \in \mathcal{R}(A)$ e $b_2 \in \mathcal{R}(A)^\perp$. Temos que b_1 pode ser escrito como $b_1 = Ay$ para algum $y \in C^n$ e pelo item 4 do Teorema 1.5 temos $\mathcal{R}(A)^\perp = \mathcal{N}(A^*)$, assim $A^* b_2 = 0$, portanto

$$AXb_1 = AXAy = Ay = b_1 \quad \text{ e } \quad AXb_2 = (AX)^* b_2 = X^* A^* b_2 = 0,$$

de onde segue que

$$AXb = AX(b_1 + b_2) = AXb_1 + AXb_2 = b_1. \tag{1.11}$$

Do Teorema de Pitágoras, temos que

$$\|AXb-b\|^2 = \|AXb-(b_1+b_2)\|^2 = \|AXb-b_1+(-b_2)\|^2 = \|AXb-b_1\|^2+\|b_2\|^2,$$

como a segunda parcela não depende de X, temos que a norma de $AXb-b$ será mínima se $AXb=b_1$, o que segue de (1.11).

$(\Rightarrow)$:

Seja Y uma matriz que satisfaça $AYA=A$ e $(AY)^*=AY$. Vimos acima que Xb minimiza $\|Ax-b\|$ se e somente se $AXb=b_1$, e como Yb é um minimizador pela demonstração acima, temos que $AYb=b_1$. Assim $AXb=AYb$, $\forall\ b\in\mathbb{C}^m$, de onde segue que $AX=AY$, implicando em

$$AXA = AYA = A \quad \text{e} \quad (AX)^* = (AY)^* = AY = AX.$$

■

1.3 Inversas generalizadas com núcleo e imagem prescritos

1.3.1 Projetores e matrizes idempotentes

Definição 1.15. Uma matriz $A\in\mathbb{C}^{n\times n}$ é chamada de idempotente se satisfaz

$$A^2 = A.$$

O lema a seguir nos fornece algumas propriedades de matrizes idempotentes.

Lema 1.16. *[29] Se $A\in\mathbb{C}^{n\times n}$ é uma matriz idempotente, então as seguintes afirmações são válidas.*

1. *A^* e $I-A$ são idempotentes;*
2. *Os autovalores de A são 0 e 1, e a multiplicidade do autovalor 1 é posto(A);*
3. *posto(A) =tr(A);*
4. *$A(I-A)=(I-A)A=0$;*
5. *$Ax=x$ se e somente se $x\in\mathcal{R}(A)$;*
6. *A é uma inversa-$\{1,2\}$ de A;*
7. *$\mathcal{N}(A)=\mathcal{R}(I-A)$.*

Demonstração:

1. Segue de $(A^*)^2 = (A^2)^* = A^*$.

2. Se $\lambda \in \mathbb{C}$ é um autovalor de A, então $Ax = \lambda x$ para algum $x \neq 0$, multiplicando por A a esquerda temos
$$Ax = A^2x = \lambda Ax = \lambda^2 x,$$
ou seja, se λ é autovalor de A então λ^2 também é, mas isso só é possível se $\lambda^2 = \lambda$, de onde segue que $\lambda \in \{0, 1\}$.

3. Como o traço de matrizes semelhantes é igual, segue da forma de Jordan de A que posto(A) =tr(A).

4. Segue de $A^2 = A$ que
$$A(I - A) = A - A^2 = A - A = 0, \quad \text{e} \quad (I - A)A = A - A^2 = A - A = 0.$$

5. Se $Ax = x$ segue diretamente que $x \in \mathcal{R}(A)$. Reciprocamente, se $x \in \mathcal{R}(A)$, então existe y tal que $Ay = x$, assim
$$Ax = AAy = A^2y = Ay = x.$$

6. Segue do fato de que $AAA = A^2A = AA = A^2 = A$.

7. Se $x \in \mathcal{R}(I - A)$, então existe y tal que $(I - A)y = x$ e assim
$$Ax = A(I - A)y = (A - A^2)y = (A - A)y = 0.$$
Portanto $\mathcal{R}(I - A) \subset \mathcal{N}(A)$, além disso, os autovalores de $I - A$ são da forma $1 - \lambda$, onde λ é autovalor de A, portanto os autovalores de $I - A$ são 1 e 0, e o autovalor 1 de $I - A$ possui a mesma multiplicidade do autovalor 0 de A, ou seja
$$\dim(\mathcal{R}(I - A)) = \dim(\mathcal{N}(A)),$$
de onde segue que $\mathcal{N}(A) = \mathcal{R}(I - A)$.

■

Lema 1.17. *[29] Seja $A \in \mathbb{C}^{n\times n}$ com posto$(A) = k$, se $A = FG$, com $F \in \mathbb{C}^{n\times k}$ e $G \in \mathbb{C}^{k\times n}$, então A é idempotente se e somente se $GF = I$.*

Demonstração:

Se $GF = I$, claramente temos

$$A^2 = (FG)^2 = FGFG = FIG = FG = A.$$

Por outro lado, tomando $F^{(1)}$ e $G^{(1)}$ inversas-$\{1\}$ de F e G respectivamente, como posto(F) = posto$(G) = k$ segue que F^*F e GG^* são não singulares, o que implica em

$$\begin{aligned} FF^{(1)}F &= F \\ F^*FF^{(1)}F &= F^*F \\ F^{(1)}F &= (F^*F)^{-1}F^*F \\ F^{(1)}F &= I \end{aligned}$$

e

$$\begin{aligned} GG^{(1)}G &= G \\ GG^{(1)}GG^* &= GG^* \\ GG^{(1)} &= GG^*(GG^*)^{-1} \\ GG^{(1)} &= I. \end{aligned}$$

Assim, multiplicando $FGFG = FG$ a esquerda por $F^{(1)}$ e a direita por $G^{(1)}$ obtemos

$$\begin{aligned} FGFG &= FG \\ F^{(1)}FGFGG^{(1)} &= F^{(1)}FGG^{(1)} \\ GF &= I. \end{aligned}$$

■

Dados dois subespaços de $\mathbb{C}^n$ complementares L e M, ou seja, $\mathbb{C}^n = L \oplus M$, então temos que $x \in \mathbb{C}^n$ pode ser expresso de maneira única

$$x = y + z \quad (y \in L,\ z \in M).$$

Neste caso, o vetor y é chamado de projeção de x sobre L ao longo de M.

Definição 1.18. Denotaremos por $P_{L,M}$ a transformação linear que leva x na sua projeção sobre L ao longo de M. Tal transformação é unicamente representada por uma matriz, uma vez fixada a base de $\mathbb{C}^n$. E por $P_{L,M}$ denotaremos tanto a transformação linear quanto a sua representação matricial, chamada de projetor sobre L ao longo de M. Neste caso $P_{L,M}x = y$.

O próximo teorema estabelece uma relação biunívoca entre uma matriz idempotente de ordem n e um projetor $P_{L,M}$ onde $L \oplus M = \mathbb{C}^n$. Além disso, na demonstração do teorema identificamos uma maneira de construir $P_{L,M}$, conhecendo bases de L e M.

Teorema 1.19. *[29] Para toda matriz idempotente $A \in \mathbb{C}^{n\times n}$, $\mathcal{R}(A)$ e $\mathcal{N}(A)$ são subespaços complementares de $\mathbb{C}^n$, com*

$$A = P_{\mathcal{R}(A),\mathcal{N}(A)}.$$

Reciprocamente, se L e M são subespaços complementares de $\mathbb{C}^n$, existe uma única matriz idempotente $P_{L,M} \in \mathbb{C}^{n\times n}$ tal que

$$\mathcal{R}(P_{L,M}) = L \quad e \quad \mathcal{N}(P_{L,M}) = M.$$

Demonstração:

Se $A \in \mathbb{C}^{n\times n}$ é uma matriz idempotente, então podemos escrever $x \in \mathbb{C}^n$ da forma

$$x = Ax + (I - A)x,$$

e do Lema 1.16 segue que $\mathbb{C}^n$ é soma de $\mathcal{R}(A)$ e $\mathcal{N}(A)$. Para mostrar que $\mathcal{R}(A) \cap \mathcal{N}(A) = \{0\}$, note que do fato de $x \in \mathcal{R}(A)$ segue que $Ax = x$ pelo item 5 do Lema 1.16, e como $x \in \mathcal{N}(A)$ então $Ax = 0$, portanto $x = 0$. Assim, $\mathbb{C}^n = \mathcal{R}(A) \oplus \mathcal{N}(A)$, e como $x = Ax + (I - A)x$ para todo $x \in \mathbb{C}^n$, então Ax é a projeção de x sobre $\mathcal{R}(A)$ ao longo de $\mathcal{N}(A)$.

Reciprocamente, sejam L e M subespaços complementares de $\mathbb{C}^n$ com bases dadas respectivamente por $\{x_1, x_2, \ldots, x_l\}$ e $\{y_1, y_2, \ldots, y_m\}$. Sabemos que um projetor sobre L ao longo de M satisfaz

$$\begin{aligned} P_{L,M}x_i &= x_i, \quad i = 1, 2, \ldots, l, \\ P_{L,M}y_i &= 0, \quad i = 1, 2, \ldots, m. \end{aligned}$$

Vamos construir uma matriz idempotente $P_{L,M}$ que satisfaça as condições acima, assim tal matriz será uma projeção sobre L ao longo de M. Se $X = (x_1, x_2, \ldots, x_l)$ e $Y = (y_1, y_2, \ldots, y_m)$, então $P_{L,M}$ deve satisfazer

$$P_{L,M}(X, Y) = (X, 0),$$

e como (X, Y) é não singular, pois $\{x_1, x_2, \ldots, x_l, y_1, y_2, \ldots, y_m\}$ é uma base de $\mathbb{C}^n$, então defina

$$P_{L,M} = (X, 0)(X, Y)^{-1}.$$

Vejamos que $P_{L,M}$ é idempotente. De fato, como $P_{L,M}(X,Y) = (X,0)$, ou seja, $P_{L,M}X = X$ e $P_{L,M}Y = 0$ temos

$$P_{L,M}^2 = P_{L,M}(X,0)(X,Y)^{-1} = (X,0)(X,Y)^{-1} = P_{L,M}.$$

Basta mostrar que $P_{L,M}$ é única, para isso suponha que existe outra matriz idempotente E tal que $\mathcal{R}(E) = L$ e $\mathcal{N}(E) = M$, então

$$\begin{aligned} Ex_i &= x_i, \quad i = 1, 2, \ldots, l, \\ Ey_i &= 0, \quad i = 1, 2, \ldots, m, \end{aligned}$$

de onde segue que $E(X,Y) = (X,0)$ e portanto $E = (X,0)(X,Y)^{-1} = P_{L,M}$. ■

Corolário 1.20. *Sejam $A \in \mathbb{C}^{m\times n}$ e $X \in \mathbb{C}^{n\times m}$ inversas-$\{1,2\}$ uma da outra. Temos que AX é o projetor sobre $\mathcal{R}(A)$ ao longo de $\mathcal{N}(X)$, e XA é o projetor sobre $\mathcal{R}(X)$ ao longo de $\mathcal{N}(A)$, ou seja,*

$$AX = P_{\mathcal{R}(A),\mathcal{N}(X)}, \qquad e \qquad XA = P_{\mathcal{R}(X),\mathcal{N}(A)}.$$

Demonstração:

Como A e X são inversas-$\{1,2\}$ uma da outra, então

$$AXA = A \quad \text{e} \quad XAX = X.$$

Assim, se $y \in \mathcal{R}(A)$ e $w \in \mathcal{N}(X)$, então existe $z \in \mathbb{C}^n$ tal que $Az = y$, e portanto

$$\begin{aligned} &AXy = AXAz = Az = y; \\ &AXw = A0 = 0; \\ &(AX)^2 = AXAX = AX. \end{aligned}$$

E segue do Teorema 1.19 que $AX = P_{\mathcal{R}(A),\mathcal{N}(X)}$.

Da mesma forma, se $y \in \mathcal{R}(X)$ e $w \in \mathcal{N}(A)$, então existe $z \in \mathbb{C}^n$ tal que $Xz = y$, e portanto

$$\begin{aligned} &XAy = XAXz = Xz = y; \\ &XAw = X0 = 0; \\ &(XA)^2 = XAXA = XA. \end{aligned}$$

E do Teorema 1.19 $XA = P_{\mathcal{R}(X),\mathcal{N}(A)}$. ■

O próximo corolário mostra uma correspondência biunívoca entre projetores ortogonais e matrizes idempotentes Hermitianas.

Corolário 1.21. *Seja $\mathbb{C}^n = L \oplus M$. Temos que $M = L^\perp$ se e somente se $P_{L,M}$ é Hermitiana.*

Demonstração:

Pelo Lema 1.16, $P^*_{L,M}$ é idempotente, e segue do Teorema 1.19 que $P^*_{L,M}$ é um projetor. Como para uma matriz qualquer A, $\mathcal{N}(A^*) = \mathcal{R}(A)^\perp$, então

$$\mathcal{R}(P^*_{L,M}) = \mathcal{N}(P_{L,M})^\perp = M^\perp \quad \text{e} \quad \mathcal{N}(P^*_{L,M}) = \mathcal{R}(P_{L,M})^\perp = L^\perp,$$

portanto $P^*_{L,M} = P_{M^\perp,L^\perp}$. Assim, se $P^*_{L,M} = P_{L,M}$, temos

$$M = \mathcal{N}(P_{L,M}) = \mathcal{N}(P^*_{L,M}) = \mathcal{N}(P_{M^\perp,L^\perp}) = L^\perp.$$

Reciprocamente, se $M = L^\perp$, então

$$P^*_{L,M} = P_{M^\perp,L^\perp} = P_{L,M}.$$

■

Denotaremos por P_L o projetor ortogonal sobre L ao longo de $L^\perp$, ou seja, $P_{L,L^\perp}$. Os próximos resultados mostram condições para que a soma, diferença e produto de projetores sejam também projetores.

Teorema 1.22. *[29] Sejam P_1 o projetor sobre R_1 ao longo de N_1 e P_2 o projetor sobre R_2 ao longo de N_2. Temos que $P = P_1 + P_2$ é um projetor se e somente se*

$$P_1P_2 = P_2P_1 = 0.$$

Neste caso, P é o projetor sobre $R = R_1 \oplus R_2$ ao longo de $N = N_1 \cap N_2$.

Demonstração:

$(\Rightarrow)$:

Se $P_1P_2 = P_2P_1 = 0$, então

$$P^2 = (P_1 + P_2)^2 = P_1^2 + P_1P_2 + P_2P_1 + P_2^2 = P_1^2 + P_2^2 = P_1 + P_2 = P,$$

e do Teorema 1.19 segue que P é um projetor sobre o espaço R ao longo de N.

Vejamos que $R = R_1 \oplus R_2$. De fato, primeiramente vejamos que $R_1 \cap R_2 = \{0\}$, para isso considere $u \in R_1 \cap R_2$, então $P_1u = u$ e $P_2u = u$ de onde temos

$$u = P_2u = P_2P_1u = 0,$$

logo $R_1 \cap R_2 = \{0\}$. Considerando $u \in R_1 \oplus R_2$, temos que $u = u_1 + u_2$, com $u_1 \in R_1$ e

$u_2 \in R_2$, além disso temos

$$P_1u_1 = u_1 \quad \text{e} \quad P_2u_2 = u_2,$$

o que implica em

$$P_2u_1 = P_2P_1u_1 = 0 \quad \text{e} \quad P_1u_2 = P_1P_2u_2 = 0.$$

Assim

$$Pu = (P_1 + P_2)(u_1 + u_2) = P_1u_1 + P_1u_2 + P_2u_1 + P_2u_2 = u_1 + u_2 = u,$$

e portanto $R_1 \oplus R_2 \subset R$. Por outro lado, tomando $u \in R$ temos

$$u = Pu = (P_1 + P_2)u = P_1u + P_2u, \quad (P_1u \in R_1 \;\text{ e }\; P_2u \in R_2),$$

de onde temos que $R \subset R_1 \oplus R_2$, e portanto $R = R_1 \oplus R_2$.

Vejamos agora que $N = N_1 \cap N_2$. De fato, se $u \in N$ então $Pu = 0$ e temos que

$$0 = P_1Pu = P_1(P_1 + P_2)u = P_1^2u + P_1P_2u = P_1u,$$

logo $u \in N_1$. Da mesma forma

$$0 = P_2Pu = P_2(P_1 + P_2)u = P_2P_1u + P_2^2u = P_2u,$$

e temos que $u \in N_2$, assim $N \subset N_1 \cap N_2$. Por outro lado, se $u \in N_1 \cap N_2$, então $P_1u = 0$ e $P_2u = 0$, assim

$$Pu = (P_1 + P_2)u = P_1u + P_2u = 0,$$

e portanto $u \in N$, assim $N = N_1 \cap N_2$.

$(\Leftarrow)$:

Como $P_1^2 = P_1$ e $P_2^2 = P_2$, se $P^2 = P$ então

$$P_1 + P_2 = P = P^2 = (P_1 + P_2)^2 = P_1^2 + P_2^2 + P_1P_2 + P_2P_1 = P_1 + P_2 + P_1P_2 + P_2P_1,$$

de onde segue que

$$P_1P_2 + P_2P_1 = 0. \tag{1.12}$$

Multiplicando (1.12) a esquerda por P_1 obtemos

$$\begin{aligned} P_1P_2 + P_2P_1 &= 0 \\ P_1^2P_2 + P_1P_2P_1 &= 0 \\ P_1P_2 + P_1P_2P_1 &= 0. \end{aligned} \tag{1.13}$$

Multiplicando (1.12) a direita por P_1 obtemos

$$\begin{aligned} P_1P_2 + P_2P_1 &= 0 \\ P_1P_2P_1 + P_2P_1^2 &= 0 \\ P_1P_2P_1 + P_2P_1 &= 0. \end{aligned} \tag{1.14}$$

E por fim, multiplicando (1.13) a direita por P_1 obtemos

$$\begin{aligned} P_1P_2 + P_1P_2P_1 &= 0 \\ P_1P_2P_1 + P_1P_2P_1^2 &= 0 \\ P_1P_2P_1 + P_1P_2P_1 &= 0 \\ 2P_1P_2P_1 &= 0, \end{aligned}$$

de onde concluímos que $P_1P_2P_1 = 0$, e substituindo em (1.13) e (1.14) teremos

$$P_1P_2 = P_2P_1 = 0.$$

■

Teorema 1.23. *[29] Sob as hipóteses do Teorema 1.22, $P = P_1 - P_2$ é um projetor se e somente se*

$$P_1P_2 = P_2P_1 = P_2.$$

Neste caso, P é um projetor sobre $R = R_1 \cap R_2$ ao longo de $N = N_1 \oplus R_2$.

Demonstração:

Note que P é um projetor se e somente se $(I - P)$ é um projetor, e como

$$(I - P) = (I - P_1) + P_2,$$

segue do Teorema 1.22 que P é um projetor se e somente se

$$(I - P_1)P_2 = P_2(I - P_1) = 0,$$

o que equivale a $P_1P_2 = P_2P_1 = P_2$. Quanto a R e N, segue do Lema 1.16 e do Teorema

1.22 que

$$N = \mathcal{N}(P) = \mathcal{R}(I-P) = \mathcal{R}(I-P_1) \oplus \mathcal{R}(P_2) = \mathcal{N}(P_1) \oplus \mathcal{R}(P_2) = N_1 \oplus R_2$$

e

$$R = \mathcal{R}(P) = \mathcal{N}(I-P) = \mathcal{N}(I-P_1) \cap \mathcal{N}(P_2) = \mathcal{R}(P_1) \cap \mathcal{N}(P_2) = R_1 \cap N_2.$$

■

Teorema 1.24. *[29] Sob as hipóteses do Teorema 1.22, se*

$$P_1P_2 = P_2P_1,$$

Então $P = P_1P_2$ é um projetor sobre $R = R_1 \cap R_2$ ao longo de $N = N_1 + N_2$.

Demonstração:
Como $P_1P_2 = P_2P_1$, então

$$P^2 = (P_1P_2)^2 = P_1P_2P_1P_2 = P_1P_1P_2P_2 = P_1^2P_2^2 = P_1P_2 = P.$$

Vejamos que $R = R_1 \cap R_2$. Se $u \in R$, então $Pu = u$ e assim

$$P_1u = P_1Pu = P_1P_1P_2u = P_1P_2u = Pu = u$$

e

$$P_2u = P_2Pu = P_2P_1P_2u = P_1P_2P_2u = P_1P_2u = Pu = u,$$

portanto $u \in R_1 \cap R_2$ e $R \subset R_1 \cap R_2$. Por outro lado, tomando $u \in R_1 \cap R_2$ temos

$$P_1u = u \quad \text{e} \quad P_2u = u,$$

de onde temos que

$$Pu = P_1P_2u = P_1u = u,$$

logo $u \in R$ e $R = R_1 \cap R_2$.

Vejamos que $N = N_1 + N_2$. De fato, tomando $u \in N$ temos que

$$P_1P_2u = Pu = 0,$$

de onde segue que $P_2u \in N_1$. Além disso, como $P_2(I-P_2)u = 0$, temos que $(I-P_2)u \in N_2$, e como $u = P_2u+(I-P_2)u$, temos que $N \subset N_1+N_2$. Por outro lado, tomando $u \in N_1+N_2$ podemos escrever u da forma $u = u_1 + u_2$, $u_1 \in N_1$, $u_2 \in N_2$, então

$$Pu = P_1P_2(u_1 + u_2) = P_1P_2u_1 + P_1P_2u_2 = P_1P_2u_1 = P_2P_1u_1 = 0,$$

e portanto $u \in N$ e $N = N_1 + N_2$. ∎

1.3.2 Inversa generalizada $A^{(1,2)}_{T,S}$

Seja $A \in \mathbb{C}^{m\times n}$, e seja $A^{(1)}$ uma inversa-$\{1\}$ arbitrária de A. Segue do fato de que $AA^{(1)}A = A$ que

$$(AA^{(1)})^2 = AA^{(1)}AA^{(1)} = AA^{(1)} \quad \text{e} \quad (A^{(1)}A)^2 = A^{(1)}AA^{(1)}A = A^{(1)}A,$$

ou seja, $AA^{(1)}$ e $A^{(1)}A$ são idempotentes, e portanto projetores. Sejam $\mathcal{R}(A) = L$ e $\mathcal{N}(A) = M$, considere S e T subespaços de $\mathbb{C}^m$ e $\mathbb{C}^n$ respectivamente, tais que

$$L \oplus S = \mathbb{C}^m \quad \text{e} \quad T \oplus M = \mathbb{C}^n,$$

neste caso, segue do Teorema 1.19 que

$$AA^{(1)} = P_{L,S} \quad \text{e} \quad A^{(1)}A = P_{T,M}.$$

Chamamos de inversa-$\{1,2\}$ TS, denotada por $A^{(1,2)}_{T,S}$, a única matriz X que satisfaz as equações

$$AX = P_{L,S}; \tag{1.15}$$
$$XA = P_{T,M}; \tag{1.16}$$
$$XAX = X. \tag{1.17}$$

Os próximos resultados discutem a questão de existência e unicidade de tal inversa.

Lema 1.25. *[29] Existe no máximo uma matriz X satisfazendo as equações*

$$AX = B, \quad XA = D, \quad XAX = X. \tag{1.18}$$

Demonstração:

Suponha que as equações 1.18 tenham mais que uma solução, digamos que X_1 e X_2 sejam soluções das equações 1.18, defina $U = X_1 - X_2$, então

$$\begin{aligned}
AU &= A(X_1 - X_2) = AX_1 - AX_2 = B - B = 0;\\
UA &= (X_1 - X_2)A = X_1A - X_2A = D - D = 0;\\
UB &= (X_1 - X_2)B = X_1B - X_2B = X_1AX_1 - X_2AX_2 = X_1 - X_2 = U;\\
DU &= D(X_1 - X_2) = DX_1 - DX_2 = X_1AX_1 - X_2AX_2 = X_1 - X_2 = U.
\end{aligned}$$

Assim, para $i = 1, 2$, temos

$$U^*U = (DU)^*(UB) = U^*D^*UB = U^*(X_iA)^*U(AX_i) = U^*A^*X_i^*UAX_i = 0.$$

Portanto, $U = 0$, de onde segue que $X_1 = X_2$. ■

Teorema 1.26. *[29] Se $A \in \mathbb{C}^{m\times n}$, $\mathcal{R}(A) = L$, $\mathcal{N}(A) = M$, $L \oplus S = \mathbb{C}^m$ e $T \oplus M = \mathbb{C}^n$, então*

1. *X é uma inversa-$\{1\}$ de A tal que $\mathcal{R}(XA) = T$ e $\mathcal{N}(AX) = S$ se e somente se*

$$AX = P_{L,S} \quad \text{e} \quad XA = P_{T,M}. \tag{1.19}$$

2. *A solução geral de (1.19) é dada por*

$$X = P_{T,M}A^{(1)}P_{L,S} + (I_n - A^{(1)}A)Y(I_m - AA^{(1)}),$$

onde $A^{(1)}$ denota uma inversa-$\{1\}$ de A fixa, porém arbitrária, e $Y \in \mathbb{C}^{n\times m}$ arbitrária.

3. *$P_{T,M}A^{(1)}P_{L,S} = A^{(1,2)}_{T,S}$, além disso, essa é a única inversa-$\{1,2\}$ que possui imagem T e núcleo S.*

Demonstração:

1. $(\Rightarrow)$:

Por hipótese temos $AXA = A$, e vimos que isso implica que AX e XA são idempotentes, pelo Teorema 1.19 temos que

$$AX = P_{\mathcal{R}(AX),\mathcal{N}(AX)} = P_{L,S} \quad \text{e} \quad XA = P_{\mathcal{R}(XA),\mathcal{N}(XA)} = P_{T,M}.$$

$(\Leftarrow)$:

Como $\mathcal{R}(A) \subset L$ então $P_{L,S}A = A$. De fato, seja $A = (a_1, a_2, \ldots, a_n)$, claramente $a_i \in \mathcal{R}(A) \subset L$, assim

$$P_{L,S}(a_1, a_2, \ldots, a_n) = (a_1, a_2, \ldots, a_n),$$

ou seja, $P_{L,S}A = A$. Porém $P_{L,S} = AX$, portanto

$$AXA = A,$$

e X é uma inversa-$\{1\}$ de A. Além disso, temos

$$\mathcal{N}(AX) = \mathcal{N}(P_{L,S}) = S \quad \text{e} \quad \mathcal{R}(XA) = \mathcal{R}(P_{T,M}) = T.$$

2. Defina $X_0 = P_{T,M}A^{(1)}P_{L,S}$, onde $A^{(1)}$ é uma inversa-$\{1\}$ de A. Vejamos que X_0 é solução comum as equações

$$AX = P_{L,S} \quad \text{e} \quad XA = P_{T,M}. \tag{1.20}$$

Para tanto, note que

$$\mathcal{R}(P_{L,S}) = L = \mathcal{R}(A) \quad \text{e} \quad \mathcal{R}(P^*_{T,M}) = \mathcal{R}(P_{M^\perp,T^\perp}) = M^\perp = \mathcal{N}(A)^\perp = \mathcal{R}(A^*),$$

de onde segue que existem Z e W tais que

$$P_{L,S} = AZ \quad \text{e} \quad P_{M^\perp,T^\perp} = A^*W \quad \equiv \quad P_{T,M} = W^*A. \tag{1.21}$$

Além disso, da demonstração do item anterior vimos que $\mathcal{R}(A) \subset L$ implica em

$$P_{L,S}A = A. \tag{1.22}$$

Vejamos também que $M \subset \mathcal{N}(A)$ implica em $AP_{T,M} = A$. De fato, pois se $M \subset \mathcal{N}(A)$, então $\mathcal{N}(A)^\perp \subset M^\perp$, e escrevendo $A^* = (\widehat{a}_1, \widehat{a}_2, \ldots, \widehat{a}_m)$, temos que $\widehat{a}_i \in \mathcal{N}(A)^\perp \subset M^\perp$, portanto

$$P_{M^\perp,T^\perp}A^* = P_{M^\perp,T^\perp}(\widehat{a}_1, \widehat{a}_2, \ldots, \widehat{a}_m) = (\widehat{a}_1, \widehat{a}_2, \ldots, \widehat{a}_m) = A^*,$$

ou seja,

$$AP_{T,M} = A. \tag{1.23}$$

Assim, de (1.21) e (1.23) temos

$$AX_0 = AP_{T,M}A^{(1)}P_{L,S} = AA^{(1)}AZ = AZ = P_{L,S},$$

e de (1.21) e (1.22) temos

$$X_0A = P_{T,M}A^{(1)}P_{L,S}A = W^*AA^{(1)}A = W^*A = P_{T,M}.$$

Portanto, X_0 é uma solução em comum de (1.20), e segue do Teorema 1.12 que a solução geral é

$$\begin{aligned} X &= X_0 + (I_n - A^{(1)}A)Y(I_m - AA^{(1)}) \\ &= P_{T,M}A^{(1)}P_{L,S} + (I_n - A^{(1)}A)Y(I_m - AA^{(1)}). \end{aligned}$$

3. Seja $X = P_{T,M}A^{(1)}P_{L,S}$ para $A^{(1)}$ alguma inversa-$\{1\}$ de A. Do item 2 temos que

$$AXA = P_{L,S}A = A,$$

logo X é uma inversa-$\{1\}$ de A, e da equação $AXA = A$ temos posto$(A) \leq$posto(X), e como

$$\text{posto}(X) = \text{posto}(P_{T,M}A^{(1)}P_{L,S}) \leq \text{posto}(P_{L,S}) = \text{posto}(A),$$

segue que posto(A) =posto(X). Além disso,

$$\text{posto}(A) = \text{posto}(AXA) \leq \text{posto}(XA) \quad \text{e} \quad \text{posto}(XA) \leq \text{posto}(A),$$

de onde temos que posto(XA) = posto(A) = posto(X), e como $\mathcal{R}(XA) \subset \mathcal{R}(X)$, segue que $\mathcal{R}(XA) = \mathcal{R}(X)$, portanto existe Y tal que $XAY = X$, de onde temos

$$\begin{aligned} XAY &= X \\ AXAY &= AX \\ AY &= AX \\ XAY &= XAX \\ X &= XAX, \end{aligned}$$

logo X é uma inversa-$\{1, 2\}$ de A. Pelo Corolário 1.20, temos que

$$AX = P_{\mathcal{R}(A),\mathcal{N}(X)} \quad \text{e} \quad XA = P_{\mathcal{R}(X),\mathcal{N}(A)},$$

e pelos itens 1 e 2 segue que $\mathcal{R}(X) = \mathcal{R}(XA) = T$ e $\mathcal{N}(X) = \mathcal{N}(AX) = S$, e como X satisfaz

$$AX = P_{L,S}, \quad XA = P_{T,M}, \quad XAX = X,$$

pelo Lema 1.25 segue que X é a única matriz satisfazendo as três equações acima.

■

Corolário 1.27. *Se $A \in \mathbb{C}^{m\times n}$, $\mathcal{R}(A) = L$ e $\mathcal{N}(A) = M$, então*

$$A^\dagger = A^{(1,2)}_{\mathcal{R}(A^*),\mathcal{N}(A^*)}.$$

Demonstração:

Como $A^\dagger$ é uma inversa-$\{1, 2\}$ com $\mathcal{R}(A^\dagger) = \mathcal{R}(A^*)$ e $\mathcal{N}(A^\dagger) = \mathcal{N}(A^*)$, segue da unicidade mostrada no teorema anterior que $A^\dagger = A^{(1,2)}_{\mathcal{R}(A^*),\mathcal{N}(A^*)}$. ■

1.3.3 Inversa generalizada $A^{(2)}_{T,S}$

Lema 1.28. *Sejam $A \in \mathbb{C}^{m\times n}$ e $B \in \mathbb{C}^{n\times p}$. Se*

$$posto(AB) = posto(A) = posto(B),$$

então

$$AB(AB)^{(1)}A = A \quad e \quad B(AB)^{(1)}AB = B,$$

onde $(AB)^{(1)}$ denota uma inversa-$\{1\}$ de AB.

Demonstração:

Como posto(AB) =posto(A) =posto(B), segue que

$$\dim(\mathcal{R}(AB)) = \dim(\mathcal{R}(A)) \quad \text{e} \quad \dim(\mathcal{N}(AB)) = \dim(\mathcal{N}(B)),$$

e como

$$\mathcal{R}(AB) \subset \mathcal{R}(A) \quad \text{e} \quad \mathcal{N}(B) \subset \mathcal{N}(AB)$$

segue que

$$\mathcal{R}(AB) = \mathcal{R}(A) \quad \text{e} \quad \mathcal{N}(AB) = \mathcal{N}(B).$$

Vimos anteriormente que

$$\mathcal{R}(A) \subset L \Rightarrow P_{L,M}A = A \quad \text{e} \quad M \subset \mathcal{N}(B) \Rightarrow BP_{L,M} = B,$$

e como

$$AB(AB)^{(1)} = P_{\mathcal{R}(AB),\mathcal{N}((AB)^{(1)})} \quad \text{e} \quad (AB)^{(1)}AB = P_{\mathcal{R}((AB)^{(1)}),\mathcal{N}(AB)}$$

então

$$AB(AB)^{(1)}A = A \quad \text{e} \quad B(AB)^{(1)}AB = B.$$

■

Teorema 1.29. *[29] Se $A \in \mathbb{C}^{m\times n}$ com posto$(A) = r$, T um subespaço de $\mathbb{C}^n$ de dimensão $t \leq r$ e S um subespaço de $\mathbb{C}^m$ de dimensão $m - t$, então A tem uma inversa-$\{2\}$ X tal que*

$$\mathcal{R}(X) = T \quad e \quad \mathcal{N}(X) = S$$

se e somente se

$$AT \oplus S = \mathbb{C}^m,$$

neste caso, X é única, e é denotada por $A^{(2)}_{T,S}$.

Demonstração:

$(\Rightarrow)$:

Como $XAX = X$ então A é uma inversa-$\{1\}$ de X e portanto AX é idempotente e pelo Teorema 1.19 temos que

$$\mathcal{R}(AX) \oplus \mathcal{N}(AX) = \mathbb{C}^m.$$

Note que $\mathcal{N}(AX) = \mathcal{N}(X)$. De fato, claramente $\mathcal{N}(X) \subset \mathcal{N}(AX)$, e por outro lado, se $x \in \mathcal{N}(AX)$ então $AXx = 0$, multiplicando por X a esquerda obtemos $Xx = XAXx = 0$. Assim

$$\mathcal{R}(AX) = \{y \in \mathbb{C}^m; AXx = y\} = A\{y \in \mathbb{C}^m; Xx = y\} = A\mathcal{R}(X) = AT,$$

portanto $AT \oplus S = \mathbb{C}^m$.

$(\Leftarrow)$:

Sejam $U \in \mathbb{C}^{n\times t}$ e $V^* \in \mathbb{C}^{m\times t}$ matrizes cujas colunas formam uma base para T e $S^\perp$ respectivamente, assim $\mathcal{R}(U) = T$ e $\mathcal{N}(V) = S$. Dessa maneira, as colunas de AU geram AT, cuja dimensão é t, portanto AU tem posto coluna completo, ou seja, posto$(AU) = t$.

Vejamos que posto$(VAU) = t$. De fato, se $VAUy = 0$ então $AUy \in \mathcal{N}(V)$ e $AUy \in \mathcal{R}(AU) = AT$, e segue de $AT \oplus S = \mathbb{C}^m$ que $AUy = 0$, mas como AU tem posto coluna completo, então $y = 0$, portanto posto$(VAU) = t$, de onde temos

$$\text{posto}(VAU) = \text{posto}(U) = \text{posto}(V) = t \leq r.$$

Defina $X = U(VAU)^{(1)}V = U(VAU)^{-1}V$, então pelo Lema 1.28 temos que

$$XAU = U(VAU)^{(1)}VAU = U \quad \text{e} \quad VAX = VAU(VAU)^{(1)}V = V,$$

logo

$$XAX = XAU(VAU)^{(1)}V = U(VAU)^{(1)}V = X,$$

assim X é uma inversa-$\{2\}$ de A, além disso

$$\begin{aligned} XAU = U &\Rightarrow \mathcal{R}(U) \subset \mathcal{R}(X); \\ X = U(VAU)^{(1)}V &\Rightarrow \mathcal{R}(X) \subset \mathcal{R}(U); \\ VAX = V &\Rightarrow \mathcal{N}(X) \subset \mathcal{N}(V); \\ X = U(VAU)^{(1)}V &\Rightarrow \mathcal{N}(V) \subset \mathcal{N}(X); \end{aligned}$$

de onde segue que

$$\mathcal{R}(X) = \mathcal{R}(U) = T \quad \text{e} \quad \mathcal{N}(X) = \mathcal{N}(V) = S.$$

Unicidade: Sejam X_1 e X_2 inversas-$\{2\}$ de A com imagem T e núcleo S. Temos que A é uma inversa-$\{1\}$ de X_1, e A também é inversa-$\{1\}$ de X_2, de onde segue que

$$\begin{aligned} X_1A &= P_{\mathcal{R}(X_1A),\mathcal{N}(X_1A)} = P_{\mathcal{R}(X_1),\mathcal{N}(X_1A)} = P_{T,\mathcal{N}(X_1A)}; \\ AX_2 &= P_{\mathcal{R}(AX_2),\mathcal{N}(AX_2)} = P_{\mathcal{R}(AX_2),\mathcal{N}(X_2)} = P_{\mathcal{R}(AX_2),S}; \end{aligned}$$

Vimos que $\mathcal{R}(X_2) = T$ e $\mathcal{N}(X_1) = S$ implica em

$$P_{T,\mathcal{N}(X_1A)}X_2 = X_2 \quad \text{e} \quad X_1P_{\mathcal{R}(AX_2),S} = X_1,$$

de onde segue que

$$X_2 = P_{T,\mathcal{N}(X_1A)}X_2 = X_1AX_2 = X_1P_{\mathcal{R}(AX_2),S} = X_1.$$

■

Corolário 1.30. *Sejam $A \in \mathbb{C}^{m\times n}$ com posto$(A) = r$, T um subespaço de $\mathbb{C}^n$ de dimensão r e S um subespaço de $\mathbb{C}^m$ de dimensão $m - r$. Temos que*

$$AT \oplus S = C^m$$

se e somente se existe X inversa-$\{1, 2\}$ de A tal que $\mathcal{R}(X) = T$ e $\mathcal{N}(X) = S$.

Demonstração:

$(\Rightarrow)$:

Supondo que $AT \oplus S = C^m$, pelo teorema anterior temos que

$$\text{posto}(VAU) = \text{posto}(U) = \text{posto}(V) = r,$$

logo

$$r = \text{posto}(VAU) \leq \text{posto}(AU) \leq \text{posto}(A) = r,$$

portanto posto(VAU) =posto(AU) =posto(A), e como $\mathcal{R}(AU) \subset \mathcal{R}(A)$ então

$$\mathcal{R}(AU) = \mathcal{R}(A),$$

portanto existe Y tal que $AUY = A$.

Como posto(VAU) =posto(AU), pelo Lema 1.28 temos que

$$AU(VAU)^{(1)}VAU = AU.$$

Definindo $X = U(VAU)^{(1)}V$ como na demonstração do teorema anterior, segue que

$$AXA = AU(VAU)^{(1)}VAUY = AUY = A,$$

portanto X é uma inversa-$\{1\}$ de A, juntando com o resultado do teorema anterior segue que X é uma inversa-$\{1,2\}$ de A, com $\mathcal{R}(X) = T$ e $\mathcal{N}(X) = S$.

$(\Leftarrow)$:

Segue diretamente do teorema anterior. ■

O corolário acima mostra que no caso em que $t = r$ no teorema 1.29, temos $A^{(1,2)}_{T,S} = A^{(2)}_{T,S}$, a principal implicação desse resultado é que no caso $T = \mathcal{R}(A^*)$ temos $t = r$, e portanto

$$A^\dagger = A^{(2)}_{\mathcal{R}(A^*),\mathcal{N}(A^*)}.$$

1.4 Inversa de Moore-Penrose com peso

Nas seções anteriores vimos que a inversa de Moore-Penrose, a inversa-$\{1,4\}$ e a inversa-$\{1,3\}$ possuem uma estreita relação com a solução de sistemas lineares da forma $Ax = b$, o qual considera as soluções usando a norma-2 (euclidiana). Uma pergunta que surge naturalmente é em relação a essa norma, se usarmos outra norma na solução de $Ax = b$, existe uma matriz X tal que a solução do sistema é dado por Xb?

Vejamos que no caso do produto interno com peso $\langle \cdot, \cdot \rangle_M$ é possível encontrar tal matriz X.

1.4.1 Produto interno e norma com peso

Sejam $M \in \mathbb{C}^{m\times m}$ e $N \in \mathbb{C}^{n\times n}$ matrizes hermitianas e positivas definidas. Definimos os produtos internos com peso em $\mathbb{C}^m$ e $\mathbb{C}^n$ respectivamente por

$$\langle x, y\rangle_M = x^*My, \quad \text{para } x, y \in \mathbb{C}^m$$

e

$$\langle x, y\rangle_N = x^*Ny, \quad \text{para } x, y \in \mathbb{C}^n.$$

A partir desses produtos internos, definimos as normas vetoriais com peso em $\mathbb{C}^m$ e $\mathbb{C}^n$ respectivamente por

$$\|x\|_M = \langle x, x\rangle_M^{\frac{1}{2}} = (x^*Mx)^{\frac{1}{2}} = \|M^{\frac{1}{2}}x\|_2, \quad x \in \mathbb{C}^m$$

e

$$\|x\|_N = \langle x, x\rangle_N^{\frac{1}{2}} = (x^* N x)^{\frac{1}{2}} = \|N^{\frac{1}{2}} x\|_2, \quad x \in \mathbb{C}^n.$$

Dizemos que $x, y \in \mathbb{C}^m$ são M-ortogonais se $\langle x, y\rangle_M = 0$, da mesma forma, dizemos que $x, y \in \mathbb{C}^n$ são N-ortogonais se $\langle x, y\rangle_N = 0$, em ambos os casos é válido o Teorema de Pitágoras.

A partir das normas vetoriais, podemos induzir a norma matricial com peso, obtendo a seguinte norma matricial

$$\|A\|_{MN} = \max_{\|x\|_N = 1} \|Ax\|_M, \quad A \in \mathbb{C}^{m\times n}$$

e

$$\|B\|_{NM} = \max_{\|x\|_M = 1} \|Bx\|_N, \quad B \in \mathbb{C}^{n\times m}.$$

Definição 1.31. Se $A \in \mathbb{C}^{m\times n}$, $M \in \mathbb{C}^{m\times m}$ e $N \in \mathbb{C}^{n\times n}$, com M e N matrizes hermitianas e positivas definidas, então a matriz $X \in \mathbb{C}^{n\times m}$ satisfazendo

$$\langle Ax, y\rangle_M = \langle x, Xy\rangle_N, \quad \forall\ x \in \mathbb{C}^n, y \in \mathbb{C}^m$$

é chamada matriz adjunta com peso, e denotada por $X = A^\#$.

Segue da definição que $A^\# = N^{-1}A^*M$. Para o caso especial em que $A \in \mathbb{C}^{n\times n}$ se

$$\langle Ax, y\rangle_N = \langle x, Ay\rangle_N$$

então $A^\# = A$, neste caso dizemos que A é uma matriz hermitiana com peso.

O lema a seguir mostra algumas propriedades da adjunta com peso de uma matriz.

Lema 1.32. *[29] As seguintes propriedades são válidas:*

1. $(A + B)^\# = A^\# + B^\#, \quad A, B \in \mathbb{C}^{m\times n}$;
2. $(AB)^\# = B^\# A^\#, \quad A \in \mathbb{C}^{m\times n}, B \in \mathbb{C}^{n\times m}$;
3. $(A^\#)^\# = A, \quad A \in \mathbb{C}^{m\times n}$;
4. $(A^\#)^{-1} = (A^{-1})^\#, \quad A \in \mathbb{C}^{n\times n}$ *não singular;*
5. $\|A\|_{MN} = \|A^\#\|_{NM}$;
6. $\|A\|_{MN}^2 = \|AA^\#\|_{MM} = \|A^\# A\|_{NN}$.
7. *Se L é um subespaço de $\mathbb{C}^n$, e $\mathcal{R}(B) = L$, então*

$$\begin{aligned} \|x\|_N &\leq \|x + y\|_N\ \forall\ y \in L \\ &\Leftrightarrow\ x^* N y = 0,\ \forall\ y \in L \\ &\Leftrightarrow\ x^* N B = 0. \end{aligned}$$

Demonstração:

Todos os resultados seguem da definição de $A^{\#}$. ■

Agora, podemos generalizar as equações de Penrose para o caso com pesos.

Definição 1.33. Se $A \in \mathbb{C}^{m\times n}$, $M \in \mathbb{C}^{m\times m}$ e $N \in \mathbb{C}^{n\times n}$ com M e N hermitianas e positivas definidas, então a matriz X satisfazendo as equações

$$AXA = A \tag{1.24}$$
$$XAX = X \tag{1.25}$$
$$(MAX)^* = MAX \tag{1.26}$$
$$(NXA)^* = NXA. \tag{1.27}$$

é chamada de inversa de Moore-Penrose com peso, e denotada por $A^{\dagger}_{MN}$.

Da mesma forma, podemos definir inversas-$\{i,j,k\}$ com peso. Dada uma matriz $A \in \mathbb{C}^{m\times n}$, dizemos que a matriz $X \in \mathbb{C}^{n\times m}$ é uma inversa-$\{i,j,k\}$ com peso de A se X satisfaz a i-ésima, a j-ésima e a k-ésima equação de Penrose com peso, dadas por (1.24)-(1.27).

Vejamos algumas aplicações das inversas com peso.

1.4.2 Inversa-$\{1,4\}$ com peso e a solução de norma-N mínima de sistemas de equações lineares consistentes

De maneira bastante similar ao caso em que a norma utilizada para escolha da solução era a norma-2, podemos escrever a solução de $Ax = b$ com norma-N mínima usando a inversa-$\{1,4\}$ com peso.

Teorema 1.34. *[29] Se $A \in \mathbb{C}^{m\times n}$, $b \in \mathbb{C}^m$ e $N \in \mathbb{C}^{n\times n}$, com $b \in \mathcal{R}(A)$ e N sendo uma matriz hermitiana positiva definida, então $x = Xb$ é a solução de norma-N mínima se e somente se X satisfaz*

$$AXA = A \quad e\ (NXA)^* = NXA.$$

Neste caso, o conjunto solução de $Ax = b$ é dado por

$$\{Xb + (I - XA)y : y \in \mathbb{C}^n\}. \tag{1.28}$$

Demonstração:

$(\Leftarrow)$:

Como $AXA = A$ e $b \in \mathcal{R}(A)$, ou seja, existe $x \in \mathbb{C}^n$ tal que $Ax = b$, então

$$AXb = AXAx = Ax = b,$$

portanto Xb satisfaz o sistema. Além disso, como $(I-XA)y \in \mathcal{N}(A)$, pois $A(I-XA)y = (A-AXA)y = (A-A)y = 0$ para todo $y \in \mathbb{C}^n$ então

$$A(Xb + (I-XA)y) = AXb + A(I-XA)y = AXb = b, \ \forall \ y \in \mathbb{C}^n.$$

Reciprocamente, se x é solução de $Ax = b$ então

$$x = Xb + x - Xb = Xb + x - XAx = Xb + (I-XA)x,$$

portanto a solução geral do sistema é dado por (1.28). Para mostrar que Xb tem norma-N mínima, precisamos mostrar que

$$\|Xb\|_N \leq \|Xb + (I-XA)y\|_N \quad \forall \ b \in \mathcal{R}(A), y \in \mathbb{C}^n,$$

e segue do item 7 do Lema 1.32 que

$$\begin{aligned}
\|XAu\|_N &\leq \|XAu + (I-XA)y\|_N, \ \ \forall \ u, y \in \mathbb{C}^n \\
&\Leftrightarrow \langle XAu, (I-XA)y\rangle_N = 0, \ \ \forall \ u, y \in \mathbb{C}^n \\
&\Leftrightarrow \langle u, (XA)^{\#}(I-XA)y\rangle_N = 0, \ \ \forall \ u, y \in \mathbb{C}^n \\
&\Leftrightarrow (XA)^{\#}(I-XA) = 0. \qquad (1.29)
\end{aligned}$$

Pelas propriedades da adjunta com peso temos

$$\begin{aligned}
(XA)^{\#}(I-XA) &= 0 \\
(XA)^{\#} - (XA)^{\#}XA &= 0 \\
(XA)^{\#} &= (XA)^{\#}XA \qquad (1.30) \\
((XA)^{\#})^{\#} &= ((XA)^{\#}XA)^{\#} \\
XA &= (XA)^{\#}XA, \qquad (1.31)
\end{aligned}$$

e segue de (1.30) e (1.31) que $(XA)^{\#} = XA$. Por outro lado, se $(XA)^{\#} = XA$ então do fato de $AXA = A$, temos

$$\begin{aligned}
(XA)^{\#} &= N^{-1}(XA)^*N \\
&= N^{-1}A^*X^*N \\
&= N^{-1}A^*X^*A^*X^*N \\
&= N^{-1}A^*X^*NN^{-1^*}A^*X^*N \\
&= (XA)^{\#}(XA)^{\#}
\end{aligned}$$

e assim

$$(XA)^{\#}(I-XA) = (XA)^{\#}-(XA)^{\#}XA = (XA)^{\#}-(XA)^{\#}(XA)^{\#} = (XA)^{\#}-(XA)^{\#} = 0.$$

Portanto, de (1.29) temos

$$\begin{aligned} \|XAu\|_N &\leq \|XAu + (I - XA)y\|_N, \quad \forall\, u, y \in \mathbb{C}^n \\ &\Leftrightarrow (XA)^{\#} = XA \\ &\Leftrightarrow (NXA)^* = NXA. \end{aligned} \tag{1.32}$$

De onde segue que Xb é a solução de norma-N mínima.

($\Rightarrow$) :

como Xb satisfaz o sistema para todo $b \in \mathcal{R}(A)$, tomando $A = [a_1, a_2, \ldots, a_n]$, temos que $a_i \in \mathcal{R}(A),\ i = 1, 2, \ldots, n$, e assim

$$AX[a_1, a_2, \ldots, a_n] = [a_1, a_2, \ldots, a_n],$$

ou seja, $AXA = A$. E como Xb é a solução de norma mínima, segue de (1.32) que X satisfaz $(NXA)^* = NXA$. ■

1.4.3 Inversa-$\{1,3\}$ com peso e a solução de mínimos quadrados com norma-M de sistemas de equações lineares inconsistentes

Também podemos utilizar a inversa-$\{1,3\}$ com peso para a escolha da solução do problema de mínimos quadrados com a norma-M, como mostra o seguinte teorema.

Teorema 1.35. *[29] Sejam $A \in \mathbb{C}^{m\times n}$, $b \in \mathbb{C}^m$ e $M \in \mathbb{C}^{m\times m}$, com $b \notin \mathcal{R}(A)$ e M sendo uma matriz hermitiana positiva definida. Temos que $x = Xb$ minimiza $\|Ax - b\|_M$ se e somente se X satisfaz*

$$AXA = A \quad e \quad (MAX)^* = MAX.$$

Demonstração:

($\Rightarrow$) :

Como

$$\begin{aligned} \|Ax - b\|_M &= \|AXb - b + Ax - AXb\|_M \\ &= \|(AX - I)b + A(x - Xb)\|_M \end{aligned}$$

então segue do item 7 do Lema 1.32 que

$$
\begin{aligned}
\|AXb - b\|_M &\leq \|Ax - b\|_M = \|(AX - I)b + A(x - Xb)\|_M \ \forall\ x \in \mathbb{C}^n, b \in \mathbb{C}^m \\
&\Leftrightarrow \langle A(x - Xb), (AX - I)b\rangle_M = 0 \ \forall\ x \in \mathbb{C}^n, b \in \mathbb{C}^m \\
&\Leftrightarrow \langle x - Xb, A^{\#}(AX - I)b\rangle_M = 0 \ \forall\ x \in \mathbb{C}^n, b \in \mathbb{C}^m \\
&\Leftrightarrow A^{\#}(AX - I) = 0 \\
&\Leftrightarrow A^{\#}AX - A^{\#} = 0 \\
&\Leftrightarrow A^{\#}AX = A^{\#}
\end{aligned} \tag{1.33}
$$

de onde temos que

$$(AX)^{\#} = X^{\#}A^{\#} = X^{\#}A^{\#}AX = (AX)^{\#}AX,$$

e tomando a adjunta com peso dos dois lados temos

$$AX = (AX)^{\#}AX,$$

portanto

$$(AX)^{\#} = AX.$$

Além disso,

$$A = (A^{\#}AX)^{\#} = X^{\#}A^{\#}A = (AX)^{\#}A = AXA.$$

$(\Leftarrow)$:

Como $(AX)^{\#} = AX$ então

$$A^{\#}AX = A^{\#}(AX)^{\#} = A^{\#}X^{\#}A^{\#} = (AXA)^{\#} = A^{\#},$$

e segue de (1.33) que

$$\|AXb - b\|_M \leq \|Ax - b\|_M \ \forall\ x \in \mathbb{C}^n, b \in \mathbb{C}^m.$$

■

1.4.4 Inversa de Moore-Penrose com peso e a solução do sistema $Ax = b$

O teorema a seguir relaciona a inversa de Moore-Penrose com peso e a solução do sistema $Ax = b$ para norma com peso.

Teorema 1.36. *[29] Dado um sistema de equações lineares $Ax = b$, com $A \in \mathbb{C}^{m\times n}$ e $b \in \mathbb{C}^m$, e dados $M \in \mathbb{C}^{m\times m}$ e $N \in \mathbb{C}^{n\times n}$ com M e N hermitianas e positivas definidas, então $x = Xb$ é um minizador de $\|Ax - b\|_M$, e além disso é o minimizador de norma-N mínima se e somente se $X = A^\dagger_{MN}$*

Demonstração:

($\Rightarrow$) :

Segue do Teorema 1.35 que X satisfaz $AXA = A$ e $(MAX)^* = MAX$, e portanto a solução geral de mínimos quadrados é dada por

$$Xb + (I - XA)z, \quad z \in \mathbb{C}^n.$$

Do item 7 do Lema 1.32 segue que

$$\begin{aligned}
\|Xb\|_N &\leq \|Xb + (I - XA)z\|_N, \ \forall\, b \in \mathbb{C}^m, z \in \mathbb{C}^n \\
&\Leftrightarrow \langle Xb, (I - XA)z\rangle_N = 0, \ \forall\, b \in \mathbb{C}^m, z \in \mathbb{C}^n \\
&\Leftrightarrow \langle b, X^\#(I - XA)z\rangle_N = 0, \ \forall\, b \in \mathbb{C}^m, z \in \mathbb{C}^n \\
&\Leftrightarrow X^\#(I - XA) = 0 \\
&\Leftrightarrow X^\# - X^\# XA = 0 \\
&\Leftrightarrow X^\# = X^\# XA, \qquad (1.34)
\end{aligned}$$

e assim

$$(XA)^\# = A^\# X^\# = A^\# X^\# XA = (XA)^\# XA,$$

tomando a ajunta com peso em ambos os lados segue que

$$XA = (XA)^\# XA,$$

de onde segue que $(XA)^\# = XA$. Além disso,

$$X = (X^\# XA)^\# = (XA)^\# X = XAX,$$

portanto $X = A^\dagger_{MN}$.

($\Leftarrow$) :

Como X satisfaz $AXA = A$ e $(MAX)^* = MAX$, então pelo Teorema 1.35 Xb minimiza $\|Ax - b\|_M$, para mostrar que Xb é o minimizador de norma-N mínima, veja que das equações $XAX = X$ e $(XA)^\# = XA$ segue que

$$X^\# XA = X^\#(XA)^\# = X^\# A^\# X^\# = (XAX)^\# = X^\#,$$

e por (1.34) temos que

$$\|Xb\|_N \leq \|Xb + (I - XA)z\|_N, \ \forall \ b \in \mathbb{C}^m, z \in \mathbb{C}^n.$$

■

1.5 Inversa de Drazin

Nas seções anteriores trabalhamos com as inversas-$\{i, j, k\}$, as quais incluem também a inversa de Moore-Penrose (inversa-$\{1, 2, 3, 4\}$), essas inversas possuem várias propriedades interessantes, assim como importantes aplicações, porém algumas propriedades da inversa usual não são satisfeitas por nenhuma classe de inversas-$\{i, j, k\}$. Por exemplo se $A, B \in \mathbb{C}^{n\times n}$, e A^- e B^- denotam inversas-$\{i, j, k\}$ de A e B respectivamente para alguma classe $\{i, j, k\}$ de inversas, então nenhuma das propriedades abaixo são válidas:

1. $AA^- = A^-A$;
2. $(A^-)^p = (A^p)^-$, para p inteiro positivo;
3. $\lambda \in \sigma(A)$ se e somente se $\lambda^\dagger \in \sigma(A^-)$, sendo $\sigma(A)$ o espectro da matriz A, ou seja, o conjunto de todos os autovalores de A;
4. $A^{p+1}A^- = A^p$, para p inteiro positivo;
5. Se $P \in \mathbb{C}^{n\times n}$ é não singular e $P^{-1}AP = B$, então $P^{-1}A^-P = B^-$.

A inversa de Drazin satisfaz todas as propriedades acima.

1.5.1 A inversa de Drazin

Definição 1.37. Dada $A \in \mathbb{C}^{n\times n}$, o menor inteiro positivo k para o qual

$$\text{posto}(A^{k+1}) = \text{posto}(A^k),$$

onde A^k denota a k-ésima potência da matriz A, é chamado de index de A, e denotado por $\text{Ind}(A) = k$. Se A é não singular, então definimos $\text{Ind}(A) = 0$.

De agora em diante, vamos considerar apenas o caso A singular. Vejamos algumas propriedades do index de uma matriz.

Teorema 1.38. *[29] Considere $A \in \mathbb{C}^{n\times n}$.*

(i) Se Ind(A) = k, então

$$\begin{aligned} posto(A^l) &= posto(A^k), \quad para\ l \geq k; && (1.35)\\ \mathcal{R}(A^l) &= \mathcal{R}(A^k) \quad para\ l \geq k; && (1.36)\\ \mathcal{N}(A^l) &= \mathcal{R}(A^k) \quad para\ l \geq k. && (1.37) \end{aligned}$$

(ii) Ind(A) = k $\Leftrightarrow$ k é o menor inteiro positivo tal que

$$A^k = A^{k+1}X,$$

para alguma matriz X.

(iii) Ind(A) = k $\Leftrightarrow$ $\mathcal{R}(A^k)$ e $\mathcal{N}(A^k)$ são subespaços complementares, ou seja,

$$\mathcal{R}(A^k) \oplus \mathcal{N}(A^k) = \mathbb{C}^n.$$

Demonstração:

(i) Por definição de index, temos que posto(A^k) =posto(A^{k+1}), e pelo item 5 do Teorema 1.5 isso vale se e somente se $\mathcal{R}(A^k) = \mathcal{R}(A^{k+1})$.

Agora, note que se $\mathcal{R}(A^k) = \mathcal{R}(A^{k+1})$ então para $A^k = [a_1^k, a_2^k, \ldots, a_n^k]$ temos que $a_i^k \in \mathcal{R}(A^k), i = 1, 2, \ldots, n$, e existe $X = [x_1, x_2, \ldots, x_n]$ tal que $a_i^k = A^{k+1}x_i, i = 1, 2, \ldots, n$, logo

$$A^k = A^{k+1}X.$$

Reciprocamente se $A^k = A^{k+1}X$ para algum X, então para $y \in \mathcal{R}(A^k)$, existe z tal que $A^k z = y$, e assim

$$A^{k+1}Xz = A^k z = y,$$

o que implica que $y \in \mathcal{R}(A^{k+1})$.

Por outro lado, se $y \in \mathcal{R}(A^{k+1})$ então existe z tal que $A^{k+1}z = y$ e portanto

$$y = A^{k+1}z = A^k Az,$$

ou seja, $y \in \mathcal{R}(A^k)$.

Assim, $\mathcal{R}(A^k) = \mathcal{R}(A^{k+1})$ se e somente se existe X tal que $A^k = A^{k+1}X$.

Agora, multiplicando $A^k = A^{k+1}X$ a esquerda por A^{l-k}, para $l \geq k$ temos

$$A^l = A^{l+1}X,$$

o que pelo mostrado acima é equivalente a

$$\text{posto}(A^l) = \text{posto}(A^{l+1}),$$

que por sua vez, pelos itens 5 e 6 do teorema 1.5, é equivalente a

$$\mathcal{R}(A^l) = \mathcal{R}(A^{l+1}) \quad \text{e} \quad \mathcal{N}(A^l) = \mathcal{N}(A^{l+1}).$$

Como o resultado vale para todo $l \geq k$, segue por indução as afirmações (1.35) - (1.37).

(ii) Segue da equivalência entre posto(A^k) =posto(A^{k+1}) e $A^k = A^{k+1}X$ para alguma matriz X, mostrada no item anterior.

(iii) Pelo Teorema do núcleo e imagem, e o fato de que o posto de uma matriz coincide com a dimensão da imagem dessa matriz, então posto(A^k) >posto(A^{k+1}) se e somente se existe $x \in \mathbb{C}^n$ tal que

$$A^{k+1}x = 0, \quad \text{e} \quad A^k x \neq 0. \tag{1.38}$$

Além disso, se existe x satisfazendo (1.38), então $y = A^k x \neq 0$ pertence a imagem de A^k, e

$$A^k y = A^k A^k x = A^{k-1} A^{k+1} x = 0,$$

ou seja y pertence ao núcleo de A^k, logo

$$\mathcal{R}(A^k) \cap \mathcal{N}(A^k) \neq \{0\}.$$

Por outro lado, se $\mathcal{R}(A^k) \cap \mathcal{N}(A^k) \neq \{0\}$ então existe $y \neq 0$ tal que $A^k x = y$ para algum x e $A^k y = 0$. Assim

$$\begin{aligned} A^k x &= y \neq 0, \\ A^{k+1} x &= A A^k x = Ay. \end{aligned}$$

Se $Ay = 0$ então x satisfaz (1.38). Caso $Ay \neq 0$, então consideramos o vetor Ax, o qual satisfaz

$$\begin{aligned} A^k(Ax) &= A^{k+1}x = Ay \neq 0, \\ A^{k+1}(Ax) &= A^2 y. \end{aligned}$$

Se $A^2 y = 0$ então o vetor Ax satisfaz (1.38). Caso $A^2 y \neq 0$, consideramos o vetor $A^2 x$. Continuando o processo até que $A^q y = 0$, o que ocorrerá para $q \leq k$ uma vez que $y \in \mathcal{N}(A^k)$, teremos que o vetor $A^{q-1}x$ satisfaz (1.38).

Portanto $\text{Ind}(A) = k$ se e somente se $\mathcal{R}(A^k) \cap \mathcal{N}(A^k) = \{0\}$, o que pelo Teorema do núcleo e imagem é equivalente a $\mathcal{R}(A^k) \oplus \mathcal{N}(A^k) = \mathbb{C}^n$.

■

Definição 1.39. Se $A \in \mathbb{C}^{n\times n}$, com $\text{Ind}(A) = k$, então a matriz X satisfazendo

$$A^k X A = A^k \tag{1.39}$$

$$XAX = X \tag{1.40}$$

$$AX = XA \tag{1.41}$$

é chamada a inversa de Drazin de A, a qual é denotada por A_d.

Note que usando a equação (1.41) nas equações (1.39) e (1.40) podemos reescrever as equações (1.39)-(1.41) como

$$A^{k+1} X = A^k, \tag{1.42}$$

$$AX^2 = X, \tag{1.43}$$

$$AX = XA. \tag{1.44}$$

Vejamos que a inversa de Drazin existe e é única.

Teorema 1.40. *[29] Se $A \in \mathbb{C}^{n\times n}$, com Ind$(A) = k$, então a inversa de Drazin A_d de A existe e é única.*

Demonstração:

Existência:

Segue do item (iii) do Teorema 1.38 que

$$\mathcal{R}(A^k) \oplus \mathcal{N}(A^k) = \mathbb{C}^n.$$

Sejam $\{p_1, p_2, \dots, p_r\}$ uma base de $\mathcal{R}(A^k)$ e $\{p_{r+1}, p_{r+2}, \dots, p_n\}$ uma base de $\mathcal{N}(A^k)$, pelo fato de que $\mathcal{R}(A^k)$ e $\mathcal{N}(A^k)$ são invariantes por A, se considerarmos

$$P = (P_1, P_2), \quad \text{sendo} \quad P_1 = (p_1, p_2, \dots, p_r) \quad \text{e} \quad P_2 = (p_{r+1}, p_{r+2}, \dots, p_n),$$

temos que existem matrizes $C \in \mathbb{C}^{r\times r}$ e $N \in \mathbb{C}^{(n-r)\times(n-r)}$ tal que

$$AP_1 = P_1 C \quad \text{e} \quad AP_2 = P_2 N,$$

de onde segue que

$$AP = A(P_1, P_2) = (AP_1, AP_2) = (P_1C, P_2N) = (P_1, P_2)\begin{pmatrix} C & 0 \\ 0 & N \end{pmatrix} = P\begin{pmatrix} C & 0 \\ 0 & N \end{pmatrix},$$

e portanto

$$A = P\begin{pmatrix} C & 0 \\ 0 & N \end{pmatrix}P^{-1}.$$

Assim,

$$A^k = P\begin{pmatrix} C^k & 0 \\ 0 & N^k \end{pmatrix}P^{-1},$$

porém, como

$$\begin{aligned} AP_2 &= P_2N \\ A^2P_2 &= AP_2N = P_2NN = P_2N^2 \\ &\vdots \\ A^kP_2 &= P_2N^k, \end{aligned}$$

e como P_2 é base de $\mathcal{N}(A^k)$, então $A^kP_2 = 0$, o que implica $N^k = 0$, e assim

$$A^k = P\begin{pmatrix} C^k & 0 \\ 0 & 0 \end{pmatrix}P^{-1},$$

de onde podemos concluir que

$$r = \text{posto}(A^k) = \text{posto}(C^k) \leq r,$$

ou seja, posto(C^k) =posto$(C) = r$, e portanto C é não singular.

Defina

$$X = P\begin{pmatrix} C^{-1} & 0 \\ 0 & 0 \end{pmatrix}P^{-1},$$

vejamos que X satisfaz as equações (1.39)-(1.41).

$$
\begin{aligned}
A^kXA &= P\begin{pmatrix} C^k & 0 \\ 0 & 0 \end{pmatrix} P^{-1}P\begin{pmatrix} C^{-1} & 0 \\ 0 & 0 \end{pmatrix} P^{-1}P\begin{pmatrix} C & 0 \\ 0 & 0 \end{pmatrix} P^{-1} \\
&= P\begin{pmatrix} C^k & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} C^{-1} & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} C & 0 \\ 0 & 0 \end{pmatrix} P^{-1} \\
&= P\begin{pmatrix} C^k & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix} P^{-1} \\
&= P\begin{pmatrix} C^k & 0 \\ 0 & 0 \end{pmatrix} P^{-1} \\
&= A^k;
\end{aligned}
$$

$$
\begin{aligned}
XAX &= P\begin{pmatrix} C^{-1} & 0 \\ 0 & 0 \end{pmatrix} P^{-1}P\begin{pmatrix} C & 0 \\ 0 & 0 \end{pmatrix} P^{-1}P\begin{pmatrix} C^{-1} & 0 \\ 0 & 0 \end{pmatrix} P^{-1} \\
&= P\begin{pmatrix} C^{-1} & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} C & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} C^{-1} & 0 \\ 0 & 0 \end{pmatrix} P^{-1} \\
&= P\begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} C^{-1} & 0 \\ 0 & 0 \end{pmatrix} P^{-1} \\
&= P\begin{pmatrix} C^{-1} & 0 \\ 0 & 0 \end{pmatrix} P^{-1} \\
&= X;
\end{aligned}
$$

e

$$
\begin{aligned}
AX &= P\begin{pmatrix} C & 0 \\ 0 & 0 \end{pmatrix} P^{-1}P\begin{pmatrix} C^{-1} & 0 \\ 0 & 0 \end{pmatrix} P^{-1} \\
&= P\begin{pmatrix} C & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} C^{-1} & 0 \\ 0 & 0 \end{pmatrix} P^{-1} \\
&= P\begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix} P^{-1},
\end{aligned}
$$

$$\begin{aligned}
XA &= P\begin{pmatrix} C^{-1} & 0 \\ 0 & 0 \end{pmatrix} P^{-1}P\begin{pmatrix} C & 0 \\ 0 & 0 \end{pmatrix} P^{-1} \\
&= P\begin{pmatrix} C^{-1} & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} C & 0 \\ 0 & 0 \end{pmatrix} P^{-1} \\
&= P\begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix} P^{-1},
\end{aligned}$$

o que mostra que $AX = XA$.

Unicidade:

Sejam X e Y inversas de Drazin de A, chamaremos

$$E = AX = XA \quad \text{e} \quad F = AY = YA.$$

Neste caso, temos que

$$\begin{aligned}
E^2 &= EE = AXAX = AX = E \\
F^2 &= FF = AYAY = AY = F
\end{aligned}$$

assim

$$\begin{aligned}
E &= AX = AXAX = AAXX = A^2X^2 = \ldots = A^kX^k \\
&= A^kYAX^k = A^{k-1}AYAX^k = A^{k-1}YAAX^k = A^{k-1}YA^2X^k = \ldots = AYA^kX^k \\
&= FA^kX^k = FAX = FE,
\end{aligned}$$

e

$$\begin{aligned}
F &= YA = YAYA = YYAA = Y^2A^2 = \ldots = Y^kA^k \\
&= Y^kA^kXA = YAXA = YAE = FE.
\end{aligned}$$

Portanto $E = F$, e

$$X = AX^2 = EX = FX = YAX = YE = YF = YYA = YAY = AYY = AY^2 = Y.$$

■

Definição 1.41. Seja $A \in \mathbb{C}^{n\times n}$. Se X satisfaz as equações

$$AXA = A; \tag{1.45}$$

$$XAX = X; \tag{1.46}$$

$$AX = XA, \tag{1.47}$$

então X é chamada inversa de grupo de A, e denotada por A_g.

Note que se $\mathrm{Ind}(A) = 1$, então A_d é uma inversa de grupo da matriz A, e nesse caso vimos que ela existe e é única. O teorema a seguir estabelece condições de existência e unicidade de A_g.

Teorema 1.42. *[29] Seja $A \in \mathbb{C}^{n\times n}$ uma matriz singular. Temos que A possui uma inversa de grupo se e somente se $\mathrm{Ind}(A) = 1$, neste caso ela é única.*

Demonstração:

Ver Teorema 2.2.1 de [29]. ■

Para $A \in \mathbb{C}^{n\times n}$ não singular, se

$$p(\lambda) = \det(\lambda I - A) = \lambda^n + a_1\lambda^{n-1} + \cdots + a_{n-1}\lambda + a_n$$

é o polinômio característico de A, então pelo Teorema de Cayley-Hamilton $p(A) = 0$, de onde segue que

$$A^{-1} = -\frac{1}{a_n}A^{n-1} - \frac{a_1}{a_n}A^{n-2} - \cdots - \frac{a_{n-2}}{a_n}A - \frac{a_{n-1}}{a_n}I.$$

Essa propriedade de escrever a inversa como um polinômio não é preservada pelas inversas-$\{i, j, k\}$, porém podemos expressar a inversa de Drazin como um polinômio, como mostra o próximo teorema.

Teorema 1.43. *[29] Se $A \in \mathbb{C}^{n\times n}$ com $\mathrm{Ind}(A) = k$, então existe um polinômio $q(\lambda)$ tal que*

$$A_d = A^l(q(A))^{l+1}, \quad l \geq k.$$

Demonstração:

Vimos que A pode ser escrita como

$$A = P\begin{pmatrix} C & 0 \\ 0 & N \end{pmatrix}P^{-1},$$

com P e C não singulares e N satisfazendo $N^k = 0$. Como C é não singular, então existe um polinômio $q(\lambda)$ tal que $C^{-1} = q(C)$, de onde segue que para $l \geq k$

$$
\begin{aligned}
A_d &= P\begin{pmatrix} C^{-1} & 0 \\ 0 & 0 \end{pmatrix} P^{-1} \\
&= P\begin{pmatrix} q(C) & 0 \\ 0 & 0 \end{pmatrix} P^{-1} \\
&= P\begin{pmatrix} C^l q(C)^{l+1} & 0 \\ 0 & 0 \end{pmatrix} P^{-1} \\
&= P\begin{pmatrix} C^l & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} q(C)^{l+1} & 0 \\ 0 & q(N)^{l+1} \end{pmatrix} P^{-1} \\
&= P\begin{pmatrix} C^l & 0 \\ 0 & 0 \end{pmatrix} P^{-1}P \begin{pmatrix} q(C)^{l+1} & 0 \\ 0 & q(N)^{l+1} \end{pmatrix} P^{-1} \\
&= P\begin{pmatrix} C^l & 0 \\ 0 & 0 \end{pmatrix} P^{-1}P \begin{pmatrix} q(C) & 0 \\ 0 & q(N) \end{pmatrix}^{l+1} P^{-1} \\
&= A^l q(A)^{l+1}.
\end{aligned}
$$

■

A inversa de Drazin também possui algumas propriedades em relação ao núcleo e a imagem da matriz

Teorema 1.44. *[29] Se $A \in \mathbb{C}^{n\times n}$, com Ind$(A) = k$, então*

$$\mathcal{R}(A_d) = \mathcal{R}(A^l), \quad e \quad \mathcal{N}(A_d) = \mathcal{N}(A^l), \quad l \geq k.$$

Demonstração:

Considere $A_d = X$. Temos que

$$A^k = A^{k+1}X = XA^{k+1},$$

multiplicando a direita por A^{l-k} temos

$$A^l = XA^{l+1}, \qquad l \geq k,$$

assim se $y \in \mathcal{R}(A^l)$ temos que existe x tal que $A^l x = y$, e então

$$y = A^l x = XA^{l+1}x,$$

de onde temos $\mathcal{R}(A^l) \subset \mathcal{R}(X)$. Por outro lado se $y \in \mathcal{R}(X)$ então existe x tal que $Xx = y$

e pelo Teorema 1.43 segue que

$$y = Xx = A^l q(A)^{l+1} x,$$

portanto $\mathcal{R}(X) \subset \mathcal{R}(A^l)$, de onde segue que

$$\mathcal{R}(X) = \mathcal{R}(A^l).$$

Para ver que os núcleos coincidem, pelo item 5 do Teorema 1.5 temos que

$$\text{posto}(X) = \text{posto}(A^l),$$

de onde obtemos que dim($\mathcal{N}(X)$) =dim($\mathcal{N}(A^l)$).

Além disso, se $x \in \mathcal{N}(A^l)$ então pelo Teorema 1.43 temos

$$Xx = A^l q(A)^{l+1} x = 0,$$

portanto $\mathcal{N}(A^l) \subset \mathcal{N}(X)$ e como os dois espaços possuem a mesma dimensão temos

$$\mathcal{N}(X) = \mathcal{N}(A^l).$$

■

1.5.2 Aplicação da inversa de Drazin em equações de diferenças

Vejamos a seguir uma maneira de aplicar a inversa de Drazin em equações de diferenças. Para tanto considere o seguinte resultado.

Lema 1.45. *[5] Seja $A \in \mathbb{C}^{n\times n}$, com Ind$(A) = k$. Se o problema*

$$Ax = b, \quad \textit{sujeito a} \quad x \in \mathcal{R}(A^k) \tag{1.48}$$

tem solução, então a solução é unicamente determinada por

$$x = A_d b.$$

Demonstração:

Seja x uma solução do problema (1.48). Neste caso existe $y \in \mathbb{C}^n$ tal que $x = A^k y$, assim segue da definição de A_d que

$$\begin{aligned} x &= A^k y = A^{k+1} A_d y = A^k A A_d y = A^k A_d A y = \ldots = A_d A^{k+1} y \\ &= A_d A A^k y = A_d A x = A_d b, \end{aligned}$$

portanto toda solução de (1.48) pode ser escrita como $A_d b$, e da unicidade de A_d segue que $A_d b$ é a única solução. ∎

O exemplo a seguir mostra uma aplicação da inversa de Drazin em equações de diferenças.

Exemplo 1.1. Seja $A \in \mathbb{C}^{n\times n}$ uma matriz singular de index k. Considere a equação de diferenças

$$Ax_{t+1} = x_t, \quad t = 0, 1, 2, \dots .$$

Como

$$x_{t+1} = Ax_{t+2} = A^2 x_{t+3} = A^k x_{t+k+1},$$

então para um dado x_t, a solução x_{t+1} pertence a $\mathcal{R}(A^k)$, e pelo lema acima temos

$$x_{t+1} = A_d x_t, \quad t = 0, 1, 2, \dots .$$

Vejamos mais alguns detalhes sobre esse tipo de equação de diferenças. Vamos considerar equações da forma

$$Ax_{t+1} = Bx_t + f_t, \;\; \text{sujeito a} \;\; x_0 = c, \tag{1.49}$$

onde $A, B \in \mathbb{C}^{n\times n}$, com A singular, $c, f_t \in \mathbb{C}^n$ para $t = 0, 1, 2, \dots$. A equação (1.49) é dita homogênea se $f_t = 0$ para todo t.

Definição 1.46. O vetor inicial c é dito ser consistente se, dado x_t a equação (1.49) tem solução.

Definição 1.47. A equação (1.49) é dita tratável se existe uma única solução para cada c consistente.

O próximo teorema caracteriza equações tratáveis de maneira simples para o caso homogêneo.

Teorema 1.48. *[5] A equação homogênea*

$$Ax_{t+1} = Bx_t, \tag{1.50}$$

é tratável se e somente se existe um escalar $\lambda \in \mathbb{C}$ tal que $(\lambda A - B)$ é não singular.

Demonstração:

$(\Leftarrow)$:

Supondo que $(\lambda A - B)$ é não singular, então verificar se

$$Ax_{t+1} = Bx_t$$

é tratável, é equivalente a verificar que

$$(\lambda A - B)^{-1} A x_{t+1} = (\lambda A - B)^{-1} B x_t$$

é tratável. Para isso, baseado na forma canônica de Jordan, podemos escrever $(\lambda A - B)^{-1}A$ da forma

$$(\lambda A - B)^{-1} A = X \begin{pmatrix} J_1 & 0 \\ 0 & J_0 \end{pmatrix} X^{-1},$$

onde J_1 é não singular e J_0 é o bloco de Jordan referente ao autovalor 0. Como

$$\lambda(\lambda A - B)^{-1} A - (\lambda A - B)^{-1} B = (\lambda A - B)^{-1}(\lambda A - B) = I,$$

segue que

$$(\lambda A - B)^{-1} B = \lambda(\lambda A - B)^{-1} A - I = X \begin{pmatrix} \lambda J_1 - I & 0 \\ 0 & \lambda J_0 - I \end{pmatrix} X^{-1},$$

e particionando x_t da forma

$$x_t = \begin{pmatrix} x_t^{(1)} \\ x_t^{(0)} \end{pmatrix}, \quad t = 0, 1, 2, \ldots,$$

precisamos analisar a tratabilidade da equação

$$\begin{pmatrix} J_1 & 0 \\ 0 & J_0 \end{pmatrix} \begin{pmatrix} x_{t+1}^{(1)} \\ x_{t+1}^{(0)} \end{pmatrix} = \begin{pmatrix} \lambda J_1 - I & 0 \\ 0 & \lambda J_0 - I \end{pmatrix} \begin{pmatrix} x_t^{(1)} \\ x_t^{(0)} \end{pmatrix}$$

que é equivalente a

$$\begin{cases} J_1 x_{t+1}^{(1)} & = & (\lambda J_1 - I) x_t^{(1)}, \\ J_0 x_{t+1}^{(0)} & = & (\lambda J_0 - I) x_t^{(0)}. \end{cases}$$

Como J_1 é não singular, então a primeira equação é tratável, e precisamos mostrar que

$$J_0 x_{t+1}^{(0)} = (\lambda J_0 - I) x_t^{(0)} \tag{1.51}$$

é tratável. Seja $k = \text{Ind}(J_0)$, como J_0 é nilpotente, e J_0 é o bloco de Jordan referente ao autovalor 0, temos que $(\lambda J_0 - I)$ possui -1 na diagonal principal e λ na diagonal acima, o que nos possibilita escrever

$$(\lambda J_0 - I)^{-1} = -\left(\sum_{i=1}^{k-1} \lambda^i J_0^i + I \right),$$

de onde temos que

$$\begin{aligned}
(\lambda J_0 - I)^{-1} J_0 &= -\left(\sum_{i=1}^{k-1} \lambda^i J_0^i + I\right) J_0 \\
&= -\left(\sum_{i=1}^{k-1} \lambda^i J_0^{i+1} + J_0\right) \\
&= -J_0 \left(\sum_{i=1}^{k-1} \lambda^i J_0^i + I\right) \\
&= J_0(\lambda J_0 - I)^{-1}.
\end{aligned}$$

Assim,

$$\begin{aligned}
x_t^{(0)} &= (\lambda J_0 - I)^{-1} J_0 x_{t+1}^{(0)} \\
&= [(\lambda J_0 - I)^{-1} J_0]^2 x_{t+2}^{(0)} \\
&= \ldots \\
&= [(\lambda J_0 - I)^{-1} J_0]^k x_{t+k}^{(0)} \\
&= (\lambda J_0 - I)^{-k} J_0^k x_{t+k}^{(0)} \\
&= 0.
\end{aligned}$$

Portanto, (1.51) é trivialmente tratável.

($\Rightarrow$) :

Suponha por absurdo que $(\lambda A - B)$ é singular para todo $\lambda \in \mathbb{C}$, então para cada $\lambda \in \mathbb{C}$ existe um autovalor nulo de $(\lambda A - B)$, ou seja,

$$\exists\ v^\lambda \in \mathbb{C}^n,\ v^\lambda \neq 0,\ (\lambda A - B) v^\lambda = 0, \quad \forall\ \lambda \in \mathbb{C}.$$

Considere um conjunto linearmente dependente de s elementos dados por

$$\{v^{\lambda_1}, v^{\lambda_2}, \ldots, v^{\lambda_s}\},$$

de onde temos que existe um conjunto

$$\{\alpha_1, \alpha_2, \ldots, \alpha_s\}$$

não todos nulos, tal que

$$\alpha_1 v^{\lambda_1} + \alpha_2 v^{\lambda_2} + \ldots + \alpha_s v^{\lambda_s} = 0.$$

Defina

$$x_t^{\lambda_i} = \lambda_i^t v^{\lambda_i},$$

então, usando que $(\lambda_i A - B)v^{\lambda_i} = 0$ temos

$$Ax_{t+1}^{\lambda_i} = \lambda_i^{t+1} A v^{\lambda_i} = \lambda^t B v^{\lambda_i} = B x_t^{\lambda_i}, \quad t = 0, 1, 2, \ldots \ .$$

Agora, definindo

$$x_t = \sum_{i=1}^{s} \alpha_i x_t^{\lambda_i},$$

temos que

$$\begin{aligned} Ax_{t+1} &= A(\alpha_1 x_{t+1}^{\lambda_1} + \alpha_2 x_{t+1}^{\lambda_2} + \ldots + \alpha_s x_{t+1}^{\lambda_s}) \\ &= \alpha_1 A x_{t+1}^{\lambda_1} + \alpha_2 A x_{t+1}^{\lambda_2} + \ldots + \alpha_s A x_{t+1}^{\lambda_s} \\ &= \alpha_1 B x_t^{\lambda_1} + \alpha_2 B x_t^{\lambda_2} + \ldots + \alpha_s B x_t^{\lambda_s} \\ &= B(\alpha_1 x_t^{\lambda_1} + \alpha_2 x_t^{\lambda_2} + \ldots + \alpha_s x_t^{\lambda_s}) \\ &= B x_t, \end{aligned}$$

ou seja, a sequência $\{x_1, x_2, \ldots\}$ é uma solução de (1.50), não identicamente nula e satisfaz a condição inicial $x_0 = \sum_{i=1}^{s} \alpha_i v^{\lambda_i} = 0$. No entanto, a sequência $\{x_t = 0; t = 0, 1, 2, \ldots\}$ é outra solução de (1.50) com a mesma condição inicial, absurdo pois a equação é tratável. Portanto, segue que existe $\lambda \in \mathbb{C}$ tal que $(\lambda A - B)$ é não singular. ■

Teorema 1.49. *[5] Se a equação homogênea*

$$Ax_{t+1} = Bx_t, \quad t = 0, 1, 2, \ldots$$

é tratável, então a solução geral é dada por

$$x_t = \begin{cases} \widehat{A}\widehat{A}_d y, & t = \ \ 0, \\ (\widehat{A}_d \widehat{B})^t \widehat{A}\widehat{A}_d y, & t = \ \ 1, 2, 3, \ldots, \end{cases}$$

onde $\widehat{A} = (\lambda A - B)^{-1} A$, $\widehat{B} = (\lambda A - B)^{-1} B$, com $\lambda \in \mathbb{C}$ tal que $(\lambda A - B)$ é não singular, e $y \in \mathbb{C}^n$ arbitrário. Além disso, uma condição inicial c é consistente se e somente se $c \in \mathcal{R}(\widehat{A}^k)$.

Demonstração:

Usando o mesmo procedimento da demonstração do teorema anterior, podemos escrever

$$\widehat{A} = X \begin{pmatrix} J_1 & 0 \\ 0 & J_0 \end{pmatrix} X^{-1}$$

e

$$\widehat{B} = X \begin{pmatrix} \lambda J_1 - I & 0 \\ 0 & \lambda J_0 - I \end{pmatrix} X^{-1},$$

e então temos o sistema equivalente

$$\left\{ \begin{array}{rcl} J_1 x_{t+1}^{(1)} & = & (\lambda J_1 - I) x_t^{(1)}, \\ J_0 x_{t+1}^{(0)} & = & (\lambda J_0 - I) x_t^{(0)}. \end{array} \right.$$

onde J_1 é não singular e J_0 nilpotente de index k. Pode-se mostrar que neste caso, a inversa de Drazin da matriz $\widehat{A}$ é dada por

$$\widehat{A}_d = X \begin{pmatrix} J_1^{-1} & 0 \\ 0 & 0 \end{pmatrix} X^{-1}.$$

Na demonstração do teorema anterior vimos que $x_t^{(0)} = 0$, $t = 0, 1, 2, \ldots$, além disso

$$\begin{array}{rcl} x_t^{(1)} & = & J_1^{-1}(\lambda J_1 - I) x_{t-1}^{(1)} \\ & = & [J_1^{-1}(\lambda J_1 - I)]^2 x_{t-2}^{(1)} \\ & = & \ldots \\ & = & [J_1^{-1}(\lambda J_1 - I)]^t x_0^{(1)}, \end{array}$$

de onde temos que

$$x_t = \begin{pmatrix} x_t^{(1)} \\ x_t^{(0)} \end{pmatrix} = \left[\begin{pmatrix} J_1^{-1} & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} (\lambda J_1 - I) & 0 \\ 0 & (\lambda J_0 - I) \end{pmatrix} \right]^t \begin{pmatrix} x_0^{(1)} \\ x_0^{(0)} \end{pmatrix} = (\widehat{A}_d \widehat{B})^t x_0$$

para $t = 1, 2, \ldots$.

Para $t = 0$, como $x_0^{(0)} = 0$ então a condição inicial é da forma

$$\begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix} y, \quad y \in \mathbb{C}^n,$$

a qual pode ser escrita como

$$\widehat{A}\widehat{A}_d y = \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix} y, \quad y \in \mathbb{C}^n,$$

de onde segue que

$$x_0 = \widehat{A}\widehat{A}_d y, \quad y \in \mathbb{C}^n,$$

portanto

$$x_t = \left\{ \begin{array}{ll} \widehat{A}\widehat{A}_d y, & t = \ 0, \\ (\widehat{A}_d \widehat{B})^t \widehat{A}\widehat{A}_d y, & t = \ 1, 2, 3, \ldots \ . \end{array} \right.$$

Além disso, segue que uma condição inicial é consistente se e somente se é da forma $\widehat{A}\widehat{A}_d y$, com $y \in \mathbb{C}^n$ arbitrário, e se $c \in \mathcal{R}(\widehat{A}^k)$, então

$$c = \widehat{A}^k z = \widehat{A}^k \widehat{A}_d \widehat{A} z = \widehat{A}\widehat{A}_d \widehat{A}^k z,$$

por outro lado, se $c = \widehat{A}\widehat{A}_d y$ então

$$c = \widehat{A}\widehat{A}_d y = \widehat{A}_d \widehat{A} y,$$

portanto $c \in \mathcal{R}(\widehat{A}_d) = \mathcal{R}(\widehat{A}^k)$.

■

O resultado do teorema anterior pode ser generalizado para o caso não homogêneo, onde a solução geral da equação

$$Ax_{t+1} = Bx_t + f_t,$$

é dada por

$$x_t = (\widehat{A}_d\widehat{B})^t \widehat{A}\widehat{A}_d y + \widehat{A}_d \sum_{i=0}^{t-1} (\widehat{A}_d\widehat{B})^{t-i-1} \widehat{f}_i - (I - \widehat{A}\widehat{A}_d) \sum_{i=0}^{k-1} (\widehat{A}\widehat{B}_d)^i \widehat{B}_d \widehat{f}_{t+i},$$

onde $\widehat{A} = (\lambda A - B)^{-1} A$, $\widehat{B} = (\lambda A - B)^{-1} B$, $\widehat{f}_i = (\lambda A - B)^{-1} f_i$, $k = \text{Ind}(\widehat{A})$ e $y \in \mathbb{C}^n$ é arbitrário.

1.5.3 A inversa de Drazin com peso

Em 1980, Cline e Greville [9] estenderam a definição de inversa de Drazin para matrizes retangulares.

Definição 1.50. Se $A \in \mathbb{C}^{m\times n}$ e $W \in \mathbb{C}^{n\times m}$, então a matriz $X \in \mathbb{C}^{m\times n}$ satisfazendo

$$(AW)^{k+1}XW = (AW)^k, \quad \text{para algum } k \text{ inteiro não negativo,} \tag{1.52}$$
$$XWAWX = X, \tag{1.53}$$
$$AWX = XWA \tag{1.54}$$

é chamada inversa de Drazin com peso W de A, e denotada por $X = A_{d,W}$.

Se $W = I$ e $A \in \mathbb{C}^{n\times n}$ então recaímos na inversa de Drazin A_d. Os resultados seguintes mostram questões de existência e unicidade para a inversa de Drazin com peso.

Teorema 1.51. *[29] Seja $A \in \mathbb{C}^{m\times n}$. Se existir uma matriz $X \in \mathbb{C}^{m\times n}$ satisfazendo as equações (1.52) - (1.54) para alguma matriz $W \in \mathbb{C}^{n\times m}$, então ela deve ser única.*

Demonstração:

Suponha que X_1 e X_2 satisfazem as equações (1.52) - (1.54), para inteiros não nulos k_1 e k_2 respectivamente, defina $k = \max\{k_1, k_2\}$. Temos que

$$\begin{aligned}
X_1 &= X_1WAWX_1 \\
&= AWX_1WX_1 \\
&= AWX_1WAWX_1WX_1 \\
&= AWAWX_1WX_1WX_1 \\
&= (AW)^2X_1(WX_1)^2 \\
&= \dots \\
&= (AW)^kX_1(WX_1)^k \\
&= (AW)^{k+1}AWX_2WX_1(WX_1)^k \\
&= (AW)^{k+1}X_2WAWX_1(WX_1)^k \\
&= (AW)^kAWX_2WAWX_1(WX_1)^k \\
&= (AW)^kX_2WAWAWX_1(WX_1)^k \\
&= (AW)^kX_2(WA)^2WX_1(WX_1)^k \\
&= \dots
\end{aligned}$$

$$\begin{aligned}
&= X_2(WA)^{k+1}WX_1(WX_1)^k \\
&= X_2(WA)^kWAWX_1WX_1(WX_1)^{k-1} \\
&= X_2(WA)^kWX_1WAWX_1(WX_1)^{k-1} \\
&= X_2(WA)^kWX_1(WX_1)^{k-1} \\
&= \dots \\
&= X_2WAWX_1.
\end{aligned}$$

O procedimento acima também vale para X_2 no lugar de X_1, de onde podemos escrever

$$X_2 = (AW)^{k+1}X_2(WX_2)^{k+1},$$

e da equação 1.54, temos que

$$AWX_2W = X_2WAW,$$

portanto

$$\begin{aligned}
X_2W &= (AW)^{k+1}X_2(WX_2)^{k+1}W \\
&= (AW)^{k+1}(X_2W)^{k+2} \\
&= (X_2W)^{k+2}(AW)^{k+1},
\end{aligned}$$

de onde segue que

$$\begin{aligned}
X_1 &= X_2WAWX_1 \\
&= (X_2W)^{k+2}(AW)^{k+1}AWX_1 \\
&= (X_2W)^{k+2}(AW)^{k+1}X_1WA \\
&= (X_2W)^{k+2}(AW)^{k}AWX_1WA \\
&= (X_2W)^{k+2}(AW)^{k}A \\
&= (X_2W)^{k+1}X_2W(AW)^{k}A \\
&= (X_2W)^{k+1}X_2(WA)^{k+1} \\
&= (X_2W)^{k}X_2WX_2WA(WA)^{k} \\
&= (X_2W)^{k}X_2WAWX_2(WA)^{k} \\
&= (X_2W)^{k}X_2(WA)^{k} \\
&= \ldots
\end{aligned}$$

$$\begin{aligned}
&= X_2WX_2WA \\
&= X_2WAWX_2 \\
&= X_2.
\end{aligned}$$

■

Teorema 1.52. *[29] Se $A \in \mathbb{C}^{m\times n}$, $W \in \mathbb{C}^{n\times m}$ e Ind$(WA) = k$, então*

$$(AW)_d = A(WA)_d^2W, \quad e \quad Ind(AW) \leq k+1.$$

Demonstração:

Defina $X = A(WA)_d^2W$. Como $\mathrm{Ind}(WA) = k$, segue da definição de inversa de Drazin que

$$\begin{aligned}
(AW)^{k+2}X &= (AW)^{k+2}A(WA)_d^2W \\
&= AWA(WA)^{k+1}(WA)_d(WA)_dW \\
&= AWA(WA)^k(WA)_dW \\
&= A(WA)^{k+1}(WA)_dW \\
&= A(WA)^kW \\
&= (AW)^{k+1},
\end{aligned} \tag{1.55}$$

além disso, da definição de inversa de Drazin temos que

$$(WA)_d^2(WA) = (WA)_d,$$

de onde segue que

$$\begin{aligned}
X^2(AW) &= (A(WA)_d^2W)^2AW \\
&= A(WA)_d^2WA(WA)_d^2WAW \\
&= A(WA)_d(WA)_dW \\
&= A(WA)_d^2W \\
&= X,
\end{aligned} \tag{1.56}$$

e

$$\begin{aligned}
X(AW) &= A(WA)_d^2WAW \\
&= A(WA)_d(WA)_d(WA)W \\
&= A(WA)_d(WA)(WA)_dW \\
&= A(WA)(WA)_d(WA)_dW \\
&= AWA(WA)_d^2W \\
&= (AW)X.
\end{aligned} \tag{1.57}$$

Das equações (1.55), (1.56) e (1.57), segue que $(AW)_d = X$ e $\mathrm{Ind}(AW) \leq k+1$. ■

Corolário 1.53. *Sob as hipóteses do teorema anterior, vale*

$$W(AW)_d^p = (WA)_d^pW, \quad e \quad A(WA)_d^p = (AW)_d^pA,$$

para p inteiro positivo.

Demonstração:

Considere $p = 1$, pelo teorema anterior temos

$$W(AW)_d = WA(WA)_d^2 W = (WA)_d W,$$

e

$$A(WA)_d = A(WA)_d^2 WA = (AW)_d A.$$

Portanto o resultado vale para $p = 1$. Suponha que as afirmações valem para todo inteiro positivo menor que p. Temos que

$$\begin{aligned} W(AW)_d^p &= W(AW)_d^{p-1}(AW)_d \\ &= (WA)_d^{p-1} W(AW)_d \\ &= (WA)_d^{p-1}(WA)_d W \\ &= (WA)_d^p W, \end{aligned}$$

e

$$\begin{aligned} A(WA)_d^p &= A(WA)_d^{p-1}(WA)_d \\ &= (AW)_d^{p-1} A(WA)_d \\ &= (AW)_d^{p-1}(AW)_d A \\ &= (AW)_d^p A. \end{aligned}$$

■

O próximo teorema nos mostra que dadas as matrizes A e W, a inversa de Drazin com peso $A_{d,W}$ existe e pode ser escrita em função da inversa de Drazin de AW ou de WA.

Teorema 1.54. *[29] Se $A \in \mathbb{C}^{m\times n}$, $X \in \mathbb{C}^{n\times m}$ e Ind$(AW) = k$, então*

$$A_{d,W} = A(WA)_d^2 = (AW)_d^2 A.$$

Demonstração:

Vejamos primeiro que $A_{d,W} = A(WA)_d^2$, para isso defina $X = A(WA)_d^2$, vejamos que X satisfaz as equações (1.52) - (1.54). Usando a definição de inversa de Drazin e o corolário 1.53 temos

$$\begin{aligned}
(AW)^{k+1}XW &= (AW)^{k+1}A(WA)_d^2W \\
&= A(WA)^k(WA)(WA)_d^2W \\
&= A(WA)^k(WA)_dW \\
&= A(WA)^kW(WA)_d \\
&= (AW)^{k+1}(WA)_d \\
&= (AW)^k;
\end{aligned}$$

$$\begin{aligned}
XWAWX &= A(WA)_d^2WAWA(WA)_d^2 \\
&= A(WA)_d^2(WA)(WA)(WA)_d^2 \\
&= A(WA)_d(WA)_d \\
&= A(WA)_d^2 \\
&= X;
\end{aligned}$$

$$\begin{aligned}
AWX &= AWA(WA)_d^2 \\
&= A(WA)(WA)_d(WA)_d \\
&= A(WA)_d(WA)(WA)_d \\
&= A(WA)_d(WA)_d(WA) \\
&= A(WA)_d^2WA \\
&= XWA.
\end{aligned}$$

De onde segue que $A_{d,W} = A(WA)_d^2$.

Para mostrar que $A_{d,W} = (AW)_d^2A$, procedemos da mesma maneira, ou seja, vamos mostrar que as equações (1.52) - (1.54) valem para $X = (AW)_d^2A$.

$$\begin{aligned}
(AW)^{k+1}XW &= (AW)^{k+1}(AW)_d^2AW \\
&= (AW)^{k+1}A(WA)_d^2W \\
&= (AW)^{k+1}AW(AW)_d^2 \\
&= (AW)^{k+1}(AW)_d \\
&= (AW)^k;
\end{aligned}$$

$$\begin{aligned} XWAWX &= (AW)_d^2AWAW(AW)_d^2A \\ &= (AW)_d^2(AW)(AW)(AW)_d^2A \\ &= (AW)_d(AW)_dA \\ &= (AW)_d^2A \\ &= X; \end{aligned}$$

$$\begin{aligned} AWX &= AW(AW)_d^2A \\ &= (AW)(AW)_d(AW)_dA \\ &= (AW)_d(AW)(AW)_dA \\ &= (AW)_d(AW)_d(AW)A \\ &= (AW)_d^2AWA \\ &= XWA. \end{aligned}$$

Portanto, $A_{d,W} = (AW)_d^2A$. ∎

Por fim, vejamos um resultado que permite verificar a existência de uma matriz peso W que vincule duas matrizes A e X.

Teorema 1.55. *[29] Dadas as matrizes $A, X \in \mathbb{C}^{m\times n}$, existe uma matriz $W \in \mathbb{C}^{n\times m}$ tal que $X = A_{d,W}$ se e somente se a matriz X pode ser decomposta da forma*

$$X = AYAYA,$$

com $Y \in \mathbb{C}^{n\times m}$ satisfazendo Ind(AY) =Ind$(YA) = 1$.

Demonstração:

$(\Rightarrow)$:

Supondo que $X = A_{d,W}$ para alguma matriz $W \in \mathbb{C}^{n\times m}$, então pelo Teorema 1.54 e Corolário 1.53 temos

$$\begin{aligned} X &= A(WA)_d^2 \\ &= A(WA)_d(WA)_d \\ &= (AW)_dA(WA)_d \\ &= (AW)(AW)_d^2A(WA)_d^2(WA) \\ &= A(WA)_d^2WA(WA)_d^2WA \\ &= AYAYA, \end{aligned}$$

onde $Y = (WA)_d^2W$.

Agora note que dada uma matriz quadrada B, então $BB_dB = B$ se e somente se

$\text{Ind}(B) = 1$. De fato, se $\text{Ind}(B) = 1$ segue da definição de inversa de Drazin que $BB_dB = B$, reciprocamente se $BB_dB = B$ da definição de inversa de Drazin temos também que $B_dBB_d = B_d$ e $BB_d = B_dB$, logo $B_d = B_g$, e pelo Teorema 1.42 segue que $\text{Ind}(B) = 1$. Neste caso, temos que $(B_d)_d = B$.

Assim, pelo fato de que

$$AY = A(WA)_d^2W = (AW)_d^2AW = (AW)_d$$

e

$$YA = (WA)_d^2WA = (WA)_d,$$

temos que

$$\text{Ind}(AY) = \text{Ind}((AW)_d) = 1$$

e

$$\text{Ind}(YA) = \text{Ind}((WA)_d) = 1.$$

$(\Leftarrow)$:

Supondo que $X = AYAYA$, defina $W = Y(AY)_d^2$, então

$$WA = Y(AY)_d^2A = (YA)_d^2YA = (YA)_d,$$

e como $\text{Ind}(YA) = 1$, então

$$\begin{aligned} WX &= WAYAYA \\ &= (YA)_d(YA)(YA) \\ &= (YA)(YA)_d(YA) \\ &= YA \\ &= ((YA)_d)_d \\ &= (WA)_d. \end{aligned}$$

Similarmente, como $\text{Ind}(AY) = 1$ e $X = AYAYA$ temos

$$AW = AY(AY)_d^2 = (AY)_d,$$

e assim

$$\begin{aligned}
XW &= AYAYAW \\
&= (AY)(AY)(AY)_d \\
&= (AY)(AY)_d(AY) \\
&= AY \\
&= ((AY)_d)_d \\
&= (AW)_d.
\end{aligned}$$

Com base nisso, podemos mostrar que X satisfaz as equações (1.52) - (1.54).

$$\begin{aligned}
(AW)^{k+1}XW &= (AW)^{k+1}(AW)_d = (AW)^k; \\
XWAWX &= (AW)_dA(WA)_d = (AW)_d(AW)_dA = AYAYA = X; \\
AWX &= A(WA)_d = (AW)_dA = XWA.
\end{aligned}$$

Portanto, $X = A_{d,W}$. ∎

Capítulo 2

Métodos computacionais para inversas generalizadas

No capítulo anterior definimos algumas inversas generalizadas e vimos também aplicações dessas inversas em diversas áreas, porém dada uma matriz qualquer, como encontramos uma inversa generalizada dessa matriz?

Neste capítulo veremos os principais métodos computacionais existentes na literatura para computar algumas inversas generalizadas. Basicamente, os métodos para computar inversas generalizadas são divididos em duas classes: métodos diretos e métodos iterativos.

Os métodos diretos são algoritmos que encontram a inversa generalizada de uma matriz em uma quantidade finita de operações aritméticas, considerando que estamos trabalhando com aritmética exata. Os métodos diretos são geralmente baseados em decomposições de matrizes, e o seu desempenho pode variar bastante dependendo do tipo de decomposição usada e da maneira como é efetuada a decomposição.

Os métodos iterativos são algoritmos que partem de uma aproximação inicial X_0, e constrõem uma sequência de matrizes $\{X_1, X_2, X_3 \ldots\}$ de tal forma que no limite a sequência converge para uma matriz X que é a inversa generalizada que procuramos. Tais métodos são da forma

$$X_{k+1} = f(X_0, X_1, \ldots, X_k), \quad k = 0, 1, 2, \ldots,$$

onde f é alguma função. Um método iterativo completo deve conter informações de como selecionar a aproximação inicial X_0, como construir X_{k+1} a partir de $X_0, X_1, \ldots, X_k$ e quando parar de iterar, ou seja, um critério de avaliação para decidir se a aproximação construída X_k é boa o suficiente.

Grande parte dos métodos diretos para encontrar inversas generalizadas são baseados em decomposições de matrizes, por isso, veremos algumas decomposições que serão fundamentais para o estudo dos métodos.

Definição 2.1. Sejam $A \in \mathbb{C}^{m\times n}$, $u \in \mathbb{C}^m$, $v \in \mathbb{C}^n$ e $\sigma > 0$ tais que

$$Av = \sigma u \quad \text{e} \quad A^*u = \sigma v.$$

Neste caso, σ é chamado valor singular de A, enquanto u e v são chamados vetores singulares a esquerda e a direita de A respectivamente.

Note que

$$A^*Av = A^*\sigma u = \sigma^2 v \quad \text{e} \quad AA^*u = A\sigma v = \sigma^2 u,$$

ou seja, σ^2 é autovalor de A^*A e de AA^*.

Teorema 2.2. *[12] Se $A \in \mathbb{C}^{m\times n}$ com posto$(A) = r$, então existem matrizes unitárias $U \in \mathbb{C}^{m\times m}$ e $V \in \mathbb{C}^{n\times n}$ tais que*

$$A = U \begin{pmatrix} \Sigma & 0 \\ 0 & 0 \end{pmatrix} V^*, \tag{2.1}$$

sendo $\Sigma = diag(\sigma_1, \sigma_2, \ldots, \sigma_r)$, com $\sigma_i = \sqrt{\lambda_i}$ e $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_r > 0$ os autovalores não nulos de A^*A. Então $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_r > 0$ são os valores singulares de A, e a decomposição (2.1) é chamada de decomposição em valores singulares de A, comumente referida como SVD (Singular Value Decomposition).

Demonstração:

Ver Teorema 2.5.2 de [12]. ∎

Teorema 2.3. *[12] Se $A \in \mathbb{C}^{m\times n}$, com $m \geq n$, então existem $Q \in \mathbb{C}^{m\times n}$ e $R \in \mathbb{C}^{n\times n}$, sendo Q uma matriz com colunas ortonormais e R triangular superior de forma que*

$$A = QR.$$

Demonstração:

Ver capítulo 5.2 de [12]. ∎

A decomposição QR, como mostrada acima, permite encontrar a inversa de Moore-Penrose apenas de matrizes de posto completo, porém uma adaptação na decomposição QR nos permite ampliar o processo para matrizes com posto incompleto.

Teorema 2.4. *[12] Se $A \in \mathbb{C}^{m\times n}$, com $m \geq n$ e posto$(A) = k$, então existem $Q \in \mathbb{C}^{m\times n}$, $R \in \mathbb{C}^{n\times n}$ e $E \in \mathbb{C}^{n\times n}$, sendo Q uma matriz com colunas ortonormais, E uma matriz de permutação e R da forma*

$$R = \begin{pmatrix} R_1 & R_2 \\ 0 & 0 \end{pmatrix},$$

sendo $R_1 \in \mathbb{C}^{k\times k}$ triangular superior e não singular, tais que

$$AE = QR.$$

Demonstração:

Ver capítulo 5.2 de [12]. ■

Definição 2.5. Seja $A \in \mathbb{C}^{m\times n}$ com posto$(A) = k$. $A = FG$ é uma decomposição de posto completo de A se $F \in \mathbb{C}^{m\times k}$, $G \in \mathbb{C}^{k\times n}$ e posto(F) =posto$(G) = k$.

Uma decomposição de posto completo de uma matriz A pode ser feita de várias maneiras diferentes, como por exemplo usando eliminação de Gauss ou transformações de Householder.

As decomposições acima são bastante usadas na obtenção de inversas generalizadas, a seguir veremos métodos computacionais para calcular tipos específicos de inversas generalizadas.

2.1 Cálculo de inversas-$\{i, j, k\}$

2.1.1 Cálculo de inversa-$\{1\}$

Definição 2.6. Seja $A \in \mathbb{C}^{m\times n}$ com posto$(A) = r$. Se A é da forma

$$A = \begin{pmatrix} C \\ 0 \end{pmatrix},$$

onde 0 é uma matriz nula de ordem $(m-r)\times n$, e $C \in \mathbb{C}^{r\times n}$ satisfaz as seguintes condições

1. $c_{ij} = 0$ quando $i > j$;
2. A primeira entrada não nula em cada linha de C é 1;
3. Se $c_{ij} = 1$ é a primeira entrada não nula da linha i de C, então a coluna j possui todos os elementos nulos, exceto c_{ij}.

Então A está na forma normal de Hermite.

A forma normal de Hermite pode ser obtida usando eliminação de Gauss. Note que se A está na forma normal de Hermite, então existe uma matriz de permutação P tal que

$$AP = \begin{pmatrix} I_r & K \\ 0 & 0 \end{pmatrix},$$

para alguma matriz $K \in \mathbb{C}^{r\times(n-r)}$.

Sendo assim, dada uma matriz $A \in \mathbb{C}^{m\times n}$ com posto$(A) = r$, é possível escrever

$$EAP = \begin{pmatrix} I_r & K \\ 0 & 0 \end{pmatrix},$$

onde E é uma matriz não singular obtida pela eliminação Gaussiana e P é uma matriz de permutação.

Teorema 2.7. *Seja $A \in \mathbb{C}^{m\times n}$ com posto$(A) = r$. Se $E \in \mathbb{C}^{m\times m}$ é não singular e $P \in \mathbb{C}^{n\times n}$ é uma matriz de permutação tal que*

$$EAP = \begin{pmatrix} I_r & K \\ 0 & 0 \end{pmatrix},$$

então

$$X = P\begin{pmatrix} I_r & 0 \\ 0 & L \end{pmatrix}E$$

é uma inversa-$\{1\}$ de A, sendo $L \in \mathbb{C}^{(n-r)\times(m-r)}$ uma matriz arbitrária.

Demonstração:

Vejamos que X satisfaz $AXA = A$,

$$\begin{aligned} AXA &= E^{-1}\begin{pmatrix} I_r & K \\ 0 & 0 \end{pmatrix}P^*P\begin{pmatrix} I_r & 0 \\ 0 & L \end{pmatrix}EE^{-1}\begin{pmatrix} I_r & K \\ 0 & 0 \end{pmatrix}P^* \\ &= E^{-1}\begin{pmatrix} I_r & K \\ 0 & 0 \end{pmatrix}\begin{pmatrix} I_r & 0 \\ 0 & L \end{pmatrix}\begin{pmatrix} I_r & K \\ 0 & 0 \end{pmatrix}P^* \\ &= E^{-1}\begin{pmatrix} I_r & KL \\ 0 & 0 \end{pmatrix}\begin{pmatrix} I_r & K \\ 0 & 0 \end{pmatrix}P^* \\ &= E^{-1}\begin{pmatrix} I_r & K \\ 0 & 0 \end{pmatrix}P^* \\ &= A. \end{aligned}$$

■

2.1.2 Cálculo de inversa-$\{1, 2\}$

Lema 2.8. *[5] Seja $A \in \mathbb{C}^{m\times n}$ uma matriz particionada da forma*

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix},$$

onde $A_{11} \in \mathbb{C}^{r\times r}$ *é não singular. Temos que posto*$(A) = r$ *se e somente se* $A_{22} = A_{21}A_{11}^{-1}A_{12}$.

Demonstração:

$(\Rightarrow)$:

Como posto$(A) = r$ e A_{11} é não singular, então as $m - r$ últimas linhas de A são combinações lineares das r primeiras, logo existe uma matriz C tal que

$$C(A_{11}, A_{12}) = (A_{21}, A_{22}),$$

ou seja, $CA_{11} = A_{21}$ e $CA_{12} = A_{22}$, assim temos que

$$C = A_{21}A_{11}^{-1}$$

e portanto

$$A_{22} = CA_{12} = A_{21}A_{11}^{-1}A_{12}.$$

$(\Leftarrow)$:

Como A_{11} é não singular, então

$$\text{posto}(A) \geq \text{posto}(A_{11}) = r. \tag{2.2}$$

Além disso, como $A_{22} = A_{21}A_{11}^{-1}A_{12}$, definindo $C = A_{21}A_{11}^{-1}$ temos

$$CA_{11} = A_{21} \quad \text{e} \quad A_{22} = CA_{12},$$

assim

$$C(A_{11}, A_{12}) = (A_{21}, A_{22}),$$

de onde segue que as $m - r$ últimas linhas de A são combinações lineares das primeiras r, assim

$$\text{posto}(A) \leq r. \tag{2.3}$$

De (2.2) e (2.3) segue que posto$(A) = r$. ∎

Teorema 2.9. *[5] Seja* $A \in \mathbb{C}^{m\times n}$ *com posto*$(A) = r$. *Se* $P \in \mathbb{C}^{m\times m}$ *e* $Q \in \mathbb{C}^{n\times n}$ *são matrizes de permutação tal que*

$$PAQ = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix},$$

onde $A_{11} \in \mathbb{C}^{r\times r}$ *é não singular, então*

$$X = Q \begin{pmatrix} A_{11}^{-1} & 0 \\ 0 & 0 \end{pmatrix} P$$

é uma inversa-$\{1,2\}$ *de* A.

Demonstração:

Basta verificar que X satisfaz $AXA = A$ e $XAX = X$. Usando o Lema 2.8 temos

$$\begin{aligned} AXA &= P^* \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} Q^*Q \begin{pmatrix} A_{11}^{-1} & 0 \\ 0 & 0 \end{pmatrix} PP^* \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} Q^* \\ &= P^* \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} A_{11}^{-1} & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} Q^* \\ &= P^* \begin{pmatrix} I & 0 \\ A_{21}A_{11}^{-1} & 0 \end{pmatrix} \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} Q^* \\ &= P^* \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{21}A_{11}^{-1}A_{12} \end{pmatrix} Q^* = P^* \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} Q^* = A \end{aligned}$$

e

$$\begin{aligned} XAX &= Q \begin{pmatrix} A_{11}^{-1} & 0 \\ 0 & 0 \end{pmatrix} PP^* \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} Q^*Q \begin{pmatrix} A_{11}^{-1} & 0 \\ 0 & 0 \end{pmatrix} P \\ &= Q \begin{pmatrix} A_{11}^{-1} & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} A_{11}^{-1} & 0 \\ 0 & 0 \end{pmatrix} P \\ &= Q \begin{pmatrix} I & A_{11}^{-1}A_{12} \\ 0 & 0 \end{pmatrix} \begin{pmatrix} A_{11}^{-1} & 0 \\ 0 & 0 \end{pmatrix} P \\ &= Q \begin{pmatrix} A_{11}^{-1} & 0 \\ 0 & 0 \end{pmatrix} P \\ &= X. \end{aligned}$$

■

Teorema 2.10. *[5] Seja* $A \in \mathbb{C}^{m\times n}$ *com posto*$(A) = r$. *Se* $E \in \mathbb{C}^{m\times m}$ *é não singular e* $P \in \mathbb{C}^{n\times n}$ *é uma matriz de permutação tal que* $EAP = \begin{pmatrix} I_r & K \\ 0 & 0 \end{pmatrix}$, *então*

$$X = P \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} E,$$

é uma inversa-$\{1,2\}$ *de* A.

Demonstração:

Segue do Teorema 2.7 que X satisfaz $AXA = A$, restando apenas mostrar que X satisfaz $XAX = X$. De fato,

$$\begin{aligned} XAX &= P\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} EE^{-1}\begin{pmatrix} I_r & K \\ 0 & 0 \end{pmatrix} P^*P\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} E \\ &= P\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} I_r & K \\ 0 & 0 \end{pmatrix}\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} E \\ &= P\begin{pmatrix} I_r & K \\ 0 & 0 \end{pmatrix}\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} E \\ &= P\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} E \\ &= X. \end{aligned}$$

■

2.1.3 Cálculo de inversa-$\{1, 2, 3\}$

Lema 2.11. *[5] Se $A \in \mathbb{C}^{m\times n}$ e $A = FG$ é uma decomposição de posto completo de A, então*

$$G^{(1)}F^{(j)}$$

é uma inversa-$\{j\}$ de A, onde $A^{(j)}$ denota uma inversa-$\{j\}$ de A para $j = 1, 2, 3, 4$.

Demonstração:

Vimos na demonstração do Lema 1.17 que

$$F^{(1)}F = GG^{(1)} = I,$$

assim, para $j = 1$ defina $X_1 = G^{(1)}F^{(1)}$, e temos que

$$AX_1A = FGG^{(1)}F^{(1)}FG = FG = A.$$

Para $j = 2$, defina $X_2 = G^{(1)}F^{(2)}$ então

$$X_2AX_2 = G^{(1)}F^{(2)}FGG^{(1)}F^{(2)} = G^{(1)}F^{(2)}FF^{(2)} = G^{(1)}F^{(2)} = X_2.$$

E para $j = 3$, defina $X_3 = G^{(1)}F^{(3)}$, então

$$(AX_3)^* = (FGG^{(1)}F^{(3)})^* = (FF^{(3)})^* = FF^{(3)} = FGG^{(1)}F^{(3)} = AX_3.$$

■

Teorema 2.12. *[5] Se $A \in \mathbb{C}^{m\times n}$ e $A = FG$ é uma decomposição de posto completo de A, então*

$$X = G^{(1)}F^{\dagger}$$

é uma inversa-$\{1,2,3\}$ de A.

Demonstração:

Como $F^{\dagger}$ é uma inversa-$\{1,2,3\}$ de F, segue do Lema 2.11 que $X = G^{(1)}F^{\dagger}$ satisfaz as equações

$$AXA = A, \quad XAX = X, \quad (AX)^* = AX,$$

portanto $G^{(1)}F^{\dagger}$ é uma inversa-$\{1,2,3\}$ de A. ■

2.2 Cálculo da inversa-$\{2\}$ com imagem e núcleo prescritos $A^{(2)}_{T,S}$

No capítulo anterior vimos que a inversa $A^{(2)}_{T,S}$ existe sob certas condições dos espaços T e S, e nesse caso é única. Além disso, vimos algumas inversas generalizadas que são únicas, e possuem imagem e núcleo determinados, com isso, podemos escrever algumas inversas generalizadas como inversa $A^{(2)}_{T,S}$, de onde segue a importância de calcular uma inversa $A^{(2)}_{T,S}$.

Segue do capítulo anterior que

$$\begin{aligned} A^{\dagger} &= A^{(2)}_{\mathcal{R}(A^*),\mathcal{N}(A^*)}; \\ A^{\dagger}_{M,N} &= A^{(2)}_{\mathcal{R}(N^{-1}A^*M),\mathcal{N}(N^{-1}A^*M)}; \\ A_d &= A^{(2)}_{\mathcal{R}(A^k),\mathcal{N}(A^k)}, \quad \text{onde } \mathrm{Ind}(A) = k. \end{aligned}$$

2.2.1 Inversa $A^{(2)}_{T,S}$ via decomposição de posto completo

Seja $A \in \mathbb{C}^{m\times n}$ com posto$(A) = r$. Considere dois subespaços $T \subset \mathbb{C}^n$ e $S \subset \mathbb{C}^m$, com dim(T) =dim$(S^{\perp}) = t < r$, e seja $B \in \mathbb{C}^{m\times n}$ uma matriz tal que

$$\mathcal{R}(B) = T \quad \text{e} \quad \mathcal{N}(B) = S.$$

Considere

$$B = FG$$

uma decomposição de posto completo de B, neste caso temos que

$$T = \mathcal{R}(B) = \mathcal{R}(F) \quad \text{e} \quad S = \mathcal{N}(B) = \mathcal{N}(G).$$

De fato, se $y \in \mathcal{R}(B)$ então existe x tal que $Bx = y$, logo $FGx = y$ e portanto $y \in \mathcal{R}(F)$, e como

$$\dim(\mathcal{R}(B)) = \text{posto}(B) = \text{posto}(F) = \dim(\mathcal{R}(F)),$$

então $\mathcal{R}(B) = \mathcal{R}(F)$. Agora tome $x \in \mathcal{N}(G)$, então $Gx = 0$, logo

$$Bx = FGx = 0,$$

e $x \in \mathcal{N}(B)$, portanto $\mathcal{N}(G) \subset \mathcal{N}(B)$, e como

$$\dim(\mathcal{N}(B)) = n - \text{posto}(B) = n - \text{posto}(G) = \dim(\mathcal{N}(G)),$$

segue que $\mathcal{N}(B) = \mathcal{N}(G)$.

Assim, do Teorema 1.29 temos que

$$A^{(2)}_{T,S} = F(GAF)^{-1}G.$$

2.3 Cálculo da inversa de Moore-Penrose

Vejamos inicialmente os principais métodos diretos para computar a inversa de Moore-Penrose.

2.3.1 Inversa de Moore-Penrose via decomposição de posto completo

Vimos que

$$A^{\dagger} = A^{(2)}_{\mathcal{R}(A^*),\mathcal{N}(A^*)},$$

e vimos também que a inversa $A^{(2)}_{T,S}$ pode ser calculada via decomposição de posto completo assim se $G^* \in \mathbb{C}^{n\times r}$ e $F \in \mathbb{C}^{m\times r}$ são matrizes cujas colunas formam uma base para $\mathcal{R}(A^*)$ e $\mathcal{N}(A^*)^{\perp}$ respectivamente, então

$$A^{\dagger} = G^*(F^*AG^*)^{-1}F^*.$$

Agora, precisamos encontrar matrizes F e G nas condições acima. Veja que se

$$A = FG$$

é uma decomposição de posto completo de A, então

$$A^* = G^*F^*$$

é uma decomposição de posto completo de A^*. Além disso, $\mathcal{R}(A^*) = \mathcal{R}(G^*)$ e $\mathcal{N}(A^*) = \mathcal{N}(F^*)$.

Sendo assim, se $A = FG$ é uma decomposição de posto completo de A, então

$$A^\dagger = G^*(F^*AG^*)^{-1}F^* = G^*(F^*FGG^*)^{-1}F^* = G^*(GG^*)^{-1}(F^*F)^{-1}F^*.$$

Uma decomposição de posto completo pode ser feita de diversas maneiras, para cada método usado para encontrar F e G temos um método diferente para encontrar $A^\dagger$. Além disso, o problema de encontrar a inversa de Moore-Penrose de A foi trocado por um problema que necessita encontrar a inversa usual de matrizes, o que também pode ser executado de diferentes maneiras.

2.3.2 Inversa de Moore-Penrose via SVD

O teorema a seguir mostra que podemos obter a inversa de Moore-Penrose a partir da decomposição em valores singulares.

Teorema 2.13. *[29] Se $A \in \mathbb{C}^{m\times n}$, com decomposição em valores singulares*

$$A = U \begin{pmatrix} \Sigma & 0 \\ 0 & 0 \end{pmatrix} V^*,$$

então

$$A^\dagger = V \begin{pmatrix} \Sigma^{-1} & 0 \\ 0 & 0 \end{pmatrix} U^*.$$

O cálculo de uma decomposição SVD pode ser feito de diversas maneiras, porém necessita do cálculo de autovalores, o que deixa o método acima pouco interessante.

2.3.3 Inversa de Moore-Penrose via decomposição QR

A decomposição QR também pode ser usada para encontrar a inversa de Moore-Penrose, mas não de maneira tão direta quanto a decomposição SVD, pelo menos não para matrizes que não possuem posto completo, nesse caso é necessário usar duas decomposições QR.

Vimos que a inversa de Moore-Penrose de um produto de matrizes nem sempre é o produto das inversas de Moore-Penrose dessas matrizes. Richard Bouldin em [6] mostrou uma relação entre essa propriedade e a comutação de algumas matrizes, conforme a proposição a seguir.

Proposição 2.14. *([6])*

Sejam $A \in \mathbb{C}^{m\times n}$ e $B \in \mathbb{C}^{n\times p}$. Temos que

$$(AB)^\dagger = B^\dagger A^\dagger$$

se e somente se

$$A^\dagger ABB^* = BB^*A^\dagger A \quad e \quad BB^\dagger A^*A = A^*ABB^\dagger.$$

Teorema 2.15. *Se $A \in \mathbb{C}^{m\times n}$ com $m \geq n$, então*

$$A^\dagger = R^\dagger Q^*.$$

Demonstração:

Vejamos que $Q^\dagger = Q^*$. De fato, basta provar Q^* satisfaz as quatro equações de Penrose.

$$\begin{aligned} QQ^*Q &= QI_n = Q; \\ Q^*QQ^* &= I_nQ^* = Q^*; \\ (QQ^*)^* &= (Q^*)^*Q^* = QQ^*; \\ (Q^*Q)^* &= Q^*(Q^*)^* = Q^*Q. \end{aligned}$$

Agora, note que

$$Q^\dagger QRR^* = Q^*QRR^* = I_nRR^* = RR^*I_n = RR^*Q^*Q = RR^*Q^\dagger Q$$

e

$$RR^\dagger Q^*Q = RR^\dagger I_n = I_nRR^\dagger = Q^*QRR^\dagger,$$

portanto segue da proposição acima que

$$A^\dagger = R^\dagger Q^\dagger = R^\dagger Q^*.$$

■

Corolário 2.16. *Se $A \in \mathbb{C}^{m\times n}$ com $m \geq n$ e posto$(A) = n$, então*

$$A^\dagger = R^{-1}Q^*.$$

Demonstração:

Segue do teorema anterior notando que posto$(A) = n$ implica em posto$(R) = n$, ou seja, R é não singular e consequentemente $R^\dagger = R^{-1}$. ■

Note que nos resultados acima consideramos $m \geq n$, porém para o caso em que $m < n$, basta calcular $(A^*)^\dagger$, e pelo item 3 do Teorema 1.2 temos

$$A^\dagger = ((A^*)^\dagger)^*.$$

O processo acima permite encontrar $A^\dagger$ de fato apenas para matrizes com posto completo, caso contrário não teremos R não singular, e o problema de encontrar $A^\dagger$ é apenas substituído pelo problema de encontrar $R^\dagger$. Em princípio parece que não houve melhora, porém se considerarmos uma permutação nas colunas da matriz A podemos transformar o problema de encontrar $R^\dagger$ em um problema de encontrar a inversa de Moore-Penrose de uma matriz de posto completo.

Teorema 2.17. *Seja $A \in \mathbb{C}^{m\times n}$, com $m \geq n$ e posto$(A) = r < n$. Se*

$$AE = QR$$

é a decomposição QR de AE, onde E é uma matriz de permutação, $Q \in C^{m\times n}$ satisfaz $Q^*Q = I_n$ e

$$R = \begin{pmatrix} R_1 \\ 0 \end{pmatrix},$$

sendo $R_1 \in \mathbb{C}^{r\times n}$, então

$$A^\dagger = E(\widehat{R}^{-1}\widehat{Q}^*, 0)Q^*,$$

onde $R_1 = \widehat{Q}\widehat{R}$ é a decomposição QR de R_1.

Demonstração:

Como E é uma matriz de permutação, então E é unitária, e pelo item 8 do Teorema 1.2 temos que $(AE)^\dagger = E^*A^\dagger$, de onde segue que

$$A^\dagger = E(AE)^\dagger,$$

e pelo Teorema 2.15 temos que

$$A^\dagger = ER^\dagger Q^*.$$

Vejamos que $R^\dagger = (R_1^\dagger, 0)$. De fato, como $R = \begin{pmatrix} R_1 \\ 0 \end{pmatrix}$, temos que $(R_1^\dagger, 0)$ satisfaz as equações de Penrose.

$$\begin{aligned}
\begin{pmatrix} R_1 \\ 0 \end{pmatrix}(R_1^\dagger, 0)\begin{pmatrix} R_1 \\ 0 \end{pmatrix} &= \begin{pmatrix} R_1R_1^\dagger & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} R_1 \\ 0 \end{pmatrix} = \begin{pmatrix} R_1R_1^\dagger R_1 \\ 0 \end{pmatrix} = \begin{pmatrix} R_1 \\ 0 \end{pmatrix}; \\
(R_1^\dagger, 0)\begin{pmatrix} R_1 \\ 0 \end{pmatrix}(R_1^\dagger, 0) &= R_1^\dagger R_1(R_1^\dagger, 0) = (R_1^\dagger R_1 R_1^\dagger, 0) = (R_1^\dagger, 0); \\
\left[\begin{pmatrix} R_1 \\ 0 \end{pmatrix}(R_1^\dagger, 0)\right]^* &= \left[\begin{pmatrix} R_1R_1^\dagger & 0 \\ 0 & 0 \end{pmatrix}\right]^* = \begin{pmatrix} (R_1R_1^\dagger)^* & 0 \\ 0 & 0 \end{pmatrix} \\
&= \begin{pmatrix} R_1R_1^\dagger & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} R_1 \\ 0 \end{pmatrix}(R_1^\dagger, 0); \\
\left[(R_1^\dagger, 0)\begin{pmatrix} R_1 \\ 0 \end{pmatrix}\right]^* &= (R_1^\dagger R_1)^* = R_1^\dagger R_1 = (R_1^\dagger, 0)\begin{pmatrix} R_1 \\ 0 \end{pmatrix}.
\end{aligned}$$

Assim, como posto(A) =posto(R) =posto$(R_1) = r$, então R_1 é uma matriz de posto completo, e pelo Corolário 2.16 temos que $R_1^\dagger = \widehat{R}^{-1}\widehat{Q}^*$, de onde segue que

$$A^\dagger = ER^\dagger Q^* = E(R_1^\dagger, 0)Q^* = E(\widehat{R}^{-1}\widehat{Q}^*, 0)Q^*.$$

■

2.3.4 Inversa de Moore-Penrose via Fatoração de Cholesky

Em 2005, Courrieu em [10] mostrou uma maneira de calcular a inversa de Moore-Penrose diretamente pela decomposição de Cholesky. O próximo teorema garante a existência da decomposição de Cholesky para matriz simétrica e positiva semidefinida.

Teorema 2.18. *[15] Se $B \in \mathbb{C}^{n\times n}$ é uma matriz simétrica e positiva semidefinida de posto r, então*

1. *Existe uma matriz R triangular superior, com elementos diagonais não negativos, tal que $B = R^*R$;*

2. *Existe uma matriz de permutação P tal que P^*BP tem uma única fatoração de Cholesky, a qual toma a forma*

$$P^*BP = R^*R, \quad \text{com} \quad R = \begin{pmatrix} R_{11} & R_{12} \\ 0 & 0 \end{pmatrix},$$

onde R_{11} é uma matriz triangular superior $r \times r$ com elementos diagonais positivos.

Demonstração:

Ver Teorema 10.9 de [15] ■

Sendo assim, considerando a decomposição de Cholesky de acordo com o item 2 do Teorema 2.18 para $B = A^*A$, se $G = AP$ e $W = [R_{11}\ R_{12}]$, então para $L = W^*$ temos $G^*G = LL^*$. O próximo teorema mostra como calcular $G^\dagger$ em função de L.

Teorema 2.19. *[10] Considere G e L como definidas acima. Temos que*

$$G^\dagger = L(L^*L)^{-1}(L^*L)^{-1}L^*G^*.$$

Demonstração:

Do item 5 do Teorema 1.2 temos que $G^\dagger = (G^*G)^\dagger G^*$, e como $G^*G = LL^*$ temos

$$G^\dagger = (LL^*)^\dagger G^*.$$

Pelos itens 4 e 5 do Teorema 1.2 e pelo fato de que L tem posto completo, temos que

$$G^\dagger = (LL^*)^\dagger G^* = (L^*)^\dagger L^\dagger G^* = L(L^*L)^{-1}(L^*L)^{-1}L^*G^*.$$

∎

2.3.5 Inversa de Moore-Penrose via ortonormalização de Gram-Schmidt

Em 2009, Toutounian e Ataei em [28] mostraram que a inversa de Moore-Penrose pode ser calculada usando o processo de ortonormalização de Gram-Schmidt modificado com o pseudo produto interno $\langle x, y\rangle_{A^*A} := x^*A^*Ay$. Para isso, note que pelo item 5 do Teorema 1.2 temos que $A^\dagger = (A^*A)^\dagger A^*$, e caso A seja uma matriz de posto completo, então $A^\dagger = (A^*A)^{-1}A^*$. Assim, se considerarmos a base canônica do $\mathbb{R}^n$, dada por $\{e_1, e_2, \ldots, e_n\}$, para aplicar o processo de ortonormalização, obtemos um conjunto de vetores $\{z_1, z_2, \ldots, z_n\}$, e definindo $Z = [z_1, z_2, \ldots, z_n]$ e $D = \text{diagonal}(d_1, d_2, \ldots, d_n)$, com $d_i = z_i^*A^*Az_i$, temos que $(A^*A)^{-1} = ZD^{-1}Z^*$, e portanto $A^\dagger = ZD^{-1}Z^*A^*$.

Para o caso de posto incompleto, considerando a decomposição de Cholesky de acordo com o item 2 do Teorema 2.18, se $W = [R_{11}\ R_{12}]$, então $P^*A^*AP = W^*W$, de onde segue que

$$A^\dagger = PW^\dagger(W^\dagger)^*P^*A^*.$$

Sendo assim, o problema agora é calcular $W^\dagger$, o que é analisado no seguinte resultado

Proposição 2.20. *[28] Se $W \in \mathbb{C}^{r\times n}$ é particionada da forma $W = [U\quad V]$, onde $U \in \mathbb{C}^{r\times r}$ é não singular e $V \in \mathbb{C}^{r\times(n-r)}$, então*

$$W^\dagger = \begin{pmatrix} I - BKB^* \\ KB^* \end{pmatrix} U^{-1},$$

onde $B = U^{-1}V$ *e* $K = (I + B^*B)^{-1}$.

Com base nos resultados acima, o algoritmo para calcular a inversa de Moore-Penrose de uma matriz qualquer A, via Gram-Schmidt modificado é dado por

Algoritmo 1. Inversa de Moore-Penrose por Gram-Schmidt modificado

Dado $A \in \mathbb{C}^{m\times n}$, Considere $z_i^{(0)} = e_i$, $L = I$, $s = 0$, $r = 0$
Escolha uma tolerância ϵ
Para $j = 1$ até n faça:
  $d_j = \langle z_j^{(j-1)}, z_j^{(j-1)} \rangle_{A^*A}$
  Se $|d_j| < \epsilon$, então faça:
    $s = s + 1$, $\pi_2(s) = j$
  Se não, faça:
    $r = r + 1$, $\pi_1(r) = j$
    Para $i = j + 1$ até n faça:
      $l_{ij} = \langle z_j^{(j-1)}, z_i^{(j-1)} \rangle_{A^*A}/d_j$, $z_i^{(j)} = z_i^{(j-1)} - l_{ij} z_j^{(j-1)}$
    Fim do para
  Fim do se
Fim do para
Construa a matriz de permutação P baseada em π_1 e π_2
Permute as colunas das matrizes Z, D e L por $Z = P^*ZP$, $L = P^*LP$, $D = P^*DP$
Defina D_{11} e Z_{11} sendo as r primeiras linhas e colunas de D e Z respectivamente
Compute $U^{-1} = Z_{11} D_{11}^{-\frac{1}{2}}$
Compute $L_1 = LD^{\frac{1}{2}}$
Defina V sendo as r primeiras linhas e $n - r$ últimas colunas de L_1^*
Compute $B = U^{-1}V$
Compute $K = (I + B^*B)^{-1}$
Compute $W^\dagger = \begin{bmatrix} I - BKB^* \\ KB^* \end{bmatrix} U^{-1}$
Compute $A^\dagger = PW^\dagger (W^\dagger)^* P^* A^*$.

2.3.6 Inversa de Moore-Penrose via métodos de recursão com matrizes particionadas

Alguns métodos para encontrar a inversa de Moore-Penrose trabalham de forma recursiva, tais métodos se comportam de maneira parecida com métodos iterativos, porém fornecem a solução "exata" em uma quantidade finita de passos, por isso classificamos tais métodos como métodos diretos. Vejamos alguns exemplos de métodos de recursão com matrizes particionadas.

Teorema 2.21. *[29] **Método de Greville:***

Seja $A \in \mathbb{C}^{m\times n}$, da forma $A = (a_1, a_2, \ldots, a_n)$. Defina $A_1 = (a_1)$ e para $k = 2, 3, \ldots, n$ defina

$$A_k = (A_{k-1}, a_k).$$

Se

$$\begin{aligned} d_k &= A_{k-1}^{\dagger} a_k; \\ c_k &= a_k - A_{k-1} d_k = (I - A_{k-1} A_{k-1}^{\dagger}) a_k. \end{aligned}$$

então

$$A_k^{\dagger} = (A_{k-1}, a_k)^{\dagger} = \begin{pmatrix} A_{k-1}^{\dagger} - d_k b_k^* \\ b_k^* \end{pmatrix}, \quad k = 2, 3, \ldots, n,$$

onde

$$b_k^* = \begin{cases} c_k^{\dagger}, & se \quad c_k \neq 0, \\ (1 + d_k^* d_k)^{-1} d_k^* A_{k-1}^{\dagger} & se \quad c_k = 0. \end{cases}$$

Demonstração:

Ver Teorema 5.3.2 de [29]. ∎

Dessa forma, o método de Greville consiste em particionar a matriz A em colunas, encontrar a inversa de Moore-Penrose da primeira coluna, e a cada passo aumentar a matriz em uma coluna e calcular a inversa de Moore-Penrose da nova matriz em função da inversa de Moore-Penrose da matriz anterior. Note que apesar de se comportar como um método iterativo, na iteração n teremos a solução, ou seja,

$$A_n^{\dagger} = A^{\dagger}.$$

Teorema 2.22. *[29]([8]) **Método de Cline:***

Se $A = (U, V) \in \mathbb{C}^{m\times n}$, com $U \in \mathbb{C}^{m\times p} (p < n)$ e $V \in \mathbb{C}^{m\times (n-p)}$, então

$$A^{\dagger} = (U, V)^{\dagger} = \begin{pmatrix} U^{\dagger} - U^{\dagger} V C^{\dagger} - U^{\dagger} V (I - C^{\dagger} C) K_1^{-1} V^* (U^{\dagger})^* U^{\dagger} (I - V C^{\dagger}) \\ C^{\dagger} + (I - C^{\dagger} C) K_1^{-1} V^* (U^{\dagger})^* U^{\dagger} (I - V C^{\dagger}) \end{pmatrix},$$

onde

$$\begin{aligned} C &= (I - U U^{\dagger}) V, \\ K_1 &= I + (I - C^{\dagger} C) V^* (U^{\dagger})^* U^{\dagger} V (I - C^{\dagger} C). \end{aligned}$$

Demonstração:

Ver [8]. ∎

O método de Cline é uma generalização do método de Greville, particionando a matriz

A em $A = (U_1, U_2, \ldots, U_l)$, chamamos de $A_1 = (U_1)$ e

$$A_k = (A_{k-1}, U_k),$$

assim, encontramos $A_1^\dagger$, e para $k \geq 2$ inteiro, $A_k^\dagger$ é calculada a partir de $A_{k-1}^\dagger$. Note que depois de l passos teremos

$$A_l^\dagger = A^\dagger.$$

2.3.7 Métodos do tipo $X_{k+1} = X_k + C_k(P_{\mathcal{R}(A)} - AX_k)$

Vejamos agora os principais métodos iterativos para computar a inversa de Moore-Penrose. Para tanto, vamos precisar de alguns resultados clássicos sobre normas matriciais que serão usados para análise de convergência de alguns métodos.

Definição 2.23. Uma norma $\|\cdot\|$ no espaço vetorial $\mathbb{C}^{m\times n}$, é uma função de $\mathbb{C}^{m\times n}$ em $\mathbb{R}$ que satisfaz as seguintes propriedades

(i) $\|A\| \geq 0 \ \forall \ A \in \mathbb{C}^{m\times n}$, e $\|A\| = 0 \Leftrightarrow A = 0$;

(ii) $\|\alpha A\| = |\alpha|\|A\|, \ \forall \ \alpha \in \mathbb{C}, \ A \in \mathbb{C}^{m\times n}$;

(iii) $\|A + B\| \leq \|A\| + \|B\|, \forall \ A, B \in \mathbb{C}^{m\times n}$.

Quando uma norma $\|\cdot\|$ satisfaz

$$\|AB\| \leq \|A\|\|B\|, \ \forall \ A, B \in \mathbb{C}^{m\times n},$$

dizemos que tal norma é submultiplicativa.

As normas matriciais mais comuns são as normas matriciais induzidas por normas vetoriais, as quais são da forma

$$\|A\|_p = \max_{\substack{x\in\mathbb{C}^n,\\ x\neq 0}} \frac{\|Ax\|_p}{\|x\|_p},$$

toda norma matricial induzida é submultiplicativa.

Dada uma matriz $A \in \mathbb{C}^{n\times n}$, denotamos por $\rho(A)$ o raio espectral da matriz A, que é definido por

$$\rho(A) = \max\{|\lambda|; \lambda \text{ é autovalor de } A\}$$

Lema 2.24. *[11] Se $A \in \mathbb{C}^{n\times n}$ e $\epsilon > 0$, então existe uma norma matricial induzida $\|\cdot\|_{A,\epsilon}$, dependendo de A e de ϵ, tal que*

$$\|A\|_{A,\epsilon} \leq \rho(A) + \epsilon$$

Demonstração:

Seja $S^{-1}AS = J$ a forma de Jordan da matriz A, defina $D_\epsilon = \text{diag}(1, \epsilon, \epsilon^2, \ldots, \epsilon^{n-1})$. Temos que

$$(SD_\epsilon)^{-1}ASD_\epsilon = D_\epsilon^{-1}JD_\epsilon = \left[\begin{array}{cccc|cccc|c} \lambda_1 & \epsilon & & & & & & & \\ & \ddots & \ddots & & & & & & \\ & & \ddots & \epsilon & & & & & \\ & & & \lambda_1 & & & & & \\ \hline & & & & \lambda_2 & \epsilon & & & \\ & & & & & \ddots & \ddots & & \\ & & & & & & \ddots & \epsilon & \\ & & & & & & & \lambda_2 & \\ \hline & & & & & & & & \ddots \end{array}\right].$$

Agora, usando a norma vetorial $\|\cdot\|_{A,\epsilon}$ definida por

$$\|x\|_{A,\epsilon} = \|(SD_\epsilon)^{-1}x\|_\infty$$

e definindo $y = (SD_\epsilon)^{-1}x$ geramos a norma matricial induzida por esta norma vetorial como sendo

$$\begin{aligned} \|A\|_{A,\epsilon} &= \max_{x\neq 0} \frac{\|Ax\|_{A,\epsilon}}{\|x\|_{A,\epsilon}} \\ &= \max_{x\neq 0} \frac{\|(SD_\epsilon)^{-1}Ax\|_\infty}{\|(SD_\epsilon)^{-1}x\|_\infty} \\ &= \max_{y\neq 0} \frac{\|(SD_\epsilon)^{-1}ASD_\epsilon y\|_\infty}{\|y\|_\infty} \\ &= \|(SD_\epsilon)^{-1}ASD_\epsilon\|_\infty \\ &= \max_i |\lambda_i + \epsilon| \\ &\leq \max_i |\lambda_i| + \epsilon \\ &= \rho(A) + \epsilon. \end{aligned}$$

■

Note que em geral, temos $\rho(A) \leq \|A\|$ para qualquer norma matricial.

Lema 2.25. *[26] Considere $A \in \mathbb{C}^{n\times n}$, então*

$$\lim_{k\to\infty} A^k = 0 \Leftrightarrow \rho(A) < 1$$

Demonstração:

($\Rightarrow$)

Seja λ um autovalor de A e $x \neq 0$ o autovetor associado a λ. Temos que $A^k x = \lambda^k x$, de onde segue que

$$0 = \lim_{k\to\infty} A^k x = \lim_{k\to\infty} \lambda^k x$$

como $x \neq 0$, assim

$$\lim_{k\to\infty} \lambda^k = 0$$

de onde segue que $|\lambda| < 1$. Como λ foi tomado arbitrário, então

$$\rho(A) < 1.$$

($\Leftarrow$)

Se $\rho(A) < 1$, então existe $\epsilon > 0$ tal que $\rho(A) < 1 - \epsilon$. Assim pelo Lema 2.24, existe uma norma matricial induzida $\|\cdot\|$ tal que

$$\|A\| \leq \rho(A) + \epsilon < 1.$$

Como essa norma é uma norma induzida, então satisfaz a proprieadade submultiplicativa, ou seja, $\|AB\| \leq \|A\|\|B\|$, e assim temos que $\|A^k\| \leq \|A\|^k$, tomando o limite quando k tende a infinito temos

$$\lim_{k\to\infty} \|A^k\| \leq \lim_{k\to\infty} \|A\|^k = 0$$

e portanto

$$\lim_{k\to\infty} A^k = 0.$$

■

Note que $A^\dagger$ é uma inversa-$\{1\}$ de A, e do Teorema 1.19 segue que

$$AA^\dagger = P_{\mathcal{R}(AA^\dagger),\mathcal{N}(AA^\dagger)} = P_{\mathcal{R}(A),\mathcal{N}(A^*)} = P_{\mathcal{R}(A)}.$$

Assim, definimos o resíduo na iteração k por

$$R_k = P_{\mathcal{R}(A)} - AX_k,$$

o qual converge para a matriz nula quando X_k converge para $A^\dagger$.

Definição 2.26.

1. Dizemos que um método iterativo para calcular $A^\dagger$ converge linearmente se existe $r \in [0, 1)$ tal que
$$\lim_{k\to\infty} \frac{\|R_{k+1}\|}{\|R_k\|} = r,$$
para alguma norma matricial submultiplicativa $\| \cdot \|$. Caso $r = 0$, dizemos que a convergência é superlinear.

2. Dizemos que um método iterativo para calcular $A^\dagger$ têm ordem de convergência p, para $p > 1$, se
$$\lim_{k\to\infty} \frac{\|R_{k+1}\|}{\|R_k\|^p} = M,$$
para alguma constante $M > 0$ e alguma norma matricial submultiplicativa $\| \cdot \|$.

Uma classe de métodos bastante conhecida é do tipo

$$X_{k+1} = X_k + C_k R_k,$$

onde C_k é alguma sequência de matrizes e R_k é o resíduo na iteração k.

Da maneira como exposto acima os métodos não são viáveis, pois calcular $P_{\mathcal{R}(A)}$ é comparável a calcular $A^\dagger$. Portanto, precisamos escolher C_k de maneira estratégica para evitar o cálculo de $P_{\mathcal{R}(A)}$. Uma maneira interessante de escolher a sequência C_k é impor que $C_k = C_k P_{\mathcal{R}(A)}$, pois dessa forma temos

$$\begin{aligned} C_k R_k &= C_k(P_{\mathcal{R}(A)} - AX_k) \\ &= C_k P_{\mathcal{R}(A)} - C_k A X_k \\ &= C_k - C_k A X_k \\ &= C_k(I - AX_k), \end{aligned}$$

e assim
$$X_{k+1} = X_k + C_k(I - AX_k).$$

Veremos algumas escolhas de C_k que satisfazem essa condição.

Definição 2.27. Considere $A, B \in \mathbb{C}^{m\times n}$, chamamos de imagem de (A, B) e núcleo de (A, B) respectivamente os conjuntos

$$\mathcal{R}(A, B) = \{Y = AXB \in \mathbb{C}^{m\times n}; X \in \mathbb{C}^{n\times m}\}$$

e
$$\mathcal{N}(A, B) = \{X \in \mathbb{C}^{n\times m}; AXB = 0\}.$$

Definição 2.28. Considere $A = (a_{ij}) \in \mathbb{C}^{m\times n}$ e $B \in \mathbb{C}^{p\times q}$. O produto de Kronecker $A \otimes B$ é a matriz $mp \times nq$ dada por

$$A \otimes B = \begin{pmatrix} a_{11}B & a_{12}B & \cdots & a_{1n}B \\ a_{21}B & a_{22}B & \cdots & a_{2n}B \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}B & a_{m2}B & \cdots & a_{mn}B \end{pmatrix}.$$

Lema 2.29. *[5] Considere o produto interno em $\mathbb{C}^{m\times n}$ dado por*

$$\langle X, Y\rangle = tr(Y^*X) = \sum_{i=1}^{m}\sum_{j=1}^{n} x_{ij}\overline{y_{ij}},$$

então dados $A, B \in \mathbb{C}^{m\times n}$ os conjuntos $\mathcal{R}(A, B)$ e $\mathcal{N}(A^*, B^*)$ são subespaços ortogonais complementares de $\mathbb{C}^{m\times n}$.

Demonstração:

Considere uma transformação linear que leva uma matriz $X \in \mathbb{C}^{m\times n}$ em um vetor $\text{vet}(X)$ dado por

$$\text{vet}(X) = \text{vet}\begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{m1} & x_{m2} & \cdots & x_{mn} \end{pmatrix} = \begin{pmatrix} x_{11} \\ x_{12} \\ \vdots \\ x_{1n} \\ x_{21} \\ x_{22} \\ \vdots \\ x_{2n} \\ \vdots \\ x_{m1} \\ x_{m2} \\ \vdots \\ x_{mn} \end{pmatrix},$$

claramente temos uma bijeção entre os espaços $\mathbb{C}^{m\times n}$ e $\mathbb{C}^{mn}$. Note que

$$\langle X, Y\rangle = \langle \text{vet}(X), \text{vet}(Y)\rangle = \text{vet}(Y)^*\text{vet}(X),$$

ou seja, no espaço $\mathbb{C}^{mn}$ esse produto interno corresponde ao produto interno euclidiano.

Além disso, segue da definição do produto de Kronecker, que

$$\mathcal{R}(A,B) = \mathcal{R}(A \otimes B^T) \quad \text{e} \quad \mathcal{N}(A^*,B^*) = \mathcal{R}((A \otimes B^T)^*),$$

os quais são subespaços ortogonais complementares de $\mathbb{C}^{mn}$, e como vet: $\mathbb{C}^{m\times n} \to \mathbb{C}^{mn}$ é uma bijeção, segue que $\mathcal{R}(A,B)$ e $\mathcal{N}(A^*,B^*)$ são subespaços ortogonais complementares de $\mathbb{C}^{m\times n}$. ∎

O próximo teorema mostra que se escolhermos $C_k = X_0$, para $k = 0,1,2,\ldots$ obtemos um método iterativo de convergência linear para aproximar $A^\dagger$.

Teorema 2.30. *[5] Se $A \in \mathbb{C}^{m\times n}$, $X_0 \in \mathbb{C}^{n\times m}$ e $R_0 = P_{\mathcal{R}(A)} - AX_0$ satisfazem*

$$X_0 \in \mathcal{R}(A^*,A^*) \quad e \quad \rho(R_0) < 1,$$

então a sequência

$$X_{k+1} = X_k + X_0(I - AX_k), \quad k = 0,1,2,\ldots \tag{2.4}$$

converge para $A^\dagger$. Além disso, a sequência de resíduos satisfaz

$$\|R_{k+1}\| \leq \|R_0\|\|R_k\|, \quad k = 0,1,2,\ldots$$

para alguma norma matricial submultiplicativa $\|\cdot\|$.

Demonstração:

Primeiro vejamos que $X_0P_{\mathcal{R}(A)} = X_0$. De fato, como $X_0 \in \mathcal{R}(A^*,A^*)$, então existe uma matriz $B \in \mathbb{C}^{m\times n}$ tal que $X_0 = A^*BA^*$, além disso podemos escrever $P_{\mathcal{R}(A)} = AA^\dagger$, e pelo item 6 do Teorema 1.2 temos que $A^*AA^\dagger = A^*$, assim

$$X_0P_{\mathcal{R}(A)} = X_0AA^\dagger = A^*BA^*AA^\dagger = A^*BA^* = X_0.$$

Portanto, podemos escrever a iteração como

$$X_{k+1} = X_k + X_0(P_{\mathcal{R}(A)} - AX_k) = X_k + X_0R_k,$$

de onde segue que

$$
\begin{aligned}
R_{k+1} &= P_{\mathcal{R}(A)} - AX_{k+1} \\
&= P_{\mathcal{R}(A)} - AX_k - AX_0R_k \\
&= P_{\mathcal{R}(A)} - AA^{\dagger}AX_k - AX_0R_k \\
&= P^2_{\mathcal{R}(A)} - P_{\mathcal{R}(A)}AX_k - AX_0R_k \\
&= P_{\mathcal{R}(A)}(P_{\mathcal{R}(A)} - AX_k) - AX_0R_k \\
&= P_{\mathcal{R}(A)}R_k - AX_0R_k \\
&= (P_{\mathcal{R}(A)} - AX_0)R_k \\
&= R_0R_k.
\end{aligned}
$$

Assim, considerando alguma norma matricial submultiplicativa $\|\cdot\|$, temos

$$\|R_{k+1}\| = \|R_0R_k\| \leq \|R_0\|\|R_k\|.$$

Além disso, como

$$R_k = R_0R_{k-1} = R_0^2R_{k-2} = \ldots = R_0^{k+1},$$

então

$$\|R_k\| = \|R_0^{k+1}\| \leq \|R_0\|^{k+1},$$

e tomando o limite quando k tende a infinito, segue da hipótese de que $\rho(R_0) < 1$ e do Lema 2.25 que

$$\lim_{k\to\infty} \|R_k\| = 0.$$

Portanto $P_{\mathcal{R}(A)} - AX_k \to 0$, e reescrevendo o método da forma

$$
\begin{aligned}
X_{k+1} &= X_k + X_0R_k \\
&= X_{k-1} + X_0R_{k-1} + X_0R_k \\
&= \ldots \\
&= X_0 + X_0R_0 + X_0R_1 + \ldots + X_0R_k \\
&= X_0 + X_0R_0 + X_0R_0^2 + \ldots + X_0R_0^{k+1} \\
&= X_0(I + R_0 + R_0^2 + \ldots + R_0^{k+1}),
\end{aligned}
$$

como $\rho(R_0) < 1$, do Lema 2.25 implica que X_k converge para algum limite X_∞, assim

$$AX_\infty = P_{\mathcal{R}(A)},$$

de onde segue que X_∞ é uma inversa-$\{1\}$ de A. Além disso, como $X_0 \in \mathcal{R}(A^*, A^*)$, e supondo que $X_k \in \mathcal{R}(A^*, A^*)$ podemos escrever $X_0 = A^*BA^*$ e $X_k = A^*CA^*$, o que implica em

$$\begin{aligned} X_{k+1} &= X_k + X_0 - X_0 A X_k \\ &= A^* C A^* + A^* B A^* - A^* B A^* A A^* C A^* \\ &= A^*(C + B - BA^* A A^* C) A^*, \end{aligned}$$

ou seja, $X_{k+1} \in \mathcal{R}(A^*, A^*)$ e segue por indução que

$$X_k \in \mathcal{R}(A^*, A^*), \quad k = 0, 1, 2, \ldots$$

portanto $X_\infty \in \mathcal{R}(A^*, A^*)$.

Além disso, pelo Teorema 1.8 temos que a solução geral da equação $AXA = A$ é dada por

$$\begin{aligned} X &= A^\dagger A A^\dagger + Y - A^\dagger A Y A A^\dagger \\ &= A^\dagger + Y - A^\dagger A Y A A^\dagger, \end{aligned}$$

para $Y \in \mathbb{C}^{n \times m}$ arbitrária. Porém note que $A^\dagger \in \mathcal{R}(A^*, A^*)$ e $(Y - A^\dagger A Y A A^\dagger) \in \mathcal{N}(A, A)$. De fato,

$$\begin{aligned} A^\dagger &= A^\dagger A A^\dagger A A^\dagger \\ &= (A^\dagger A)^* A^\dagger (A A^\dagger)^* \\ &= A^* (A^\dagger)^* A^\dagger (A^\dagger)^* A^* \end{aligned}$$

e

$$A(Y - A^\dagger A Y A A^\dagger) A = AYA - AA^\dagger A Y A A^\dagger A = AYA - AYA = 0,$$

sendo assim, pelo Lema 2.29 temos que $\mathcal{R}(A^*, A^*)$ e $\mathcal{N}(A, A)$ são subespaços complementares de $\mathbb{C}^{n \times m}$, o que implica que a representação

$$X = A^\dagger + Y - A^\dagger A Y A A^\dagger$$

é única, e como $X_\infty \in \mathcal{R}(A^*, A^*)$ e satisfaz $AX_\infty A = A$ segue que

$$X_\infty = A^\dagger$$

. ∎

O próximo teorema mostra que a escolha

$$C_k = X_k(I + T_k + T_k^2 + \ldots + T_k^{p-2}),$$

onde $T_k = (I - AX_k)$, para $p \geq 2$ inteiro, fornece um método iterativo de ordem p para computar a inversa de Moore-Penrose.

Teorema 2.31. *[5] Se $A \in \mathbb{C}^{m\times n}$, $X_0 \in \mathbb{C}^{n\times m}$ e $R_0 = P_{\mathcal{R}(A)} - AX_0$ satisfazem*

$$X_0 \in \mathcal{R}(A^*, A^*) \qquad e \qquad \rho(R_0) < 1,$$

então para $p \geq 2$ inteiro, a sequência

$$X_{k+1} = X_k + C_k T_k,$$

onde

$$C_k = X_k(I + T_k + T_k^2 + \ldots + T_k^{p-2}) \qquad e \qquad T_k = I - AX_k,$$

converge para $A^\dagger$, e a sequência de resíduos R_k satisfaz

$$\|R_{k+1}\| \leq \|R_k\|^p, \quad k = 0, 1, 2, \ldots$$

para alguma norma matricial submultiplicativa $\|\cdot\|$.

Demonstração:

Vejamos que $X_k \in \mathcal{R}(A^*, A^*)$ para $k = 0, 1, 2, \ldots$. De fato, procedendo por indução, temos por hipótese que $X_0 \in \mathcal{R}(A^*, A^*)$, e supondo que $X_k \in \mathcal{R}(A^*, A^*)$ temos que $X_k T_k^i = X_k(I - AX_k)^i \in \mathcal{R}(A^*, A^*)$, para $i = 1, 2, \ldots, p-1$, e como

$$\begin{aligned}
X_{k+1} &= X_k + C_k T_k \\
&= X_k + X_k(I + T_k + T_k^2 + \ldots + T_k^{p-2})T_k \\
&= X_k + X_k T_k + X_k T_k^2 + \ldots + X_k T_k^{p-1},
\end{aligned}$$

segue que $X_{k+1} \in \mathcal{R}(A^*, A^*)$.

Sendo assim, podemos escrever $X_k = A^*BA^*$ e $X_k T_k^i = A^* C_i A^*$, para $i = 1, 2, \ldots, p-1$, de onde temos que

$$\begin{aligned}
C_k P_{\mathcal{R}(A)} &= X_k(I + T_k + T_k^2 + \ldots + T_k^{p-2})AA^\dagger \\
&= X_k AA^\dagger + X_k T_k AA^\dagger + X_k T_k^2 AA^\dagger + \ldots + X_k T_k^{p-2} AA^\dagger \\
&= A^*BA^*AA^\dagger + A^*C_1A^*AA^\dagger + A^*C_2A^*AA^\dagger + \ldots + A^*C_{p-2}A^*AA^\dagger \\
&= A^*BA^* + A^*C_1A^* + A^*C_2A^* + \ldots + A^*C_{p-2}A^* \\
&= X_k + X_k T_k + X_k T_k^2 + \ldots + X_k T_k^{p-2} \\
&= X_k(I + T_k + T_k^2 + \ldots + T_k^{p-2}) \\
&= C_k, \quad k = 0, 1, 2, \ldots,
\end{aligned}$$

e então podemos reescrever o método da forma

$$X_{k+1} = X_k(I + R_k + R_k^2 + \ldots + R_k^{p-1}),$$

de onde segue que

$$\begin{aligned} R_{k+1} &= P_{\mathcal{R}(A)} - AX_{k+1} \\ &= P_{\mathcal{R}(A)} - AX_k(I + R_k + R_k^2 + \ldots + R_k^{p-1}) \\ &= P_{\mathcal{R}(A)} - AX_k - AX_k(R_k + R_k^2 + \ldots + R_k^{p-1}) \\ &= R_k - AX_k(R_k + R_k^2 + \ldots + R_k^{p-1}). \end{aligned} \tag{2.5}$$

Porém, como

$$R_k = P_{\mathcal{R}(A)} - AX_k = P_{\mathcal{R}(A)}^2 - P_{\mathcal{R}(A)}AX_k = P_{\mathcal{R}(A)}(P_{\mathcal{R}(A)} - AX_k) = P_{\mathcal{R}(A)}R_k,$$

então para $i = 1, 2, \ldots, p-1$ temos

$$\begin{aligned} R_k^i - AX_kR_k^i &= R_kR_k^{i-1} - AX_kR_k^i \\ &= P_{\mathcal{R}(A)}R_kR_k^{i-1} - AX_kR_k^i \\ &= P_{\mathcal{R}(A)}R_k^i - AX_kR_k^i \\ &= (P_{\mathcal{R}(A)} - AX_k)R_k^i \\ &= R_kR_k^i \\ &= R_k^{i+1}. \end{aligned} \tag{2.6}$$

Juntando (2.5) e (2.6) temos

$$\begin{aligned} R_{k+1} &= R_k - AX_k(R_k + R_k^2 + \ldots + R_k^{p-1}) \\ &= R_k - AX_kR_k - AX_k(R_k^2 + \ldots + R_k^{p-1}) \\ &= R_k^2 - AX_k(R_k^2 + \ldots + R_k^{p-1}) \\ &= \ldots \\ &= R_k^{p-1} - AX_kR_k^{p-1} \\ &= R_k^p, \end{aligned}$$

considerando uma norma submultiplicativa $\|\cdot\|$, temos

$$\|R_{k+1}\| = \|R_k^p\| \leq \|R_k\|^p.$$

Além disso,

$$R_k = R_{k-1}^p = R_{k-2}^{p^2} = \ldots = R_0^{p^k},$$

e como $\rho(R_0) < 1$, segue do Lema 2.25 que

$$\lim_{k\to\infty} \|R_k\| = 0,$$

ou seja, $P_{\mathcal{R}(A)} - AX_k \to 0$, e de $\rho(R_0) < 1$ e $R_{k+1} = R_k^p$ podemos concluir que $X_k \to X_\infty$, a qual satisfaz $AX_\infty = P_{\mathcal{R}(A)}$, e portanto é solução de $AXA = A$. Como $X_k \in \mathcal{R}(A^*, A^*)$ para $k = 0, 1, 2, \ldots$, segue que $X_\infty \in \mathcal{R}(A^*, A^*)$ e procedendo da mesma forma que na demonstração do Teorema 2.30 concluímos que

$$X_\infty = A^\dagger.$$

∎

O teorema anterior mostra que podemos construir métodos com a taxa de convergência desejada, porém quando aumentamos a taxa de convergência do método a iteração fica mais cara, na prática o método mais mais vantajoso dentro dessa classe é o caso $p = 2$, como mostrado em [5].

Outra questão a discutir é a aproximação inicial X_0. Nos teoremas acima pede-se que $X_0 \in \mathcal{R}(A^*, A^*)$, e $\rho(R_0) < 1$, o que não é uma condição fácil de ser checada. Frequentemente é usada como aproximação inicial a matriz αA^* para algum $\alpha \in \mathbb{R}$. Note que X_0 dessa forma pertence a $\mathcal{R}(A^*, A^*)$, pois

$$\alpha A^* = \alpha(AA^\dagger A)^* = A^*\alpha A^\dagger A^*.$$

O próximo resultado estabelece como escolher α.

Teorema 2.32. *[5] Sejam $0 \neq A \in \mathbb{C}^{m\times n}$ com posto$(A) = r$, $\alpha \in \mathbb{R}$ e*

$$R_0 = P_{\mathcal{R}(A)} - \alpha AA^*.$$

Temos que $\rho(R_0) < 1$ se e somente se

$$0 < \alpha < \frac{2}{\rho(AA^*)}.$$

Demonstração:

Como AA^* é Hermitiana positiva semidefinida, então seus autovalores são reais e não negativos. Sejam

$$\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_r > \lambda_{r+1} = \ldots = \lambda_m = 0,$$

os autovalores de AA^*, neste caso os autovalores não nulos de R_0 são

$$\mu_i = 1 - \alpha\lambda_i, \quad i = 1, 2, \ldots, r.$$

$(\Rightarrow)$:

Supondo $\rho(R_0) < 1$ então

$$-1 < \mu_i < 1, \quad i = 1, 2, \ldots, r,$$

e como $\mu_i = 1 - \alpha\lambda_i$, para $i = 1, 2, \ldots, r$ temos

$$\begin{array}{ccccc} -1 & < & 1 - \alpha\lambda_i & < & 1 \\ -2 & < & -\alpha\lambda_i & < & 0 \\ 0 & < & \alpha\lambda_i & < & 2 \\ 0 & < & \alpha & < & \frac{2}{\rho(AA^*)}. \end{array}$$

$(\Leftarrow)$:

Como os autovalores não nulos de R_0 são $\mu_i = 1 - \alpha\lambda_i$, de $\alpha > 0$ temos

$$\mu_i = 1 - \alpha\lambda_i < 1 - 0 = 1,$$

e de $\alpha < \frac{2}{\rho(AA^*)} = \frac{2}{\lambda_1}$ temos

$$\mu_i > 1 - \frac{2}{\lambda_1}\lambda_i \geq 1 - \frac{2}{\lambda_i}\lambda_i = 1 - 2 = -1,$$

portanto $-1 < \mu_i < 1$ para $i = 1, 2, \ldots, r$ de onde segue que

$$\rho(R_0) < 1.$$

■

2.3.8 Métodos baseados nas equações de Penrose

Comumente encontramos métodos iterativos para aproximar a inversa de Moore-Penrose baseados em algumas equações de Penrose (1.1)-(1.4).

O método de ordem 2 apresentado anteriormente pode ser visto com um método baseado na equação (1.2), pois se $X = XAX$ então

$$X = X + X - X = X + X - XAX = X + X(I - AX).$$

Recentemente, Petković e Stanimirović em [23] e [24] proporam um método iterativo baseado nas equações (1.2) e (1.4) usando a seguinte ideia

$$X^* = (XAX)^* = X^*(XA)^* = X^*XA,$$

assim, para $\beta \in \mathbb{R}$, temos

$$\begin{aligned}
X^* &= X^* - \beta(X^*XA - X^*) \\
&= X^*(I - \beta XA) + \beta X^* \\
X &= (I - \beta XA)^*X + \beta X,
\end{aligned}$$

o que sugere o método

$$X_{k+1} = (I - \beta X_k A)^* X_k + \beta X_k. \tag{2.7}$$

Lema 2.33. *[23] Se $A \in \mathbb{C}^{m\times n}$, $X_0 = \beta A^*$ e $X = A^\dagger$, então a sequência de matrizes gerada pelo algoritmo (2.7) satisfaz*

$$(X_kA)^* = X_kA, \quad XAX_k = X_k, \quad X_kAX = X_k, \quad k = 0, 1, 2, \dots\ . \tag{2.8}$$

Demonstração:

Procedendo por indução, para $k = 0$ segue do item 6 do Teorema 1.2 que

$$\begin{aligned}
(X_0A)^* &= (\beta A^*A)^* = \beta A^*A = X_0A; \\
XAX_0 &= \beta XAA^* = \beta A^* = X_0 \\
X_0AX &= \beta A^*AX = \beta A^* = X_0.
\end{aligned}$$

Agora, suponha que as três equações em (2.8) sejam válidas para k. Para $k+1$ temos

$$\begin{aligned}
(X_{k+1}A)^* &= (((I - \beta X_kA)^*X_k + \beta X_k)A)^* \\
&= ((I - \beta X_kA)^*X_kA + \beta X_kA)^* \\
&= (X_kA)^*(I - \beta X_kA) + \beta(X_kA)^* \\
&= X_kA(I - \beta X_kA) + \beta X_kA \\
&= (I - \beta X_kA)X_kA + \beta X_kA \\
&= (I - \beta(X_kA)^*)X_kA + \beta X_kA \\
&= (I - \beta X_kA)^*X_kA + \beta X_kA \\
&= ((I - \beta X_kA)^*X_k + \beta X_k)A \\
&= X_{k+1}A,
\end{aligned}$$

$$\begin{aligned}
XAX_{k+1} &= XA((I-\beta X_kA)^*X_k+\beta X_k)\\
&= XA(I-\beta(X_kA)^*)X_k+\beta XAX_k\\
&= XA(I-\beta X_kA)X_k+\beta XAX_k\\
&= XAX_k-\beta XAX_kAX_k+\beta XAX_k\\
&= X_k-\beta X_kAX_k+\beta X_k\\
&= (I-\beta X_kA)X_k+\beta X_k\\
&= (I-\beta(X_kA)^*)X_k+\beta X_k\\
&= (I-\beta X_kA)^*X_k+\beta X_k\\
&= X_{k+1}
\end{aligned}$$

e

$$\begin{aligned}
X_{k+1}AX &= ((I-\beta X_kA)^*X_k+\beta X_k)AX\\
&= (I-\beta X_kA)^*X_kAX+\beta X_kAX\\
&= (I-\beta X_kA)^*X_k+\beta X_k\\
&= X_{k+1}.
\end{aligned}$$

■

Do lema anterior, segue que o método pode ser escrito na forma

$$X_{k+1}=(I-\beta X_kA)X_k+\beta X_k=(1+\beta)X_k-\beta X_kAX_k. \tag{2.9}$$

Teorema 2.34. *[23] Sejam $A\in\mathbb{C}^{m\times n}$ e $X_0=\beta A^*$. Se*

$$\|(X_0-X)A\|\leq 1, \quad e \quad 0<\beta\leq 1,$$

para alguma norma matricial submultiplicativa $\|\cdot\|$, então a sequência gerada por (2.9) converge para $X=A^\dagger$.

Demonstração:

Note que, do Lema 2.33 e das equações de Penrose temos que

$$\begin{aligned}
X_{k+1}A-XA &= ((1+\beta)X_k-\beta X_kAX_k)A-XA\\
&= (1+\beta)X_kA-\beta X_kAX_kA-XA\\
&= X_kA+\beta X_kA-\beta X_kAX_kA-XA\\
&= XAX_kA+\beta X_kAXA-\beta X_kAX_kA-XAXA\\
&= -(\beta X_kA-XA)(X_kA-XA),
\end{aligned}$$

e como

$$\beta X_k A - XA = \beta X_k A - XA + \beta XA - \beta XA = \beta(X_k A - XA) - (1-\beta)XA,$$

segue que

$$\begin{aligned} X_{k+1}A - XA &= -(\beta X_k A - XA)(X_k A - XA) \\ &= -(\beta(X_k A - XA) - (1-\beta)XA)(X_k A - XA) \\ &= -\beta(X_k A - XA)^2 + (1-\beta)XA(X_k A - XA) \\ &= -\beta(X_k A - XA)^2 + (1-\beta)(XAX_k A - XAXA) \\ &= -\beta(X_k A - XA)^2 + (1-\beta)(X_k A - XA). \qquad (2.10) \end{aligned}$$

Definindo $E_k = X_k - X$, temos que (2.10) implica em

$$E_{k+1}A = -\beta(E_k A)^2 + (1-\beta)E_k A.$$

Agora, seja $t_k = \|E_k A\|$ para alguma norma submultiplicativa $\|\cdot\|$. Vejamos que $t_k \to 0$. Note que $t_k < 1$, $k = 0, 1, 2, \ldots$. De fato, procedendo por indução, para $k = 0$ temos $t_0 < 1$ por hipótese, e supondo $t_k < 1$ implica em

$$\begin{aligned} t_{k+1} &= \|E_{k+1}A\| \\ &= \|-\beta(E_k A)^2 + (1-\beta)E_k A\| \\ &\leq \beta\|(E_k A)\|^2 + (1-\beta)\|E_k A\| \\ &= \beta t_k^2 + (1-\beta)t_k \qquad (2.11) \\ &< \beta t_k + t_k - \beta t_k \\ &= t_k \\ &< 1, \end{aligned}$$

de onde concluímos que a sequência $\{t_k\}$ é decrescente e limitada inferiormente por 0, logo $t_k \to t$ para algum $t \in [0, 1)$, e tomando o limite quando k tende a infinito em (2.11) temos

$$\begin{aligned} t &\leq \beta t^2 + (1-\beta)t \\ 0 &\leq \beta t^2 - \beta t \\ t &\leq t^2, \end{aligned}$$

logo $t \geq 1$ ou $t = 0$, e como $t \in [0, 1)$ então $t = 0$. Assim, temos que

$$\|X_k - X\| = \|X_k AX - XAX\| \leq \|X_k A - XA\|\|X\| = \|E_k A\|\|X\| = t_k\|X\|,$$

de onde segue que $\|X_k - X\| \to 0$, e portanto

$$\lim_{k\to\infty} X_k = X = A^\dagger.$$

■

O próximo teorema mostra que a taxa de convergência do método (2.9) é linear se $\beta < 1$ e quadrática se $\beta = 1$. Note que quando $\beta = 1$ recaímos no método de ordem 2 do Teorema 2.31.

Teorema 2.35. *[23] Se $A \in \mathbb{C}^{m\times n}$, $X_0 = \beta A^*$ e X_k é gerado pela sequência (2.9), então o erro $E_k = X_k - X$ satisfaz*

$$E_{k+1} = (1-\beta)E_k - \beta E_k A E_k.$$

Demonstração:
Como $E_k = X_k - X$ então $X_k = X + E_k$, assim

$$\begin{aligned}
X_{k+1} &= (1+\beta)(X+E_k) - \beta(X+E_k)A(X+E_k)\\
&= (1+\beta)X + (1+\beta)E_k - \beta XAX - \beta XAE_k - \beta E_kAX - \beta E_kAE_k\\
&= X + \beta X + (1+\beta)E_k - \beta X - \beta XAE_k - \beta E_kAX - \beta E_kAE_k\\
&= X + (1+\beta)E_k - \beta XAE_k - \beta E_kAX - \beta E_kAE_k,
\end{aligned}$$

portanto

$$\begin{aligned}
E_{k+1} &= X_{k+1} - X = (1+\beta)E_k - \beta XAE_k - \beta E_kAX - \beta E_kAE_k\\
&= (1+\beta)(X_k - X) - \beta XA(X_k - X) - \beta(X_k - X)AX - \beta E_kAE_k\\
&= (1+\beta)(X_k - X) - \beta(XAX_k - XAX) - \beta(X_kAX - XAX) - \beta E_kAE_k\\
&= (1+\beta)(X_k - X) - \beta(X_k - X) - \beta(X_k - X) - \beta E_kAE_k\\
&= (1+\beta-2\beta)(X_k - X) - \beta E_kAE_k\\
&= (1-\beta)E_k - \beta E_kAE_k.
\end{aligned}$$

■

2.4 Cálculo da inversa de Drazin

2.4.1 Inversa de Drazin via decomposição de posto completo

Seja $A \in \mathbb{C}^{n\times n}$ com $\text{Ind}(A) = k$. Na seção anterior vimos que a inversa de Drazin de A pode ser vista como uma inversa-$\{2\}$ com imagem e núcleo prescritos, ou seja,

$$A_d = A^{(2)}_{\mathcal{R}(A^k),\mathcal{N}(A^k)},$$

assim, podemos usar a decomposição de posto completo para calcular A_d. Para tanto, precisamos de uma decomposição de posto completo de A^k. Porém para calcular A^k precisamos conhecer o index da matriz, o que não é um procedimente barato computacionalmente. O próximo teorema nos fornece uma maneira de encontrar uma decomposição de posto completo de A^k sem precisar conhecer o index da matriz previamente.

Teorema 2.36. *[29] Seja $A \in \mathbb{C}^{n\times n}$. Considere a sequência de decomposições de posto completo*

$$A = F_1G_1, \quad G_1F_1 = F_2G_2, \quad G_2F_2 = F_3G_3, \quad \dots,$$

onde F_iG_i é a decomposição de posto completo de $G_{i-1}F_{i-1}$ para $i = 2, 3, \dots$. Em algum momento teremos que G_kF_k é não singular ou $G_kF_k = 0$. Se k é o primeiro inteiro para o qual isso ocorre, então

$$\mathit{Ind}(A) = \begin{cases} k, & \text{caso } G_kF_k \text{ é não singular,} \\ k+1, & \text{caso } G_kF_k = 0. \end{cases}$$

Demonstração:

Se $G_iF_i \in \mathbb{C}^{p\times p}$ e $\text{posto}(G_iF_i) = q < p$ então $G_{i+1}F_{i+1} \in \mathbb{C}^{q\times q}$, ou seja, o tamanho de $G_{i+1}F_{i+1}$ deve ser estritamente menor que o tamanho de G_iF_i quando G_iF_i é singular, de onde segue que em algum momento teremos G_iF_i não singular ou $G_iF_i = 0$. Seja k o primeiro inteiro para o qual isso ocorre, podemos escrever

$$\begin{aligned} A^k &= (F_1G_1)^k = F_1(G_1F_1)^{k-1}G_1 = F_1(F_2G_2)^{k-1}G_1 \\ &= F_1F_2(G_2F_2)^{k-2}G_2G_1 = \ \dots \ = F_1F_2\dots F_{k-1}(F_kG_k)G_{k-1}\dots G_2G_1 \end{aligned} \tag{2.12}$$

e

$$A^{k+1} = F_1F_2\dots F_{k-1}F_k(G_kF_k)G_kG_{k-1}\dots G_2G_1. \tag{2.13}$$

Assim, assumindo que G_kF_k é não singular, se $F_k \in \mathbb{C}^{p\times r}, G_k \in \mathbb{C}^{r\times p}$ então temos $\text{posto}(F_k) = \text{posto}(G_k) = r$ e $\text{posto}(F_kG_k) = r$, e como $G_kF_k \in \mathbb{C}^{r\times r}$ é não singular, então

$$\text{posto}(G_kF_k) = r = \text{posto}(F_kG_k),$$

e como as matrizes F_i têm posto coluna completo e as matrizes G_i têm posto linha completo para $i = 1, 2, 3 \ldots$, segue de (2.12) e (2.13) que

$$\text{posto}(A^{k+1}) = \text{posto}(G_k F_k) = \text{posto}(F_k G_k) = \text{posto}(A^k),$$

e como k é o menor inteiro para o qual isso vale, temos que $\text{Ind}(A) = k$.

Para o caso em que $G_k F_k = 0$ temos de (2.13) que $A^{k+1} = 0$, e portanto

$$\text{posto}(A^{k+2}) = \text{posto}(A^{k+1}) = 0,$$

de onde segue que $\text{Ind}(A) = k + 1$. ■

Com base no teorema acima fazemos o seguinte procedimento. Defina $A_1 = A$, para $j = 1$ até que $G_k F_k$ seja não singular ou nula, faça a decomposição de posto completo de A_j

$$A_j = F_j G_j,$$

e construa A_{j+1} da forma

$$A_{j+1} = G_j F_j.$$

Assim, temos que

$$\begin{aligned} A^k &= (F_1 G_1)^k = F_1 (G_1 F_1)^{k-1} G_1 = F_1 (F_2 G_2)^{k-1} G_1 \\ &= F_1 F_2 (G_2 F_2)^{k-2} G_2 G_1 = \ \ldots \ = (F_1 F_2 \ldots F_{k-1} F_k)(G_k G_{k-1} \ldots G_2 G_1), \end{aligned}$$

é uma decomposição de posto completo de A^k, e segue do Teorema 1.29 que

$$\begin{aligned} A_d &= F_1 F_2 \ldots F_k (G_k G_{k-1} \ldots G_1 A F_1 F_2 \ldots F_k)^{-1} G_k G_{k-1} \ldots G_1 \\ &= F_1 F_2 \ldots F_k (G_k G_{k-1} \ldots G_1 F_1 G_1 F_1 F_2 \ldots F_k)^{-1} G_k G_{k-1} \ldots G_1 \\ &= F_1 F_2 \ldots F_k (G_k G_{k-1} \ldots G_3 G_2 F_2 G_2 F_2 G_2 F_2 F_3 \ldots F_k)^{-1} G_k G_{k-1} \ldots G_1 \\ &= \ldots \\ &= F_1 F_2 \ldots F_k ((G_k F_k)^{k+1})^{-1} G_k G_{k-1} \ldots G_1 \\ &= F_1 F_2 \ldots F_k (G_k F_k)^{-(k+1)} G_k G_{k-1} \ldots G_1. \end{aligned}$$

ou seja,

$$A_d = F_1 F_2 \ldots F_k (G_k F_k)^{-(k+1)} G_k G_{k-1} \ldots G_1.$$

Em particular a inversa de grupo A_g de uma matriz $A \in \mathbb{C}^{n \times n}$ com $\text{Ind}(A) = 1$ é dada por

$$A_g = F_1 (G_1 F_1)^{-2} G_1.$$

2.4.2 Inversa de Drazin via Fórmula de Greville

O próximo teorema mostra a relação entre a inversa de Drazin e uma inversa-$\{1\}$, tal expressão é conhecida como Fórmula de Greville.

Teorema 2.37. *[29] Se $A \in \mathbb{C}^{n\times n}$ com $Ind(A) = k$, então para todo inteiro $l \geq k$ vale*

$$A_d = A^l (A^{2l+1})^{(1)} A^l,$$

onde $(A^{2l+1})^{(1)}$ denota uma inversa-$\{1\}$ de A^{2l+1}. Em particular temos

$$A_d = A^l (A^{2l+1})^{\dagger} A^l.$$

Demonstração:

Escrevendo A da forma

$$A = P \begin{pmatrix} C & 0 \\ 0 & N \end{pmatrix} P^{-1},$$

temos que

$$A^{2l+1} = P \begin{pmatrix} C^{2l+1} & 0 \\ 0 & N^{2l+1} \end{pmatrix} P^{-1} = P \begin{pmatrix} C^{2l+1} & 0 \\ 0 & 0 \end{pmatrix} P^{-1}.$$

Se X é uma inversa-$\{1\}$ de A^{2l+1} então X satisfaz

$$\begin{aligned} A^{2l+1} X A^{2l+1} &= A^{2l+1} \\ P \begin{pmatrix} C^{2l+1} & 0 \\ 0 & 0 \end{pmatrix} P^{-1} X P \begin{pmatrix} C^{2l+1} & 0 \\ 0 & 0 \end{pmatrix} P^{-1} &= P \begin{pmatrix} C^{2l+1} & 0 \\ 0 & 0 \end{pmatrix} P^{-1} \\ \begin{pmatrix} C^{2l+1} & 0 \\ 0 & 0 \end{pmatrix} P^{-1} X P \begin{pmatrix} C^{2l+1} & 0 \\ 0 & 0 \end{pmatrix} &= \begin{pmatrix} C^{2l+1} & 0 \\ 0 & 0 \end{pmatrix}. \end{aligned}$$

Seja

$$P^{-1} X P = \begin{pmatrix} X_1 & X_2 \\ X_3 & X_4 \end{pmatrix}.$$

Temos que

$$\begin{aligned} \begin{pmatrix} C^{2l+1} & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} X_1 & X_2 \\ X_3 & X_4 \end{pmatrix} \begin{pmatrix} C^{2l+1} & 0 \\ 0 & 0 \end{pmatrix} &= \begin{pmatrix} C^{2l+1} & 0 \\ 0 & 0 \end{pmatrix} \\ \begin{pmatrix} C^{2l+1} X_1 & C^{2l+1} X_2 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} C^{2l+1} & 0 \\ 0 & 0 \end{pmatrix} &= \begin{pmatrix} C^{2l+1} & 0 \\ 0 & 0 \end{pmatrix} \\ \begin{pmatrix} C^{2l+1} X_1 C^{2l+1} & 0 \\ 0 & 0 \end{pmatrix} &= \begin{pmatrix} C^{2l+1} & 0 \\ 0 & 0 \end{pmatrix}, \end{aligned}$$

de onde segue que

$$C^{2l+1}X_1C^{2l+1} = C^{2l+1},$$

e portanto

$$X_1 = C^{-2l-1}.$$

Sendo assim,

$$P^{-1}XP = \begin{pmatrix} X_1 & X_2 \\ X_3 & X_4 \end{pmatrix} = \begin{pmatrix} C^{-2l-1} & X_2 \\ X_3 & X_4 \end{pmatrix},$$

e

$$X = P \begin{pmatrix} C^{-2l-1} & X_2 \\ X_3 & X_4 \end{pmatrix} P^{-1}.$$

Com isso

$$\begin{aligned} A^lXA^l &= P \begin{pmatrix} C^l & 0 \\ 0 & 0 \end{pmatrix} P^{-1}P \begin{pmatrix} C^{-2l-1} & X_2 \\ X_3 & X_4 \end{pmatrix} P^{-1}P \begin{pmatrix} C^l & 0 \\ 0 & 0 \end{pmatrix} P^{-1} \\ &= P \begin{pmatrix} C^l & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} C^{-2l-1} & X_2 \\ X_3 & X_4 \end{pmatrix} \begin{pmatrix} C^l & 0 \\ 0 & 0 \end{pmatrix} P^{-1} \\ &= P \begin{pmatrix} C^{-l-1} & C^lX_2 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} C^l & 0 \\ 0 & 0 \end{pmatrix} P^{-1} \\ &= P \begin{pmatrix} C^{-1} & 0 \\ 0 & 0 \end{pmatrix} P^{-1} \\ &= A_d. \end{aligned}$$

■

2.4.3 Inversa de Drazin via decomposição SVD

Os próximos resultados nos permitem construir um método para encontrar a inversa de Drazin usando decomposição em valores singulares.

Lema 2.38. *[29] Sejam $A, W \in \mathbb{C}^{n\times n}$, com Ind$(A) = k$ e W não singular. Se $R = WAW^{-1}$, então*

$$A_d = W^{-1}R_dW.$$

Demonstração:

Primeiramente, vejamos que $\text{Ind}(A) = \text{Ind}(R)$. De fato, como

$$\text{posto}(A) = \text{posto}(W^{-1}RW) \leq \text{posto}(R)$$

e

$$\text{posto}(R) = \text{posto}(WAW^{-1}) \leq \text{posto}(A),$$

então posto(A) =posto(R), e como $A^k = W^{-1}R^kW$ e $R^k = WA^kW^{-1}$ segue que Ind(A) = Ind(R).

Agora defina $X = W^{-1}R_dW$, vejamos que X satisfaz as equações

$$A^kXA = A^k, \quad XAX = X, \quad AX = XA.$$

De fato,

$$\begin{aligned}
A^kXA &= (W^{-1}RW)^kW^{-1}R_dWW^{-1}RW = W^{-1}R^kWW^{-1}R_dWW^{-1}RW \\
&= W^{-1}R^kR_dRW = W^{-1}R^kW = (W^{-1}RW)^k = A^k; \\
XAX &= W^{-1}R_dWW^{-1}RWW^{-1}R_dW = W^{-1}R_dRR_dW = W^{-1}R_dW = X; \\
AX &= W^{-1}RWW^{-1}R_dW = W^{-1}RR_dW = W^{-1}R_dRW = \\
&= W^{-1}R_dWW^{-1}RW = XA.
\end{aligned}$$

■

Lema 2.39. *[29] Seja $M \in \mathbb{C}^{n\times n}$ uma matriz triangular inferior por blocos*

$$M = \begin{pmatrix} A & 0 \\ B & C \end{pmatrix}.$$

Se A é não singular, então

$$Ind(M) = Ind(C).$$

Demonstração:

Se Ind(M) $= k$, então posto(M^k) =posto(M^{k+1}), e como

$$M^k = \begin{pmatrix} A^k & 0 \\ S(k) & C^k \end{pmatrix},$$

onde $S(k)$ é uma matriz que depende de A, B e C, então posto(M^k) = posto(A^k) + posto(C^k), e portanto posto(A^k) + posto(C^k) = posto(A^{k+1}) + posto(C^{k+1}).

Porém, por hipótese A é não singular, o que implica em posto(A^k) = posto(A^{k+1}), portanto posto(C^k) = posto(C^{k+1}), de onde segue que Ind(C) $= k$. ■

O teorema a seguir fornece uma maneira de calcular a inversa de Drazin de uma matriz triangular inferior por blocos sob algumas condições.

Teorema 2.40. *[29] Se $M \in \mathbb{C}^{n\times n}$ é triangular inferior por blocos*

$$M = \begin{pmatrix} B_1 & 0 \\ B_2 & N \end{pmatrix},$$

com $B_1 \in \mathbb{C}^{s\times s}$ *não singular e* $N \in \mathbb{C}^{t\times t}$ *uma matriz estritamente triangular inferior, então*

$$M_d = \begin{pmatrix} B_1^{-1} & 0 \\ XB_1^{-1} & 0 \end{pmatrix},$$

onde X *é solução da equação matricial* $XB_1 - NX = B_2$*. Além disso, se* r_i *é a i-ésima linha de* B_2*, então as linhas* x_i *de* X *são recursivamente dadas por*

$$\begin{aligned} x_1 &= r_1 B_1^{-1}, \\ x_i &= \left(r_i + \sum_{j=1}^{i-1} n_{ij} x_j \right) B_1^{-1}, \ \textit{para } i = 2, 3, \ldots, s. \end{aligned}$$

Demonstração:

Seja

$$Y = \begin{pmatrix} B_1^{-1} & 0 \\ XB_1^{-1} & 0 \end{pmatrix},$$

onde X é solução da equação matricial $XB_1 - NX = B_2$. Vejamos que Y satisfaz as equações da inversa de Drazin. Note que

$$M^p = \begin{pmatrix} B_1 & 0 \\ B_2 & N \end{pmatrix}^p = \begin{pmatrix} B_1^p & 0 \\ S(p) & N^p \end{pmatrix},$$

onde $S(p)$ depende de B_1, B_2 e N. Além disso, como N é estritamente triangular inferior, então $N^p = 0$ para $p \geq t$, assim temos

$$\begin{aligned} M^tYM &= \begin{pmatrix} B_1^t & 0 \\ S(t) & N^t \end{pmatrix} \begin{pmatrix} B_1^{-1} & 0 \\ XB_1^{-1} & 0 \end{pmatrix} \begin{pmatrix} B_1 & 0 \\ B_2 & N \end{pmatrix} \\ &= \begin{pmatrix} B_1^t & 0 \\ S(t) & 0 \end{pmatrix} \begin{pmatrix} I & 0 \\ X & 0 \end{pmatrix} = \begin{pmatrix} B_1^t & 0 \\ S(t) & 0 \end{pmatrix} = \begin{pmatrix} B_1^t & 0 \\ S(t) & N^t \end{pmatrix} = M^t. \end{aligned}$$

$$\begin{aligned} YMY &= \begin{pmatrix} B_1^{-1} & 0 \\ XB_1^{-1} & 0 \end{pmatrix} \begin{pmatrix} B_1 & 0 \\ B_2 & N \end{pmatrix} \begin{pmatrix} B_1^{-1} & 0 \\ XB_1^{-1} & 0 \end{pmatrix} \\ &= \begin{pmatrix} I & 0 \\ X & 0 \end{pmatrix} \begin{pmatrix} B_1^{-1} & 0 \\ XB_1^{-1} & 0 \end{pmatrix} = \begin{pmatrix} B_1^{-1} & 0 \\ XB_1^{-1} & 0 \end{pmatrix} = Y. \end{aligned}$$

Agora, como X é solução de $XB_1 - NX = B_2$, então $B_2 + NX = XB_1$ e como B_1 é não singular temos $B_2B_1^{-1} + NXB_1^{-1} = X$, assim

$$\begin{aligned} MY &= \begin{pmatrix} B_1 & 0 \\ B_2 & N \end{pmatrix} \begin{pmatrix} B_1^{-1} & 0 \\ XB_1^{-1} & 0 \end{pmatrix} = \begin{pmatrix} I & 0 \\ B_2B_1^{-1} + NXB_1^{-1} & 0 \end{pmatrix} \\ &= \begin{pmatrix} I & 0 \\ X & 0 \end{pmatrix} = \begin{pmatrix} B_1^{-1} & 0 \\ XB_1^{-1} & 0 \end{pmatrix} \begin{pmatrix} B_1 & 0 \\ B_2 & N \end{pmatrix} = YM, \end{aligned}$$

de onde segue que $M_d = Y$. Além disso, da expressão $B_2B_1^{-1} + NXB_1^{-1} = X$, e do fato de que N é estritamente triangular inferior segue que

$$\begin{aligned} x_1 &= r_1B_1^{-1}, \\ x_i &= \left(r_i + \sum_{j=1}^{i-1} n_{ij}x_j \right) B_1^{-1}, \quad i \geq 2. \end{aligned}$$

■

Com base nos resultados anteriores podemos construir o seguinte algoritmo para encontrar a inversa de Drazin.

1. Dada a matriz $A \in C^{n\times n}$, decomponha A em valores singulares

$$A = U \begin{pmatrix} \Sigma & 0 \\ 0 & 0 \end{pmatrix} V^*.$$

Se $\Sigma \in \mathbb{C}^{n\times n}$, então $A_d = A^{-1} = V\Sigma^{-1}U^*$, caso contrário escreva

$$V^*AV = V^*U \begin{pmatrix} \Sigma & 0 \\ 0 & 0 \end{pmatrix} V^*V = V^*U \begin{pmatrix} \Sigma & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} A_{11}^{(1)} & 0 \\ A_{21}^{(1)} & 0 \end{pmatrix}.$$

2. Agora decomponha $A_{11}^{(1)}$ em valores singulares

$$A_{11}^{(1)} = U_1 \begin{pmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{pmatrix} V_1^*.$$

Se $A_{11}^{(1)}$ é não singular vá para o passo 4, caso contrário escreva

$$V_1^*A_{11}^{(1)}V_1 = V_1^*U_1 \begin{pmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{pmatrix} V_1^*V_1 = V_1^*U_1 \begin{pmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} A_{11}^{(2)} & 0 \\ A_{21}^{(2)} & 0 \end{pmatrix}.$$

Assim,

$$\begin{pmatrix} V_1^* & 0 \\ 0 & I \end{pmatrix} V^*AV \begin{pmatrix} V_1 & 0 \\ 0 & I \end{pmatrix} = \left(\begin{array}{c|c} V_1^*A_{11}^{(1)}V_1 & 0 \\ \hline A_{21}^{(1)}V_1 & 0 \end{array} \right) = \left(\begin{array}{cc|c} A_{11}^{(2)} & 0 & 0 \\ A_{21}^{(2)} & 0 & 0 \\ \hline A_{31}^{(2)} & A_{32}^{(2)} & 0 \end{array} \right).$$

3. Continue decompondo $A_{11}^{(m)}$ em valores singulares e multiplicando

$$\begin{pmatrix} V_m^* & 0 \\ 0 & I \end{pmatrix}\begin{pmatrix} V_{m-1}^* & 0 \\ 0 & I \end{pmatrix}\cdots\begin{pmatrix} V_1^* & 0 \\ 0 & I \end{pmatrix}V^*AV\begin{pmatrix} V_1 & 0 \\ 0 & I \end{pmatrix}\cdots\begin{pmatrix} V_{m-1} & 0 \\ 0 & I \end{pmatrix}\begin{pmatrix} V_m & 0 \\ 0 & I \end{pmatrix}$$

   para obter $A_{11}^{(m+1)}$. Se $A_{11}^{(k)} = 0$ então A é nilpotente e $A_d = 0$. Se $A_{11}^{(k)}$ é não singular, então vá para o passo 4.

4. Nos passos anteriores construímos a matriz

$$WAW^* = \left(\begin{array}{c|cccc} A_{11}^{(k)} & 0 & \cdots & \cdots & 0 \\ \hline A_{21}^{(k)} & 0 & \cdots & \cdots & 0 \\ A_{31}^{(k)} & A_{32}^{(k)} & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \vdots \\ A_{k+1,1}^{(k)} & A_{k+1,2}^{(k)} & \cdots & A_{k+1,k}^{(k)} & 0 \end{array}\right) = \left(\begin{array}{c|c} B_1 & 0 \\ \hline B_2 & N \end{array}\right),$$

   onde B_1 é não singular, N é estritamente triangular inferior e W é uma matriz unitária dada por

$$W = \begin{pmatrix} V_{k-1}^* & 0 \\ 0 & I \end{pmatrix}\begin{pmatrix} V_{k-2}^* & 0 \\ 0 & I \end{pmatrix}\cdots\begin{pmatrix} V_1^* & 0 \\ 0 & I \end{pmatrix}V^*.$$

Segue do Lema 2.38 que $A_d = W^*(WAW^*)_d W$, e como WAW^* satisfaz as condições do Teorema 2.40, temos que

$$A_d = W^*(WAW^*)_d W = W^*\begin{pmatrix} B_1^{-1} & 0 \\ XB_1^{-1} & 0 \end{pmatrix}W,$$

onde X é a solução de $XB_1 - NX = B_2$, que é dada por

$$x_1 = r_1 B_1^{-1}, \qquad x_i = \left(r_i + \sum_{j=1}^{i-1} n_{ij}x_j\right)B_1^{-1},$$

onde r_i é a i-ésima linha de B_2 e x_i é a i-ésima linha de X.

Capítulo 3

Novos métodos computacionais para inversas generalizadas baseados em decomposição conjugada

Neste capítulo apresentamos os métodos diretos propostos neste trabalho, os quais consistem em um método direto para encontrar a inversa de Moore-Penrose, e um método direto para encontrar a inversa de Drazin, ambos baseados na decomposição conjugada reduzida de matrizes, a qual será definida adiante.

3.1 Decomposição Conjugada

A decomposição conjugada foi proposta por Wang e Yuan em [30], a qual consiste em decompor uma matriz $A \in \mathbb{C}^{m\times n}$ com posto r em $A = Q\Gamma Z^{-1}$ com $Q \in \mathbb{C}^{m\times m}$ unitária, $\Gamma \in \mathbb{C}^{m\times n}$ com elementos não nulos apenas nos primeiros r elementos da diagonal, e $Z \in \mathbb{C}^{n\times n}$ não singular. Alternativamente, A pode ser decomposta em $A = S^{-1}\Gamma U$ onde $S \in \mathbb{C}^{m\times m}$ é não singular, $\Gamma \in \mathbb{C}^{m\times n}$ possui elementos não nulos apenas nos primeiros r elementos da diagonal, e $U \in \mathbb{C}^{n\times n}$ é unitária.

Para os métodos propostos neste trabalho, usaremos a decomposição conjugada em um formato um pouco diferente do proposto em [30], conforme sugerido pelos resultados a seguir.

Definição 3.1. [30] Seja $A \in \mathbb{C}^{n\times n}$ uma matriz Hermitiana positiva definida. Temos que $z_1, z_2, \ldots, z_n$, com $z_i \in \mathbb{C}^n$ $i = 1, 2, \ldots n$, são vetores conjugados de A se satisfazem

$$\begin{cases} z_i^* A z_j = 0 & \text{se} \quad i \neq j \\ z_i^* A z_j > 0 & \text{se} \quad i = j \end{cases}$$

Lema 3.2. *[30] Se $A \in \mathbb{C}^{n\times n}$ é uma matriz Hermitiana positiva semidefinida, com posto$(A) = r$, então existe uma matriz não singular $Z \in \mathbb{C}^{n\times n}$ tal que*

$$Z^*AZ = D$$

sendo $D =$diag$(z_1^*Az_1, z_2^*Az_2, \ldots, z_r^*Az_r, 0, \ldots, 0)$, e $z_1^*Az_1 \geq z_2^*Az_2 \geq \ldots \geq z_r^*Az_r > 0$.

Demonstração:

Como A é Hermitiana positiva semidefinida, $\langle x, y\rangle = x^*Ay$ define um pseudo produto interno. Seja $\{u_1, u_2, \ldots, u_r\}$ uma base para $\mathcal{R}(A)$. Aplicando o processo de Gram-Schmidt com o pseudo produto interno $\langle x, y\rangle = x^*Ay$ em $\{u_1, u_2, \ldots, u_r\}$ obtemos $\{z_1, z_2, \ldots, z_r\}$ satisfazendo

$$\begin{cases} z_i^*Az_j = 0, & \text{se} \quad i \neq j, \\ z_i^*Az_j > 0, & \text{se} \quad i = j. \end{cases}$$

Agora, tome $n-r$ vetores $\{z_{r+1}, z_{r+2}, \ldots, z_n\}$ de $\mathbb{C}^n$ de forma que $\{z_{r+1}, z_{r+2}, \ldots, z_n\}$ seja uma base para $\mathcal{N}(A)$ e defina $Z = (z_1, z_2, \ldots, z_n)$. Temos que Z é não singular e

$$Z^*AZ = \begin{pmatrix} D_1 & 0 \\ 0 & 0 \end{pmatrix},$$

com $D_1 =$diag$(z_1^*Az_1, z_2^*Az_2, \ldots, z_r^*Az_r)$. ■

Teorema 3.3. *Se $A \in \mathbb{C}^{m\times n}$ com $m \geq n$ e posto$(A) = r$, então existe uma matriz não singular $Z \in \mathbb{C}^{n\times n}$ e uma matriz $Q \in \mathbb{C}^{m\times n}$ satisfazendo $Q^*Q = I_n$, tal que $A = Q\Gamma Z^{-1}$, com $\Gamma =$ diag$(\gamma_1, \gamma_2, \ldots, \gamma_n) \in \mathbb{C}^{n\times n}$, $\gamma_1 \geq \gamma_2 \geq \ldots \geq \gamma_r > 0 = \ldots = \gamma_n$, e $\gamma_i = \sqrt{z_i^*A^*Az_i}$, $i = 1, 2, \ldots, r$.*

Demonstração:

Como A^*A é Hermitiana e positiva semidefinida, pelo Lema 3.2 temos que existe $Z = [z_1, z_2, \ldots, z_n] \in \mathbb{C}^{n\times n}$ não singular tal que

$$Z^*A^*AZ = D, \quad D = \text{diag}(z_1^*A^*Az_1, z_2^*A^*Az_2, \ldots, z_r^*A^*Az_r, 0, \ldots, 0).$$

Defina $\gamma_i = \sqrt{z_i^*A^*Az_i}$ para $i = 1, 2, \ldots, r$, $\Gamma = \text{diag}(\gamma_1, \gamma_2, \ldots, \gamma_r, 0 \ldots, 0) \in \mathbb{C}^{n\times n}$, e $q_i = \frac{1}{\gamma_i}Az_i \in \mathbb{C}^m$, $i = 1, 2, \ldots, r$. Note que

$$q_i^*q_j = \frac{1}{\sqrt{z_i^*A^*Az_i}\sqrt{z_j^*A^*Az_j}} z_i^*A^*Az_j = \begin{cases} 1 & \text{se} \quad i = j, \\ 0 & \text{se} \quad i \neq j. \end{cases} \qquad i, j = 1, 2, \ldots, r.$$

Então, q_i, $i = 1, 2, \ldots, r$ são vetores ortonormais.

Agora, escolha $n-r$ vetores $q_{r+1},\ldots,q_n \in \mathbb{C}^m$ de forma que $q_1,\ldots,q_n$ sejam ortonormais. Como $z_i \in \mathcal{N}(A)$, $i = r+1,\ldots,n$, então

$$A[z_1, z_2, \ldots, z_n] = [q_1, q_2, \ldots, q_n]\Gamma,$$

o que resulta em $AZ = Q\Gamma$, e portanto $A = Q\Gamma Z^{-1}$, com

$$Q^*Q = \begin{bmatrix} q_1^* \\ q_2^* \\ \vdots \\ q_n^* \end{bmatrix} [q_1, q_2, \ldots, q_n] = \begin{bmatrix} q_1^*q_1 & q_1^*q_2 & \cdots & q_1^*q_n \\ q_2^*q_1 & q_2^*q_2 & \cdots & q_2^*q_n \\ \vdots & \vdots & \ddots & \vdots \\ q_n^*q_1 & q_n^*q_2 & \cdots & q_n^*q_n \end{bmatrix} = I_n.$$

■

Teorema 3.4. *Se $A \in \mathbb{C}^{m\times n}$, com $m \le n$ e posto$(A) = r$, então existe uma matriz não singular $S \in \mathbb{C}^{m\times m}$, e uma matriz $U \in \mathbb{C}^{m\times n}$ satisfazendo $UU^* = I_m$ tal que $A = S^{-1}\Gamma U$, com $\Gamma \in \mathbb{C}^{m\times m}$, $\Gamma = diag(\gamma_1, \gamma_2, \ldots, \gamma_r, 0, \ldots, 0)$, $\gamma_1 \ge \gamma_2 \ge \ldots \ge \gamma_r > 0 = \ldots = \gamma_m$, e $\gamma_i = \sqrt{s_i^* AA^* s_i}$, $i = 1, 2, \ldots, r$.*

Demonstração:

Aplicando o Teorema 3.3 em A^* segue que $A^* = Q\Gamma Z^{-1}$, assim

$$A = (A^*)^* = (Q\Gamma Z^{-1})^* = (Z^{-1})^*\Gamma^* Q^* = (Z^*)^{-1}\Gamma Q^*.$$

Considere $S = Z^*$ e $U = Q^*$, então

$$A = S^{-1}\Gamma U \quad \text{e} \quad UU^* = Q^*(Q^*)^* = Q^*Q = I_m$$

. ■

As decomposições mostradas nos Teoremas 3.3 e 3.4 são chamadas de decomposição conjugada reduzida de A. Tal decomposição pode ser executada de diversas maneiras, aqui consideramos duas delas. Um caminho é usar o processo de ortonormalização de Gram-Schmidt modificado com o pseudo produto interno $\langle x, y\rangle_{A^*A} = x^*A^*Ay$, o que é estável no sentido de manter a matriz Q (ou U quando $m < n$) com colunas ortonormais, porém esse processo utiliza $z_1, z_2, \ldots, z_i$ para encontrar z_{i+1}. Outra maneira de executar a decomposição conjugada é pelo método de gradientes conjugados aplicado no sistema $A^*Ax = b$, para algum vetor b, esse processo tem a vantagem de usar apenas o vetor z_i para computar z_{i+1}, porém esse processo não é estável, pois conforme aumentamos o tamanho das matrizes, as colunas de Q deixam de ser ortonormais. Os algoritmos 2 e 3 mostram como obter a decomposição conjugada reduzida via processo de ortonormalização de Gram-Schmidt modificado (GSM) e via método de gradientes conjugados (GC) respectivamente.

Algoritmo 2. Decomposição conjugada reduzida via GSM.

Dada $A \in \mathbb{C}^{m\times n}$ com $m \geq n$, escolha $p_1, p_2, \ldots, p_n$ em $\mathbb{C}^n$ L.I. e uma tolerância tol

$k_1 = 0, \quad k_2 = 0$

Para $i = 1$ até n faça

Para $j = 1$ até $i - 1$ faça:

Se $g_j > tol$ faça:

$p_i = p_i - \frac{p_i^* A^* A p_j}{g_j} p_j$

fim do se

fim do para

$g_i = p_i^* A^* A p_i$

Se $g_i > tol$

$k_1 = k_1 + 1, \quad \gamma_{k_1} = g_i, \quad q_{k_1} = \frac{1}{\gamma_{k_1}} A p_i, \quad z_{k_1} = p_i$

Se não, faça:

$k_2 = k_2 + 1, \quad \tilde{z}_{k_2} = p_i$

Fim do se

Fim do para

$Z = [z_1, z_2, \ldots, z_{k_1}, \tilde{z}_1, \tilde{z}_2, \ldots, \tilde{z}_{k_2}]$

Escolha $q_{k_1+1}, q_{k_1+2}, \ldots, q_n$ tal que $q_1, q_2, \ldots, q_n$ seja L.I.

Para $i = k_1 + 1$ até n faça:

Para $j = k_1 + 1$ até $i - 1$ faça:

$q_i = q_i - \frac{q_i^* q_j}{q_i^* q_i} q_j$

Fim do para

Fim do para

Algoritmo 3. Decomposição conjugada reduzida via GC.

Dada $A \in \mathbb{C}^{m\times n}$ com $m \geq n$, escolha $r_1 \neq 0$ e uma tolerância tol

$z_1 = r_1, \quad \gamma_1 = \sqrt{z_1 A^* A z_1}, \quad q_1 = \frac{1}{\gamma_1} A z_1, \quad k = 0$

Para $i = 1$ até n faça:

$r_{i+1} = r_i - \frac{r_i^* r_i}{z_i^* A^* A z_i} A^* A z_i, \quad d = r_{i+1} + \frac{r_{i+1}^* r_{i+1}}{r_i^* r_i} z_i, \quad g = d^* A^* A d$

Se $g > tol$

$z_{i+1} = d, \quad \gamma_{i+1} = \sqrt{g}, \quad q_{i+1} = \frac{1}{\gamma_{i+1}} A z_{i+1}, \quad k = k + 1$

Se não

Pare

Fim do se

Fim do Para

Para $i = k + 1$ até n faça:

Para $j = k + 1$ até $i - 1$ faça:

$q_i = q_i - \frac{q_i^* q_j}{q_i^* q_i} q_j$

Fim do para

Fim do para

Escolha $p_{k+1}, p_{k+2}, \ldots, p_n$ em $\mathcal{N}(A)$

3.2 Inversa de Moore-Penrose via decomposição conjugada

3.2.1 O método

No capítulo anterior vimos que em geral os métodos diretos para encontrar a inversa de Moore-Penrose são baseados em decomposições. Nesta seção apresentaremos um método direto para encontrar a inversa de Moore-Penrose via decomposição conjugada. O próximo teorema mostra que a inversa de Moore-Penrose de uma matriz de posto completo pode ser construída a partir da decomposição conjugada reduzida.

Teorema 3.5. *Seja $A \in \mathbb{C}^{m\times n}$, com $m \geq n$ e posto$(A) = n$, e seja $A = Q\Gamma Z^{-1}$ a decomposição conjugada reduzida de A. Temos que*

$$A^{\dagger} = Z\Gamma^{-1}Q^{*},$$

Demonstração:

Considere $A = Q\Gamma Z^{-1}$ a decomposição conjugada reduzida de A, defina $X = Z\Gamma^{-1}Q^{*}$. Precisamos verificar que X é solução das equações de Penrose (1.1)-(1.4). De fato, como posto$(A) = n$, então posto$(\Gamma) = n$, de onde segue que Γ é não singular, assim

$$\begin{aligned}
AXA &= AZ\Gamma^{-1}Q^{*}Q\Gamma Z^{-1} = AZ\Gamma^{-1}\Gamma Z = AZ^{-1}Z = A;\\
XAX &= Z\Gamma^{-1}Q^{*}Q\Gamma Z^{-1}X = Z\Gamma^{-1}\Gamma Z^{-1}X = ZZ^{-1}X = X;\\
(AX)^{*} &= (Q\Gamma Z^{-1}Z\Gamma^{-1}Q^{*})^{*} = (Q\Gamma\Gamma^{-1}Q^{*})^{*} = (QQ^{*})^{*} = QQ^{*} = AX;\\
(XA)^{*} &= (Z\Gamma^{-1}Q^{*}Q\Gamma Z^{-1})^{*} = (Z\Gamma^{-1}\Gamma Z^{-1})^{*} = (ZZ^{-1})^{*} = I_n^{*} = I_n = XA.
\end{aligned}$$

De onde temos que

$$A^{\dagger} = Z\Gamma^{-1}Q^{*}.$$

■

Note que consideramos $m \geq n$, porém caso tenhamos $m < n$ segue do fato de que $A^{\dagger} = ((A^{*})^{\dagger})^{*}$ e do Teorema 3.4 que $A^{\dagger} = U^{*}\Gamma^{-1}S$.

O teorema anterior mostra que $A^{\dagger} = Z\Gamma^{-1}Q^{*}$ para o caso em que A tem posto completo, para o caso em que A tem posto incompleto, em geral temos $A^{\dagger}$ diferente de $Z\Gamma^{-1}Q^{*}$. Pórem, com o uso de mais uma decomposição conjugada é possível construir a inversa de Moore-Penrose de uma matriz com posto incompleto, como mostrado no teorema abaixo.

Teorema 3.6. *Seja $A \in \mathbb{C}^{m\times n}$, com $m \geq n$ e posto$(A) = r < n$. Se*

$$A = Q\Gamma Z^{-1}$$

é a decomposição conjugada reduzida de A, então

$$A^{\dagger} = (U_1^* \Gamma_1^{-1} S_1, 0) Q^*,$$

onde $S_1^{-1}\Gamma_1 U_1$ *é a decomposição conjugada reduzida da submatriz formada pelas* r *primeiras linhas de* ΓZ^{-1}.

Demonstração:

Como $Q^{\dagger} = Q^*$, temos que

$$\begin{aligned} Q^{\dagger}Q(\Gamma Z^{-1})(\Gamma Z^{-1})^* &= Q^*Q(\Gamma Z^{-1})(\Gamma Z^{-1})^* = I_n(\Gamma Z^{-1})(\Gamma Z^{-1})^* = (\Gamma Z^{-1})(\Gamma Z^{-1})^* I_n \\ &= (\Gamma Z^{-1})(\Gamma Z^{-1})^* Q^* Q = (\Gamma Z^{-1})(\Gamma Z^{-1})^* Q^{\dagger} Q \end{aligned}$$

e

$$(\Gamma Z^{-1})(\Gamma Z^{-1})^{\dagger} Q^* Q = (\Gamma Z^{-1})(\Gamma Z^{-1})^{\dagger} I_n = I_n (\Gamma Z^{-1})(\Gamma Z^{-1})^{\dagger} = Q^* Q (\Gamma Z^{-1})(\Gamma Z^{-1})^{\dagger},$$

portanto segue da proposição 2.14 que $A^{\dagger} = (\Gamma Z^{-1})^{\dagger} Q^{\dagger} = (\Gamma Z^{-1})^{\dagger} Q^*$.

Porém, como posto$(A) = r < n$, então

$$\Gamma = \begin{pmatrix} \widetilde{\Gamma} & 0 \\ 0 & 0 \end{pmatrix},$$

sendo $\widetilde{\Gamma} \in \mathbb{C}^{r\times r}$ uma matriz diagonal, com posto$(\widetilde{\Gamma}) = r$, e como Z é não singular, então

$$\Gamma Z^{-1} = \begin{pmatrix} \widetilde{\Gamma} & 0 \\ 0 & 0 \end{pmatrix} Z^{-1} = \begin{pmatrix} X \\ 0 \end{pmatrix},$$

onde $X \in \mathbb{C}^{r\times n}$ com posto$(X) = r$, vejamos que $(\Gamma Z^{-1})^{\dagger} = (X^{\dagger}, 0)$. De fato, basta mostrar que $(X^{\dagger}, 0)$ satisfaz as equações de Penrose (1.1)-(1.4).

$$\begin{aligned} \begin{pmatrix} X \\ 0 \end{pmatrix} (X^{\dagger}, 0) \begin{pmatrix} X \\ 0 \end{pmatrix} &= \begin{pmatrix} XX^{\dagger} & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} X \\ 0 \end{pmatrix} = \begin{pmatrix} XX^{\dagger}X \\ 0 \end{pmatrix} = \begin{pmatrix} X \\ 0 \end{pmatrix}; \\ (X^{\dagger}, 0) \begin{pmatrix} X \\ 0 \end{pmatrix} (X^{\dagger}, 0) &= X^{\dagger}X(X^{\dagger}, 0) = (X^{\dagger}XX^{\dagger}, 0) = (X^{\dagger}, 0); \\ \left[\begin{pmatrix} X \\ 0 \end{pmatrix} (X^{\dagger}, 0) \right]^* &= \left[\begin{pmatrix} XX^{\dagger} & 0 \\ 0 & 0 \end{pmatrix} \right]^* = \begin{pmatrix} (XX^{\dagger})^* & 0 \\ 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} XX^{\dagger} & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} X \\ 0 \end{pmatrix} (X^{\dagger}, 0); \\ \left[(X^{\dagger}, 0) \begin{pmatrix} X \\ 0 \end{pmatrix} \right]^* &= (X^{\dagger}X)^* = X^{\dagger}X = (X^{\dagger}, 0) \begin{pmatrix} X \\ 0 \end{pmatrix}. \end{aligned}$$

Agora, considere a decomposição conjugada reduzida de X dada por $X = S_1^{-1}\Gamma_1 U_1$, como X tem posto completo então $X^\dagger = U_1^*\Gamma_1^{-1}S_1$, portanto

$$(\Gamma Z^{-1})^\dagger = \begin{pmatrix} X \\ 0 \end{pmatrix}^\dagger = (X^\dagger, 0) = (U_1^*\Gamma_1^{-1}S_1, 0),$$

de onde segue que

$$A^\dagger = (\Gamma Z^{-1})^\dagger Q^* = (U_1^*\Gamma_1^{-1}S_1, 0)Q^*.$$

■

O teorema acima descreve um método direto para encontrar a inversa de Moore-Penrose usando decomposição conjugada reduzida. Vale ressaltar que da maneira como exposto no Teorema 3.6, não se faz necessário computar as n colunas da matriz Q, basta que sejam calculadas as r primeiras colunas. Dessa forma o produto Q^*A fornece a matriz de posto completo desejada. Além disso, não é necessário conhecer as $n - r$ últimas colunas de Q para calcular $A^\dagger = (U_1^*\Gamma_1^{-1}S_1, 0)Q^*$. O Algoritmo 4 descreve o método direto mostrado nos Teoremas 3.5 e 3.6

Algoritmo 4. Inversa de Moore-Penrose via decomposição conjugada

Dada $A \in \mathbb{C}^{m\times n}$ com $m \geq n$ e posto$(A) = r$

Compute $Q \in \mathbb{C}^{m\times r}$, $\Gamma \in \mathbb{C}^{r\times r}$, $Z \in \mathbb{C}^{m\times r}$ via decomposição conjugada reduzida de A

Se $r < n$ então faça:

  $X = Q^*A$

  Compute a decomposição conjugada reduzida de X, obtendo U_1, Γ_1 e S_1

  Compute $X^\dagger = U_1^*\Gamma_1^{-1}S_1$

  Compute $A^\dagger = X^\dagger Q^*$

Se não, faça:

  Compute $A^\dagger = Z\Gamma^{-1}Q^*$

Fim do se

3.2.2 Exemplos numéricos

Os testes abaixo foram realizados usando o software MATLAB.

Exemplo 3.1. Neste exemplo, consideramos algumas matrizes aleatórias, de posto completo e incompleto, geradas pelo comando “rand” do MATLAB, e computamos a inversa de Moore-Penrose de tais matrizes usando a decomposição conjugada executada com o método de gradientes conjugados. A tabela 3.1 mostra o erro obtido para cada matriz, comparado na norma-2 com $A^\dagger$ obtida via SVD. Além disso analisamos a ortonormalidade da matriz Q obtida na decomposição conjugada através da norma-2 da diferença $Q^*Q - I_n$, onde I_n denota a matriz identidade de ordem n.

$m \times n$	posto	$\|X - A^\dagger\|_2$	$\|Q^*Q - I_n\|_2$
8×4	4	$5,91 \times 10^{-11}$	$2,02 \times 10^{-11}$
8×4	2	$9,39 \times 10^{-16}$	$2,75 \times 10^{-15}$
20×10	10	$1,18$	$0,99$
20×10	5	$3,01 \times 10^{+15}$	$2,90$
80×50	50	$0,88$	$6,00$
80×50	30	$4,69 \times 10^{+14}$	$13,10$

Tabela 3.1: Erro cometido ao aproximar $A^\dagger$ via decomposição conjugada executada com o método de gradientes conjugados.

Como podemos notar, quando aumentamos o tamanho da matriz, rapidamente as colunas da matriz Q perdem a ortonormalidade, com isso $Q^\dagger$ difere de Q^*, e consequentemente a matriz $Z\Gamma^{-1}Q^*$ difere de $A^\dagger$.

Exemplo 3.2. Neste exemplo, assim como no exemplo anterior, consideramos matrizes aleatórias, e computamos a inversa de Moore-Penrose de tais matrizes usando a decomposição conjugada executada com o processo de ortonormalização de Gram-Schmidt modificado. A tabela 3.2 mostra o erro obtido para cada matriz, comparado na norma-2 com $A^\dagger$ obtida via SVD, e a norma-2 da diferença $Q^*Q - I_n$.

$m \times n$	posto	$\|X - A^\dagger\|_2$	$\|Q^*Q - I_n\|_2$
8×6	6	$7,76 \times 10^{-15}$	$2,74 \times 10^{-15}$
8×6	3	$4,95 \times 10^{-16}$	$1,47 \times 10^{-15}$
50×30	30	$5,09 \times 10^{-14}$	$5,61 \times 10^{-14}$
50×30	20	$6,89 \times 10^{-14}$	$4,90 \times 10^{-14}$
100×70	70	$2,79 \times 10^{-12}$	$2,16 \times 10^{-12}$
100×70	40	$2,67 \times 10^{-13}$	$2,01 \times 10^{-13}$
800×500	500	$1,14 \times 10^{-10}$	$3,56 \times 10^{-10}$
800×500	200	$8,56 \times 10^{-14}$	$1,26 \times 10^{-11}$
4000×2000	2000	$3,29 \times 10^{-11}$	$3,06 \times 10^{-10}$
4000×2000	1200	$2,97 \times 10^{-12}$	$7,75 \times 10^{-11}$

Tabela 3.2: Erro cometido ao aproximar $A^\dagger$ via decomposição conjugada executada com o processo de ortonormalização de Gram-Schmidt modificado.

Como podemos notar, mesmo para tamanhos maiores, as colunas de Q se mantém ortonormais, e as aproximações de $A^\dagger$ obtidas pela decomposição conjugada apresentam erro pequeno.

Nos exemplos abaixo usamos o processo de ortonormalização de Gram-Schmidt modificado para obter a decomposição conjugada reduzida, nesse processo, γ_i é considerado nulo se $\gamma_i < tol$, para alguma tolerância tol estabelecida. A menos que seja especificado, a tolerância padrão utilizada foi 10^{-8}.

Exemplo 3.3. Encontre a inversa de Moore-Penrose da matriz

$$A = \begin{pmatrix} 3 & -2 & 1 & 5 \\ 1 & -2 & -1 & 3 \\ 1 & 4 & 5 & -3 \\ 2 & 0 & 2 & 2 \\ 4 & 2 & 6 & 2 \\ 2 & -1 & 1 & 3 \end{pmatrix}.$$

Calculando uma decomposição conjugada reduzida de A até o posto de A, obtemos

$$Q = \begin{pmatrix} 0,3753 & -0,5287 \\ 0,0564 & -0,4146 \\ 0,3656 & 0,6581 \\ 0,3189 & -0,1141 \\ 0,7410 & 0,1294 \\ 0,2674 & -0,2929 \end{pmatrix}, \Gamma = \begin{pmatrix} 9,2165 & 0 \\ 0 & 0,3350 \end{pmatrix} \text{ e } Z = \begin{pmatrix} 0,0758 & 0,8480 \\ 0,3299 & 0,4857 \\ 0,7695 & -0,6464 \\ 0,6244 & -0,2206 \end{pmatrix}.$$

Isso nos mostra que posto$(A) = 2$, portanto precisamos de mais uma decomposição conjugada reduzida. Computando Q^*A obtemos

$$Q^*A = \begin{pmatrix} 5,6846 & 1,8136 & 7,4982 & 3,8709 \\ -1,6388 & 5,0706 & 3,4317 & -6,7094 \end{pmatrix} = \begin{pmatrix} Y \\ 0 \end{pmatrix},$$

com $Y \in \mathbb{C}^{2\times 4}$. Agora, calculamos a decomposição conjugada $Y = S_1^{-1}\Gamma_1 U_1$, obtendo

$$S_1 = \begin{pmatrix} 0,6724 & 0,3610 \\ -0,1825 & 0,4287 \end{pmatrix}, \Gamma_1 = \begin{pmatrix} 7,6952 & 0 \\ 0 & 4,3895 \end{pmatrix}$$

e

$$U_1 = \begin{pmatrix} 0,4198 & 0,3964 & 0,8162 & 0,0234 \\ -0,3964 & 0,4198 & 0,0234 & -0,8162 \end{pmatrix}.$$

Portanto, a inversa de Moore-Penrose de A é dada por

$$A^\dagger = (U_1^*\Gamma_1^{-1}S_1)Q^* = \begin{pmatrix} 0,0300 & 0,0109 & 0,0069 & 0,0191 & 0,0369 & 0,0198 \\ -0,0251 & -0,0237 & 0,0455 & -0,0013 & 0,0204 & -0,0129 \\ 0,0049 & -0,0129 & 0,0524 & 0,0178 & 0,0574 & 0,0069 \\ 0,0551 & 0,0346 & -0,0386 & 0,0204 & 0,0165 & 0,0326 \end{pmatrix},$$

com resíduo $\max\{\|AXA-A\|_2, \|XAX-X\|_2, \|(AX)^*-AX\|_2, \|(XA)^*-XA\|_2\} = 3,76\times 10^{-15}$.

Para o próximo exemplo, vamos considerar o seguinte problema teste:

Problema Teste 3.1. [14]. Neste problema teste consideramos uma equação de Fredholm da forma

$$g(s) = \int_a^b K(t,s) f(t) dt. \tag{3.1}$$

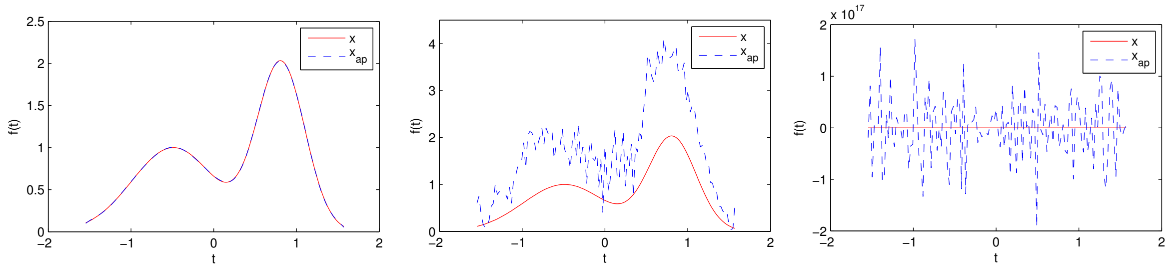
com o núcleo $K(t,s)$ e a solução $f(t)$ dadas por

$$K(t,s) = (\cos(s) + \cos(t)) \left(\frac{\operatorname{sen}(u)}{u} \right)^2 \quad \text{e} \quad f(t) = a_1 e^{-c_1(t-t_1)^2} + a_2 e^{-c_2(t-t_2)^2},$$

onde $u = \pi(\operatorname{sen}(s) + \operatorname{sen}(t))$, $a_i, c_i, t_i \in \mathbb{R}$ para $i = 1, 2$ e $t, s \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$. Neste caso foram usados os seguintes valores: $a_1 = 2$; $a_2 = 1$; $c_1 = 6$; $c_2 = 2$; $t_1 = 0,8$; $t_2 = -0,5$. A rotina do MATLAB shaw.m de [14] fornece: uma matriz $A \in \mathbb{C}^{n\times n}$ da discretização de $K(t,s)$ mal condicionada, a solução exata x baseada em $f(t)$, e o lado direito b da equação construído a partir de $b = Ax$.

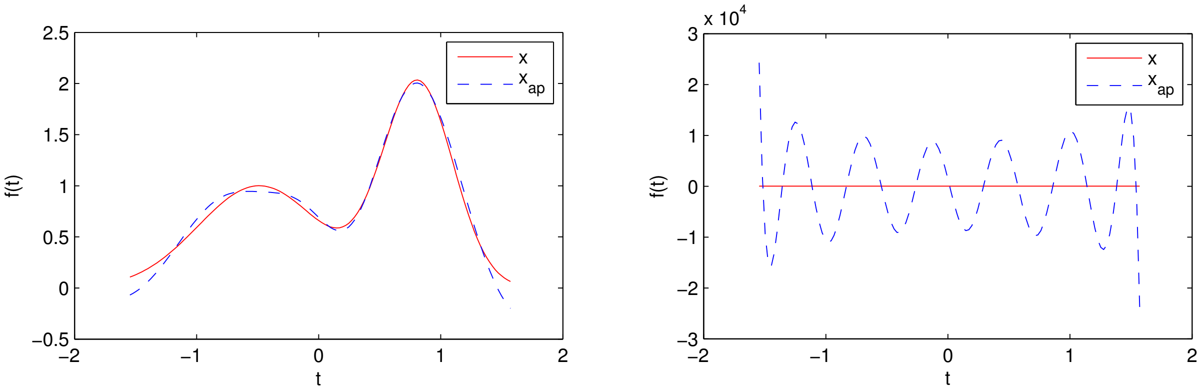
Exemplo 3.4. Neste exemplo resolvemos o Problema Teste 3.1 usando o método baseado em decomposição conjugada. A, b e x são obtidos a partir da rotina shaw.m de [14] para $n = 128$. Neste caso a matriz A obtida é muito mal condicionada (número de condição de A: $2,46 \times 10^{20}$). X é calculada via decomposição conjugada. A figura 3.1 mostra a solução exata x e a aproximação calculada $x_{ap} = Xb$ para os seguintes valores de *tol*: $tol = 10^{-16}$, $tol = 10^{-30}$ e $tol = 10^{-32}$ respectivamente. Neste caso o algoritmo identificou os seguintes valores para o posto de A: $r = 13$, $r = 25$ e $r = 128$ respectivamente. Sabemos previamente que posto$(A) = 128$, porém o que podemos perceber é que quando tentamos calcular a inversa de Moore-Penrose de A com valores muito pequenos para γ_i, o processo se torna instável. Ao considerar uma tolerância maior, na prática aproximamos a matriz A por uma matriz B, cujo posto$(B) <$posto(A), e calculamos $B^\dagger$, então consideramos que $A^\dagger \approx B^\dagger$, o que pode ser usado para obter boas aproximações da solução exata x, como mostra a figura 3.1.

Ainda no mesmo exemplo, considerando que em aplicações envolvendo matrizes mal condicionadas, em geral o vetor b contém pequenas perturbações, testamos também a solução do sistema $Ax = \tilde{b}$, onde $\tilde{b} = b + e$, com um determinado nível de ruído, representado por NL (*Noise level*). A figura 3.2 compara a solução exata x com a solução aproximada $x_{ap} = X\tilde{b}$, onde X é obtido pela decomposição conjugada usando $tol = 10^{-6}$ e $tol = 10^{-16}$ respectivamente. Note que a mesma tolerância que capturou uma boa aproximação para o caso de b sem ruídos não obteve sucesso nesse caso, porém com uma tolerância maior foi possível capturar melhores aproximações.



(a) $tol = 10^{-16}$. (b) $tol = 10^{-30}$. (c) $tol = 10^{-32}$.

Figura 3.1: Solução exata do Problema Teste 3.1 e solução aproximada construída por Xb, onde X é calculada por decomposição conjugada com tolerâncias: 10^{-16}, 10^{-30} e 10^{-32} respectivamente.



(a) $tol = 10^{-6}$. (b) $tol = 10^{-16}$.

Figura 3.2: Solução exata do Problema Teste 3.1 e solução aproximada construída por $X\tilde{b}$, onde X é calculada por decomposição conjugada com tolerâncias: 10^{-6} e 10^{-16} respectivamente, e $\tilde{b} = b + e$ com NL$= 0,0010$.

O Exemplo 3.4 mostra que o método baseado em decomposição conjugada pode ser usado em problemas mal condicionados, desde que consideremos uma tolerância adequada para identificar quando γ_i é nulo.

3.3 Inversa de Drazin via decomposição conjugada

3.3.1 O método

No capítulo anterior vimos um método direto para calcular a inversa de Drazin baseado em SVD. Considerando uma matriz $A \in \mathbb{C}^{n\times n}$ com $\text{Ind}(A) = k$, tal método faz uso de k decomposições em valores singulares. Nesta seção propomos um método direto para encontrar a inversa de Drazin baseado em decomposição conjugada reduzida, o método é descrito abaixo:

1. Dada a matriz $A \in C^{n\times n}$, faça a decomposição conjugada de A

$$A = S^{-1}\begin{pmatrix} \Gamma & 0 \\ 0 & 0 \end{pmatrix} U.$$

Se $\Gamma \in \mathbb{C}^{n\times n}$, então $A_d = A^{-1} = U^*\Gamma^{-1}S$, caso contrário escreva

$$UAU^* = US^{-1}\begin{pmatrix} \Gamma & 0 \\ 0 & 0 \end{pmatrix} UU^* = US^{-1}\begin{pmatrix} \Gamma & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} A_{11}^{(1)} & 0 \\ A_{21}^{(1)} & 0 \end{pmatrix}.$$

2. Agora faça a decomposição conjugada de $A_{11}^{(1)}$

$$A_{11}^{(1)} = S_1^{-1}\begin{pmatrix} \Gamma_1 & 0 \\ 0 & 0 \end{pmatrix} U_1.$$

Se $A_{11}^{(1)}$ é não singular vá para o passo 4, caso contrário escreva

$$U_1A_{11}^{(1)}U_1^* = U_1S^{-1}\begin{pmatrix} \Gamma_1 & 0 \\ 0 & 0 \end{pmatrix} U_1U_1^* = U_1S_1^{-1}\begin{pmatrix} \Gamma_1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} A_{11}^{(2)} & 0 \\ A_{21}^{(2)} & 0 \end{pmatrix}.$$

Assim,

$$\begin{pmatrix} U_1 & 0 \\ 0 & I \end{pmatrix} UAU^* \begin{pmatrix} U_1^* & 0 \\ 0 & I \end{pmatrix} = \left(\begin{array}{c|c} U_1A_{11}^{(1)}U_1^* & 0 \\ \hline A_{21}^{(1)}U_1^* & 0 \end{array}\right) = \left(\begin{array}{cc|c} A_{11}^{(2)} & 0 & 0 \\ A_{21}^{(2)} & 0 & 0 \\ \hline A_{31}^{(2)} & A_{32}^{(2)} & 0 \end{array}\right).$$

3. Continue fazendo a decomposição conjugada de $A_{11}^{(m)}$ e multiplicando

$$\begin{pmatrix} U_m & 0 \\ 0 & I \end{pmatrix}\begin{pmatrix} U_{m-1} & 0 \\ 0 & I \end{pmatrix}\cdots\begin{pmatrix} U_1 & 0 \\ 0 & I \end{pmatrix} UAU^* \begin{pmatrix} U_1^* & 0 \\ 0 & I \end{pmatrix}\cdots\begin{pmatrix} U_{m-1}^* & 0 \\ 0 & I \end{pmatrix}\begin{pmatrix} U_m^* & 0 \\ 0 & I \end{pmatrix},$$

para obter $A_{11}^{(m+1)}$. Se $A_{11}^{(k)} = 0$ então A é nilpotente e $A_d = 0$. Se $A_{11}^{(k)}$ é não singular, então vá para o passo 4.

4. Nos passos anteriores construímos a matriz

$$WAW^* = \left(\begin{array}{c|cccc} A_{11}^{(k)} & 0 & \cdots & \cdots & 0 \\ \hline A_{21}^{(k)} & 0 & \cdots & \cdots & 0 \\ A_{31}^{(k)} & A_{32}^{(k)} & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \vdots \\ A_{k+1,1}^{(k)} & A_{k+1,2}^{(k)} & \cdots & A_{k+1,k}^{(k)} & 0 \end{array}\right) = \left(\begin{array}{c|c} B_1 & 0 \\ \hline B_2 & N \end{array}\right)$$

Onde B_1 é não singular, N é estritamente triangular inferior e W é uma matriz unitária dada por

$$W = \begin{pmatrix} U_{k-1} & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} U_{k-2} & 0 \\ 0 & I \end{pmatrix} \cdots \begin{pmatrix} U_1 & 0 \\ 0 & I \end{pmatrix} U.$$

Segue do Lema 2.38 que $A_d = W^*(WAW^*)_d W$, e como WAW^* satisfaz as condições do Teorema 2.40, temos que

$$A_d = W^*(WAW^*)_d W = W^* \begin{pmatrix} B_1^{-1} & 0 \\ XB_1^{-1} & 0 \end{pmatrix} W$$

onde X é a solução de $XB_1 - NX = B_2$, que é dada por

$$x_1 = r_1 B_1^{-1}, \quad x_i = \left(r_i + \sum_{j=1}^{i-1} n_{ij} x_j \right) B_1^{-1}, \quad i \geq 2,$$

onde r_i é a i-ésima linha de B_2 e x_i é a i-ésima linha de X.

Esse método possui um funcionamento similar ao método baseado em SVD, porém com o uso de decomposições conjugadas ao invés de decomposições em valores singulares, o que torna o método atraente é o fato de que a execução da decomposição conjugada pode ser feita de diferentes maneiras, e em geral envolve processos de ortonormalização, enquanto a SVD necessita do cálculo de autovalores.

3.3.2 Exemplos numéricos

Nos exemplos a seguir usamos o processo de ortonormalização de Gram-Schmidt modificado para obter a decomposição conjugada reduzida, com γ_i considerado nulo se $\gamma_i < tol$. Todos os testes foram realizados usando o software MATLAB, e consideramos como padrão $tol = 10^{-8}$.

Exemplo 3.5. Encontre a inversa de Drazin da matriz

$$A = \begin{pmatrix} 0 & 0 & 0 & 2 & 0 \\ 4 & 1 & 0 & 2 & 0 \\ 0 & -2 & 0 & 1 & 0 \\ 0 & 0 & 0 & 2 & 0 \\ 2 & 1 & 4 & -3 & 1 \end{pmatrix}.$$

Primeiramente decompomos $A = S^{-1}\Gamma U$, obtendo

$$S = \begin{pmatrix} 0,2639 & 0,6610 & 0,3486 & 0,2294 & 0,4862 \\ 0,1494 & -0,0923 & 0,3750 & -0,0342 & 0,0979 \\ 0,5333 & -0,0788 & -0,2194 & 0,5693 & -0,0983 \\ -0,2857 & 0,1271 & 0,1638 & -0,0801 & -0,1111 \\ 0,1530 & 0 & 0 & -0,1530 & 0 \end{pmatrix},$$

$$\Gamma = \operatorname{diag}(4,3288 \quad 0,8737 \quad 2,2366 \quad 0,6240 \quad 0)$$

e

$$U = \begin{pmatrix} 0,8355 & 0,1040 & 0,4493 & 0,2769 & 0,1123 \\ -0,1983 & -0,8521 & 0,4483 & 0,1455 & 0,1121 \\ -0,2288 & 0,1170 & -0,1758 & 0,9493 & -0,0439 \\ 0,4586 & -0,4995 & -0,7124 & 0,0319 & -0,1781 \\ 0 & 0 & -0,2425 & 0 & 0,9701 \end{pmatrix}.$$

Como Γ é singular, então calculamos

$$UAU^* = \left(\begin{array}{cccc|c} 1,3835 & 1,0385 & 2,1057 & 0,3758 & 0 \\ -3,0865 & 1,9925 & -1,1585 & -1,0363 & 0 \\ 0,7296 & -0,2837 & 1,5440 & 0,1474 & 0 \\ -2,2835 & -0,5376 & 0,5730 & -0,9199 & 0 \\ \hline 2,7517 & -0,2350 & -3,9914 & -2,8749 & 0 \end{array}\right) = \left(\begin{array}{c|c} A_{11}^{(1)} & 0 \\ \hline A_{21}^{(1)} & 0 \end{array}\right),$$

com $A_{11}^{(1)} \in \mathbb{C}^{4\times 4}$. Agora decompomos $A_{11}^{(1)} = S_1^{-1}\Gamma_1 U_1$, obtendo

$$S_1 = \begin{pmatrix} 0,5345 & 0,5232 & 0,8513 & 0,8699 \\ 0,0365 & 0,1391 & -0,2511 & -0,0973 \\ 0,1142 & -0,1415 & -0,3699 & 0,1889 \\ 0,7618 & -0,4688 & -1,6067 & 0,5819 \end{pmatrix},$$

$$\Gamma_1 = \operatorname{diag}(3,5044 \quad 0,7697 \quad 0,2046 \quad 0)$$

e

$$U_1 = \begin{pmatrix} -0,6394 & 0,2535 & 0,6655 & -0,2899 \\ -0,4416 & 0,5699 & -0,6855 & -0,1013 \\ -0,5213 & -0,7814 & -0,2858 & -0,1896 \\ -0,3528 & -0,0182 & 0,0743 & 0,9326 \end{pmatrix}.$$

Novamente temos Γ_1 singular, então construímos

$$U_1 A_{11}^{(1)} U_1^* = \left(\begin{array}{ccc|c} -0,1947 & 0,6678 & 0,6694 & 0 \\ 0,3423 & 3,6258 & 1,5706 & 0 \\ -2,4204 & -1,5041 & 0,5689 & 0 \\ \hline 1,5988 & 0,7272 & 2,2326 & 0 \end{array}\right) = \left(\begin{array}{c|c} A_{11}^{(2)} & 0 \\ \hline A_{21}^{(2)} & 0 \end{array}\right),$$

com $A_{11}^{(2)} \in \mathbb{C}^{3\times 3}$. Agora decompomos $A_{11}^{(2)} = S_2^{-1}\Gamma_2 U_2$ obtendo

$$S_2 = \begin{pmatrix} 0,5191 & 0,0020 & 0,3232 \\ 0,0637 & 0,0724 & -0,0123 \\ -0,6247 & 0,1709 & 0,1007 \end{pmatrix}, U_2 = \begin{pmatrix} -0,8484 & -0,1270 & 0,5138 \\ 0,1172 & 0,9016 & 0,4163 \\ -0,5162 & 0,4134 & -0,7501 \end{pmatrix}$$

e

$$\Gamma_2 = \mathrm{diag}(1,0404 \quad 0,3589 \quad 0,1233).$$

Enfim obtemos Γ_2 não singular, portanto $\mathrm{Ind}(A) = 2$, e definindo

$$W = \begin{pmatrix} U_1 & 0 \\ 0 & 1 \end{pmatrix} U,$$

temos que

$$\begin{aligned} WAW^* &= \left(\begin{array}{ccc|cc} -0,1947 & 0,6678 & 0,6694 & 0 & 0 \\ 0,3423 & 3,6258 & 1,5706 & 0 & 0 \\ -2,4204 & -1,5041 & 0,5689 & 0 & 0 \\ \hline 1,5988 & 0,7272 & 2,2326 & 0 & 0 \\ -3,6417 & 1,6782 & 0,4352 & -3,9441 & 0 \end{array}\right) = \left(\begin{array}{c|cc} A_{11}^{(2)} & 0 & 0 \\ \hline A_{21}^{(2)} & 0 & 0 \\ A_{31}^{(2)} & A_{32}^{(2)} & 0 \end{array}\right) \\ &= \left(\begin{array}{c|c} B_1 & 0 \\ \hline B_2 & N \end{array}\right), \end{aligned}$$

sendo

$$B_1 = A_{11}^{(1)}, \quad B_2 = \begin{pmatrix} A_{21}^{(2)} \\ A_{21}^{(2)} \end{pmatrix} \quad \text{e} \quad N = \begin{pmatrix} 0 & 0 \\ A_{32}^{(2)} & 0 \end{pmatrix}.$$

Agora calculamos

$$B_1^{-1} = (A_{11}^{(2)})^{-1} = U_2^*\Gamma_2^{-1} S_2 = \begin{pmatrix} 2,2125 & -0,6934 & -0,6892 \\ -1,9981 & 0,7548 & 0,2675 \\ 4,1306 & -0,9546 & -0,4673 \end{pmatrix}$$

e X solução da equação $XB_1 - NX = B_2$, dada por

$$X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} r_1 B_1^{-1} \\ (r_2 + n_{21}x_1)B_1^{-1} \end{pmatrix} = \begin{pmatrix} 11,3065 & -2,6910 & -1,9507 \\ -97,7059 & 34,9638 & 32,7333 \end{pmatrix},$$

onde r_i é a i-ésima linha de B_2. E por fim, temos que

$$A_d = W^* \begin{pmatrix} B_1^{-1} & 0 \\ XB_1^{-1} & 0 \end{pmatrix} W = \begin{pmatrix} -0,0000 & -0,0000 & -0,0000 & 0,5000 & -0,0000 \\ 4,0000 & 1,0000 & 0,0000 & -7,0000 & 0,0000 \\ -8,0000 & -2,0000 & -0,0000 & 17,2500 & -0,0000 \\ -0,0000 & -0,0000 & -0,0000 & 0,5000 & -0,0000 \\ 90,0000 & 15,0000 & 4,0000 & -149,5000 & 1,0000 \end{pmatrix},$$

com resíduo $\max\{\|A^kXA - A^k\|_2, \|XAX - X\|_2, \|AX - XA\|_2\} = 2,30 \times 10^{-12}$.

Capítulo 4

Um novo método iterativo para aproximar a inversa de Moore-Penrose baseado nas equações de Penrose

No capítulo anterior vimos que alguns métodos iterativos para aproximar a inversa de Moore-Penrose são baseados em algumas das equações de Penrose (1.1)-(1.4). Nesta seção apresentamos um método iterativo baseado nas equações (1.1) e (1.2).

Sejam $A \in \mathbb{C}^{m\times n}$ e $X = A^{\dagger}$. Das equações (1.1) e (1.2) temos que

$$X = XAX = XAXAX,$$

assim, dado $\beta \in \mathbb{R}$ temos que

$$\begin{aligned} X &= X - \beta(XAXAX - X) \\ &= X - \beta XAXAX + \beta X \\ &= X[(1+\beta)I - \beta AXAX] \end{aligned}$$

o que sugere a seguinte iteração:

$$X_0 = \alpha A^*, \qquad X_{k+1} = X_k[(1+\beta)I - \beta Y_k^2], \qquad Y_k = AX_k. \tag{4.1}$$

4.1 O método

Vejamos que a iteração sugerida em (4.1) é um método iterativo consistente.

Lema 4.1. *Se $A \in \mathbb{C}^{m\times n}$ e $X = A^\dagger$, então a sequência $\{X_k\}$ gerada pela iteração (4.1) satisfaz as seguintes propriedades para todo $k = 0, 1, 2, \ldots$.*

1. $X_k A X = X_k$,
2. $X A X_k = X_k$,
3. $(AX_k)^* = AX_k$,
4. $(X_k A)^* = X_k A$.

Demonstração:

Vamos proceder por indução.

1. Para $k = 0$ temos

$$X_0 AX = \alpha A^* AX = \alpha A^*(AX)^* = \alpha A^* X^* A^* = \alpha (AXA)^* = \alpha A^* = X_0.$$

Agora, supondo que a afirmação é válida para um inteiro k, para $k+1$ obtemos

$$\begin{aligned} X_{k+1}AX &= [(1+\beta)X_k - \beta X_k AX_k AX_k]AX \\ &= (1+\beta)X_k AX - \beta X_k AX_k AX_k AX \\ &= (1+\beta)X_k - \beta X_k AX_k AX_k = X_{k+1}. \end{aligned}$$

2. Para $k = 0$ temos

$$XAX_0 = XA\alpha A^* = \alpha (XA)^* A^* = \alpha A^* X^* A^* = \alpha (AXA)^* = \alpha A^* = X_0.$$

Supondo a afirmação válida para um k inteiro, para $k+1$ obtemos

$$\begin{aligned} XAX_{k+1} &= XAX_k[(1+\beta)I - \beta AX_k AX_k] \\ &= X_k[(1+\beta)I - \beta AX_k AX_k] = X_{k+1}. \end{aligned}$$

3. Para $k = 0$,

$$(AX_0)^* = (\alpha AA^*)^* = \alpha AA^* = AX_0.$$

Assumindo válido para k, segue para $k+1$ que

$$\begin{aligned}
(AX_{k+1})^* &= [AX_k[(1+\beta)I - \beta AX_kAX_k]]^* \\
&= [(1+\beta)AX_k - \beta AX_kAX_kAX_k]^* \\
&= (1+\beta)(AX_k)^* - \beta(AX_k)^*(AX_k)^*(AX_k)^* \\
&= (1+\beta)AX_k - \beta AX_kAX_kAX_k \\
&= AX_k[(1+\beta)I - \beta AX_kAX_k] = AX_{k+1}.
\end{aligned}$$

4. Para $k=0$ temos
$$(X_0A)^* = (\alpha A^*A)^* = \alpha A^*A = X_0A.$$
Supondo válido para k, para $k+1$ segue que
$$\begin{aligned}
(X_{k+1}A)^* &= [X_k[(1+\beta)I - \beta AX_kAX_k]A]^* \\
&= [(1+\beta)X_kA - \beta X_kAX_kAX_kA]^* \\
&= (1+\beta)(X_kA)^* - \beta(X_kA)^*(X_kA)^*(X_kA)^* \\
&= (1+\beta)X_kA - \beta X_kAX_kAX_kA \\
&= X_k[(1+\beta)I - \beta AX_kAX_k]A = X_{k+1}A.
\end{aligned}$$

■

Para X_k gerado pela iteração (4.1), definimos a matriz erro E_k por

$$E_k = X_k - X, \tag{4.2}$$

onde $X = A^\dagger$, dessa forma podemos mostrar a seguinte relação de recorrência em $E_{k+1}A$.

Lema 4.2. *Se $A \in \mathbb{C}^{m\times n}$, $X = A^\dagger$ e $E_k = X_k - X$, onde X_k é gerado por (4.1), então*

$$E_{k+1}A = -\beta(E_kA)^3 - 3\beta(E_kA)^2 + (1-2\beta)E_kA.$$

Demonstração:

Como $X_{k+1} = X_k[(1+\beta)I - \beta AX_kAX_k]$, então

$$\begin{aligned}
E_{k+1}A &= (X_{k+1} - X)A \\
&= X_{k+1}A - XA \\
&= X_k[(1+\beta)I - \beta AX_kAX_k]A - XA \\
&= X_kA + \beta X_kA - \beta(X_kA)^3 - XA. \qquad (4.3)
\end{aligned}$$

Assim, segue das equações de Penrose e do Lema 4.1 que

$$\begin{aligned}
-\beta(E_kA)^3 &= -\beta(X_kA - XA)^3 \\
&= -\beta[(X_kA)^2 - X_kAXA - XAX_kA + XAXA](X_kA - XA) \\
&= -\beta[(X_kA)^2 - 2X_kA + XA](X_kA - XA) \\
&= -\beta[(X_kA)^3 - X_kAX_kAXA - 2(X_kA)^2 + 2X_kAXA + \\
&\quad XAX_kA - (XA)^2] \\
&= -\beta[(X_kA)^3 - (X_kA)^2 - 2(X_kA)^2 + 2X_kA + X_kA - XA] \\
&= -\beta(X_kA)^3 + 3\beta(X_kA)^2 - 3\beta(X_kA) + \beta XA + \\
&\quad (XA - XA + X_kA - X_kA) \\
&= [-\beta(X_kA)^3 + X_kA + \beta X_kA - XA] + [3\beta(X_kA)^2 - \\
&\quad (1+4\beta)X_kA + (1+\beta)XA]. \qquad (4.4)
\end{aligned}$$

Substituindo (4.4) na equação (4.3) resulta em

$$E_{k+1}A = -\beta(E_kA)^3 - 3\beta(X_kA)^2 + (1+4\beta)X_kA - (1+\beta)XA. \qquad (4.5)$$

Agora, analisando o termo $-3\beta(E_kA)^2$ obtemos

$$\begin{aligned}
-3\beta(E_kA)^2 &= -3\beta(X_kA - XA)^2 \\
&= -3\beta[(X_kA)^2 - X_kAXA - XAX_kA + XAXA] \\
&= -3\beta[(X_kA)^2 - 2X_kA + XA] \\
&= -3\beta(X_kA)^2 + 6\beta X_kA - 3\beta XA \\
&= -3\beta(X_kA)^2 + (1+4\beta)X_kA - (1+\beta)XA + \\
&\quad (2\beta - 1)X_kA + (1-2\beta)XA,
\end{aligned}$$

assim,

$$\begin{aligned}
&-3\beta(X_kA)^2 + (1+4\beta)X_kA - (1+\beta)XA = \\
&-3\beta(E_kA)^2 + (1-2\beta)X_kA - (1-2\beta)XA, \qquad (4.6)
\end{aligned}$$

e a combinação de (4.6) e (4.5) nos fornece

$$E_{k+1}A = -\beta(E_kA)^3 - 3\beta(E_kA)^2 + (1-2\beta)E_kA.$$

∎

Com esses resultados, podemos mostrar a convergência do método.

Teorema 4.3. *Seja $A \in \mathbb{C}^{m\times n}$. Se $0 < \alpha < \frac{2}{\rho(A^*A)}$ e $0 < \beta \leq \frac{1}{3}$, então o método iterativo (4.1) converge para a inversa de Moore-Penrose de A.*

Demonstração:

Defina $t_k = \|E_kA\|_2$, vamos mostrar via indução que $t_k < 1\ \forall\ k \in \mathbb{N}$. Para $k = 0$, como $0 < \alpha < \frac{2}{\rho(A^*A)}$, segue do Teorema 2.32 que $t_0 < 1$. Agora, assumindo que $t_k < 1$ para um inteiro k, do Lema 4.2 temos que

$$\begin{aligned} t_{k+1} &= \| -\beta(E_kA)^3 - 3\beta(E_kA)^2 + (1-2\beta)E_kA\|_2 \\ &\leq \|I - \beta[(E_kA)^2 + 3E_kA + 2I]\|_2 t_k. \end{aligned} \tag{4.7}$$

Agora, do Lema 4.1 segue que E_kA é uma matriz Hermitiana e portanto $I - \beta[(E_kA)^2 + 3E_kA + 2I]$ é Hermitiana, assim

$$\begin{aligned} \|I - \beta[(E_kA)^2 + 3E_kA + 2I]\|_2 &= \rho\left(I - \beta[(E_kA)^2 + 3E_kA + 2I]\right) \\ &= \max_i |1 - \beta(\lambda_i^2 + 3\lambda_i + 2)|, \end{aligned}$$

onde λ_i são os autovalores da matriz E_kA. Como E_kA é Hermitiana, usando a hipótese de indução temos que

$$\rho(E_kA) = \|E_kA\|_2 < 1,$$

o que implica em $-1 < \lambda_i < 1$, e portanto

$$0 < \lambda_i^2 + 3\lambda_i + 2 < 6,$$

e de $\beta \leq \frac{1}{3}$ segue que

$$0 < \beta(\lambda_i^2 + 3\lambda_i + 2) < 2,$$

implicando

$$-1 < 1 - \beta(\lambda_i^2 + 3\lambda_i + 2) < 1,$$

e portanto

$$|1 - \beta(\lambda_i^2 + 3\lambda_i + 2)| < 1,$$

o que mostra que

$$\|I - \beta[(E_kA)^2 + 3E_kA + 2I]\|_2 < 1.$$

Assim, da equação (4.7) temos que $t_{k+1} < t_k < 1$, concluindo que a sequência $\{t_k\}$ é uma sequência decrescente com limite inferior 0, e como $t_0 < 1$ então $t_k \to t$ para algum $t \in [0, 1)$.

Vejamos que $t = 0$. De fato, usando novamente o Lema 4.2 temos que

$$\begin{aligned} \|E_{k+1}A\|_2 &= \| -\beta(E_kA)^3 - 3\beta(E_kA)^2 + (1-2\beta)E_kA\|_2 \\ &\leq \|I - \beta[(E_kA)^2 + 3E_kA + 2I\|_2\|E_kA\|_2. \end{aligned} \tag{4.8}$$

Porém, como $I - \beta[(E_kA)^2 + 3E_kA + 2I]$ é Hermitiana temos

$$\|I - \beta[(E_kA)^2 + 3E_kA + 2I]\|_2 = \max_i |1 - \beta(\lambda_i^2 + 3\lambda_i + 2)|,$$

para λ_i autovalor de E_kA, com $i = i, 2, \ldots, n$. De $t_0 < 1$, existe $\epsilon > 0$ pequeno, não dependendo de k, tal que $t_0 < 1 - \epsilon$, e segue do fato de $t_k < t_{k-1} < \ldots < t_0$ que

$$-1 + \epsilon < \lambda_i < 1 - \epsilon, \quad i = 1, 2, \ldots, n,$$

assim

$$\begin{aligned} (-1+\epsilon)^2 + 3(-1+\epsilon) + 2 &< \lambda_i^2 + 3\lambda_i + 2 < (1-\epsilon)^2 + 3(1-\epsilon) + 2 \\ \epsilon^2 + \epsilon &< \lambda_i^2 + 3\lambda_i + 2 < 6 - (5\epsilon - \epsilon^2), \end{aligned}$$

multiplicando por $\beta > 0$, obtemos

$$\beta(\epsilon^2 + \epsilon) < \beta(\lambda_i^2 + 3\lambda_i + 2) < \beta[6 - (5\epsilon - \epsilon^2)],$$

e como $0 < \beta \leq \frac{1}{3}$, segue que

$$\beta(\epsilon^2 + \epsilon) < \beta(\lambda_i^2 + 3\lambda_i + 2) < 2 - \frac{1}{3}(5\epsilon - \epsilon^2)$$

e

$$-1 + \frac{1}{3}(5\epsilon - \epsilon^2) < 1 - \beta(\lambda_i^2 + 3\lambda_i + 2) < 1 - \beta(\epsilon^2 + \epsilon).$$

Agora, definindo

$$\gamma = \gamma(\epsilon) = \min\left\{\frac{1}{3}(5\epsilon - \epsilon^2), \beta(\epsilon^2 + \epsilon)\right\} > 0,$$

temos que

$$|1 - \beta(\lambda_i^2 + 3\lambda_i + 2)| < 1 - \gamma, \quad \text{para} \;\; i = 1, 2, \ldots, n.$$

assim,

$$\|I - \beta[(E_kA)^2 + 3E_kA + 2I]\|_2 < 1 - \gamma, \quad \forall\, k \in \mathbb{N}.$$

e tomando o limite quando k tende a infinito em (4.8), como γ não depende de k, obtemos

$$t \leq \lim_{k\to\infty} \|I - \beta[(E_kA)^2 + 3E_kA + 2I]\|_2\|E_kA\|_2 \leq (1-\gamma)t,$$

e como $\gamma > 0$ segue que $t = 0$.

Dessa forma, usando o Lema 4.1 e as equações de Penrose obtemos

$$\|E_k\|_2 = \|X_k - X\|_2 = \|X_k AX - XAX\|_2 \leq \|E_k A\|_2 \|X\|_2$$

e quando $k \to \infty$ temos

$$\lim_{k\to\infty} \|E_k\|_2 \leq \lim_{k\to\infty} \|E_k A\|_2 \|X\|_2 = 0,$$

de onde segue que X_k converge para $X = A^\dagger$. ■

O próximo teorema mostra que a convergência do método é linear, além de explicitar cada termo do erro.

Teorema 4.4. *Seja $A \in \mathbb{C}^{m\times n}$. Temos que a matriz E_k definida por (4.2) satisfaz*

$$E_{k+1} = (1-2\beta)E_k - 3\beta E_k AE_k - \beta E_k AE_k AE_k. \tag{4.9}$$

Demonstração:

Como $E_k = X_k - X$ então $X_k = E_k + X$, e substituindo essa expressão em (4.1) temos

$$\begin{aligned} X_{k+1} &= (E_k + X)\left[(1+\beta)I - \beta A(E_k + X)A(E_k + X)\right] \\ &= (1+\beta)(E_k + X) - \beta(E_k + X)A(E_k + X)A(E_k + X) \\ &= (1+\beta)(E_k + X) - \beta\,(E_k AE_k AE_k + E_k AE_k AX + E_k AXAE_k + \\ &\quad E_k AXAX + XAE_k AE_k + XAE_k AX + XAXAE_k + XAXAX\ ). \end{aligned}$$

Segue do Lema 4.1 e das equações de Penrose que $E_k AX = E_k$ e $XAE_k = E_k$, assim

$$\begin{aligned} X_{k+1} &= (1+\beta)(E_k + X) - \beta(E_k AE_k AE_k + 3E_k AE_k + 3E_k + X) \\ &= E_k + X + \beta E_k + \beta X - \beta E_k AE_k AE_k - 3\beta E_k AE_k - 3\beta E_k - \beta X \\ &= X + (1-2\beta)E_k - 3\beta E_k AE_k - \beta E_k AE_k AE_k, \end{aligned}$$

de onde temos que

$$E_{k+1} = X_{k+1} - X = (1-2\beta)E_k - 3\beta E_k AE_k - \beta E_k AE_k AE_k.$$

■

Como podemos ver pelo teorema acima, a taxa de convergência é linear, além disso, a constante que multiplica o termo linear é $(1-2\beta)$, o que nos mostram que para valores de β mais próximos de $\frac{1}{3}$, teremos uma constante menor, melhorando a convergência.

Daqui em diante, por $\|\cdot\|$ denotamos uma norma matricial submultiplicativa. O Algoritmo 5 resume o método iterativo 4.1.

Algoritmo 5. Algoritmo para aproximar a inversa de Moore-Penrose.

Dados A, α, β, tol,
$X_0 = \alpha A^*$,
Para $k = 0, 1$, faça
  $Y_k = AX_k$,
  $X_{k+1} = X_k[(1+\beta)I - \beta Y_k^2]$,
  $r_k = \|X_{k+1} - X_k\|$,
Fim do para
$k = 2$,
Enquanto o critério de parada não for satisfeito, faça
  $Y_k = AX_k$,
  $X_{k+1} = X_k[(1+\beta)I - \beta Y_k^2]$,
  $r_k = \|X_{k+1} - X_k\|$,
  $k = k + 1$,
Fim do enquanto

4.2 Critérios de parada

O algoritmo 5 pede que a iteração seja executada até que o critério de parada seja satisfeito, por isso, para que o método esteja completo precisamos estabelecer quando parar de iterar, ou seja, algum critério fácil de ser avaliado que permita concluir que X_k está próximo o suficiente da inversa de Moore-Penrose.

O critério mais comumente usado é verificar o decréscimo no tamanho do erro $E_k = X_k - X$, o qual pode ser medido sem o conhecimento prévio da solução $X = A^\dagger$, pois

$$\|E_{k+1} - E_k\| = \|(X_{k+1} - X) - (X_k - X)\| = \|X_{k+1} - X_k\|.$$

Assim, um critério a ser usado é

$$\|X_{k+1} - X_k\| < tol, \tag{4.10}$$

onde tol é uma tolerância estabelecida.

O próximo teorema nos fornece um critério de parada alternativo.

Teorema 4.5. *Considere o método iterativo definido por (4.1). Se $t_k = \|E_k A\|$, $e_k = \|E_k\|$ e $r_k = \|E_{k+1} - E_k\|$, para alguma norma matricial submultiplicativa $\|\cdot\|$, então*

$$\lim_{k\to\infty} \frac{t_{k+1}}{t_k} = \lim_{k\to\infty} \frac{e_{k+1}}{e_k} = \lim_{k\to\infty} \frac{r_{k+1}}{r_k} = 1 - 2\beta.$$

Demonstração:

Segue do Lema 4.2 e da desigualdade triangular que

$$t_{k+1} \leq \beta t_k^3 + 3\beta t_k^2 + (1-2\beta)t_k,$$

dividindo por t_k e tomando o limite em k temos

$$\lim_{k\to\infty} \frac{t_{k+1}}{t_k} \leq \lim_{k\to\infty} \left(\beta t_k^2 + 3\beta t_k + (1-2\beta)\right),$$

e como $\lim_{k\to\infty} t_k = 0$, temos

$$\lim_{k\to\infty} \frac{t_{k+1}}{t_k} \leq 1-2\beta. \tag{4.11}$$

Por outro lado, segue do Lema 4.2 que

$$E_{k+1}A + \beta(E_kA)^3 + 3\beta(E_kA)^2 = (1-2\beta)E_kA,$$

o que é o mesmo que

$$t_{k+1} + \beta t_k^3 + 3\beta t_k^2 \geq (1-2\beta)t_k,$$

portanto

$$t_{k+1} \geq -\beta t_k^3 - 3\beta t_k^2 + (1-2\beta)t_k.$$

Dividindo por t_k e tomando o limite quando k tende a infinito temos

$$\lim_{k\to\infty} \frac{t_{k+1}}{t_k} \geq \lim_{k\to\infty} -\beta t_k^2 - 3\beta t_k + (1-2\beta),$$

e como t_k converge para zero, obtemos

$$\lim_{k\to\infty} \frac{t_{k+1}}{t_k} \geq (1-2\beta). \tag{4.12}$$

De (4.11) e (4.12) segue que

$$\lim_{k\to\infty} \frac{t_{k+1}}{t_k} = 1-2\beta.$$

Usando a mesma ideia na relação de recorrência do Teorema 4.4 obtemos

$$e_{k+1} \leq (1-2\beta)e_k + 3\beta e_k^2\|A\| + \beta e_k^3\|A\|^2,$$

dividindo por e_k, e usando que e_k converge para zero, temos que

$$\lim_{k\to\infty} \frac{e_{k+1}}{e_k} \leq 1-2\beta. \tag{4.13}$$

Usando novamente o Teorema 4.4, temos que

$$E_{k+1} + 3\beta E_k A E_k + \beta E_k A E_k A E_k = (1-2\beta)E_k,$$

tomando a norma $\|\cdot\|$ em ambos os lados, resulta em

$$e_{k+1} + 3\beta e_k^2\|A\| + \beta e_k^3\|A\|^2 \geq (1-2\beta)e_k,$$

dividindo por e_k e tomando o limite, como e_k converge para zero, temos que

$$\lim_{k\to\infty} \frac{e_{k+1}}{e_k} \geq 1-2\beta. \tag{4.14}$$

De (4.13) e (4.14) temos

$$\lim_{k\to\infty} \frac{e_{k+1}}{e_k} = 1-2\beta.$$

Agora, para r_k basta subtrair E_k na relação de recorrência do Teorema 4.4 e tomar a norma $\|\cdot\|$ em ambos os lados que obtemos

$$r_k \leq 2\beta e_k + 3\beta e_k^2\|A\| + \beta e_k^3\|A\|^2,$$

dividindo por e_k em ambos os lados, e tomando o limite para k tendendo a infinito temos que

$$\lim_{k\to\infty} \frac{r_k}{e_k} \leq 2\beta. \tag{4.15}$$

Por outro lado, da relação de recorrência do Teorema 4.4 podemos escrever

$$E_{k+1} - E_k + 3\beta E_k A E_k + \beta E_k A E_k A E_k = -2\beta E_k$$

e tomando a norma $\|\cdot\|$ em ambos os lados temos

$$r_k + 3\beta e_k^2\|A\| + \beta e_k^3\|A\|^2 \geq 2\beta e_k,$$

assim

$$r_k \geq 2\beta e_k - 3\beta e_k^2\|A\| - \beta e_k^3\|A\|^2,$$

dividindo por e_k e tomando o limite em k temos

$$\lim_{k\to\infty} \frac{r_k}{e_k} \geq 2\beta. \tag{4.16}$$

De (4.15) e (4.16) segue que

$$\lim_{k\to\infty} \frac{r_k}{e_k} = 2\beta.$$

Repetindo o mesmo procedimento para $k+1$ no lugar de k obtemos

$$\lim_{k\to\infty} \frac{r_{k+1}}{e_{k+1}} = 2\beta,$$

assim

$$\lim_{k\to\infty} \frac{r_{k+1}}{r_k} = \lim_{k\to\infty} \left(\frac{\frac{r_{k+1}}{e_{k+1}}}{\frac{r_k}{e_k}} \frac{e_{k+1}}{e_k} \right) = \lim_{k\to\infty} \left(\frac{\frac{r_{k+1}}{e_{k+1}}}{\frac{r_k}{e_k}} \right) \lim_{k\to\infty} \left(\frac{e_{k+1}}{e_k} \right) = 1 - 2\beta.$$

■

O teorema anterior mostra que podemos usar como critério de parada a expressão

$$\left| \frac{r_{k+1}}{r_k} - (1-2\beta) \right| < tol, \tag{4.17}$$

onde tol é a tolerância previamente estabelecida.

4.3 Influência dos erros de arredondamento

Erros de arredondamento sempre ocorrem quando lidamos com aritmética de ponto flutuante, e nesta seção vamos considerar a influência desses erros no método iterativo (4.1). A consequência é que a matriz computada, a qual chamaremos de $\tilde{X}_k$, difere da matriz X_k por uma matriz erro Δ_k. Vejamos agora algumas consequências desse erro.

Considere o processo iterativo

$$\tilde{Y}_k = A\tilde{X}_k, \qquad \tilde{X}_{k+1} = \tilde{X}_k[(1+\beta)I - \beta\tilde{Y}_k^2] \tag{4.18}$$

onde $\tilde{X}_k = X_k + \Delta_k$.

Lema 4.6. *A sequência X_k gerada pelo método 4.1 satisfaz $\mathcal{R}(X_k) = \mathcal{R}(A^*)$ e $\mathcal{N}(X_k) = \mathcal{N}(A^*)$.*

Demonstração:

Para $k = 0$ temos que $X_0 = \alpha A^*$, como $\alpha \neq 0$ claramente temos $\mathcal{R}(X_0) = \mathcal{R}(A^*)$ e $\mathcal{N}(X_0) = \mathcal{N}(A^*)$.

Agora, para $k > 0$ inteiro arbitrário, tome $y \in \mathcal{N}(X_0)$, então

$$\begin{aligned}
X_1 y &= (1+\beta)X_0 y - \beta X_0 A X_0 A X_0 y = 0 \\
X_2 y &= (1+\beta)X_1 y - \beta X_1 A X_1 A X_1 y = 0 \\
&\vdots \\
X_k y &= (1+\beta)X_{k-1} y - \beta X_{k-1} A X_{k-1} A X_{k-1} y = 0,
\end{aligned}$$

assim, $\mathcal{N}(A^*) = \mathcal{N}(X_0) \subset \mathcal{N}(X_1) \subset \cdots \subset \mathcal{N}(X_k)$.

Agora tome $y \in \mathcal{R}(X_k)$, então existe $z \in \mathbb{C}^m$ tal que $X_k z = y$, assim

$$y = X_k z = (1+\beta)X_{k-1}z - \beta X_{k-1}AX_{k-1}AX_{k-1}z = X_{k-1}[(1+\beta)z - \beta AX_{k-1}AX_{k-1}z],$$

portanto $y \in \mathcal{R}(X_{k-1})$. Repetindo o processo obtemos

$$\mathcal{R}(X_k) \subset \mathcal{R}(X_{k-1}) \subset \cdots \subset \mathcal{R}(X_0) = \mathcal{R}(A^*).$$

Defina $\mathcal{U} = \bigcup_{k\in\mathbb{N}} \mathcal{N}(X_k)$. Se $y \in \mathcal{U}$, então $y \in \mathcal{N}(X_l)$ para algum $l \in \mathbb{N}$, o que implica que $y \in \mathcal{N}(X_k)$ para todo $k \geq l$, ou seja, $X_k y = 0$ $\forall k \geq l$. Pelo Teorema 4.3 temos

$$Xy = \lim_{k\to\infty} X_k y = 0,$$

portanto $y \in \mathcal{N}(X) = \mathcal{N}(A^*)$ e $\mathcal{U} \subset \mathcal{N}(A^*)$. Assim,

$$\mathcal{N}(A^*) \subset \mathcal{N}(X_k) \subset \mathcal{U} \subset \mathcal{N}(A^*).$$

portanto

$$\mathcal{N}(X_k) = \mathcal{N}(A^*) \quad \forall\ k \geq 0.$$

Como

$$\dim(\mathcal{R}(X_k)) = m - \dim(\mathcal{N}(X_k)) = m - \dim(\mathcal{N}(A^*)) = \dim(\mathcal{R}(A^*)),$$

e $\mathcal{R}(X_k) \subset \mathcal{R}(A^*)$ então

$$\mathcal{R}(X_k) = \mathcal{R}(A^*) \quad \forall\ k \geq 0.$$

■

Teorema 4.7. *Considere a sequência $\{X_k\}$ gerada pelo método iterativo (4.1), e a sequência $\{\tilde{X}_k\}$ gerada pelo processo com erros de arredondamento (4.18). Suponha que na s-ésima iteração ocorra*

$$posto(\tilde{X}_s) > posto(X_s) = posto(A).$$

Então o processo iterativo (4.18) diverge e

$$\|\tilde{X}_k\| \geq C(1+\beta)^{k-s},$$

para alguma norma matricial induzida $\|\cdot\|$, onde $C > 0$ é uma constante que depende somente de $\tilde{X}_s$.

Demonstração:

Como

$$\mathcal{N}(\tilde{X}_s) \subset \mathcal{N}(\tilde{X}_s A \tilde{X}_s A \tilde{X}_s) \quad \text{e} \quad \text{posto}(\tilde{X}_s A \tilde{X}_s A \tilde{X}_s) \leq \text{posto}(A) < \text{posto}(\tilde{X}_s),$$

então,

$$\mathcal{N}(\tilde{X}_s) \subsetneq \mathcal{N}(\tilde{X}_s A \tilde{X}_s A \tilde{X}_s).$$

Assim, existe $y \neq 0$, tal que $y \in \mathcal{N}(\tilde{X}_s A \tilde{X}_s A \tilde{X}_s) \setminus \mathcal{N}(\tilde{X}_s)$, ou seja, $\tilde{X}_s A \tilde{X}_s A \tilde{X}_s y = 0$ e $\tilde{X}_s y \neq 0$.

Vamos mostrar, via indução, que

$$\tilde{X}_k y = (1+\beta)^{k-s} \tilde{X}_s y \quad \text{e} \quad \tilde{X}_k A \tilde{X}_k A \tilde{X}_k y = 0 \quad \forall\, k \geq s. \tag{4.19}$$

Para s foi provado acima. Agora assuma que as equações de (4.19) sejam válidas para k. Para $k+1$ temos

$$\begin{aligned}
&\tilde{X}_{k+1} A \tilde{X}_{k+1} A \tilde{X}_{k+1} = \\
&= [(1+\beta)\tilde{X}_k - \beta \tilde{X}_k (A\tilde{X}_k)^2] A [(1+\beta)\tilde{X}_k - \beta \tilde{X}_k (A\tilde{X}_k)^2] A [(1+\beta)\tilde{X}_k - \beta \tilde{X}_k (A\tilde{X}_k)^2] \\
&= (1+\beta)^3 \tilde{X}_k (A\tilde{X}_k)^2 - 3\beta(1+\beta)^2 \tilde{X}_k (A\tilde{X}_k)^4 + 3\beta^2 (1+\beta) \tilde{X}_k (A\tilde{X}_k)^6 - \beta^3 \tilde{X}_k (A\tilde{X}_k)^8,
\end{aligned}$$

e da hipótese de indução seque que $\tilde{X}_{k+1} A \tilde{X}_{k+1} A \tilde{X}_{k+1} y = 0$. Além disso, temos que

$$\begin{aligned}
\tilde{X}_{k+1} y &= (1+\beta)\tilde{X}_k y - \beta \tilde{X}_k A \tilde{X}_k A \tilde{X}_k y \\
&= (1+\beta)(1+\beta)^{k-s} \tilde{X}_s y \\
&= (1+\beta)^{k+1-s} \tilde{X}_s y,
\end{aligned}$$

mostrando que (4.19) vale para todo $k \geq s$. Assim, para $\|\cdot\|$ norma matricial induzida por uma norma vetorial $\|\cdot\|$ temos

$$\|\tilde{X}_k\| = \sup_{z \neq 0} \frac{\|\tilde{X}_k z\|}{\|z\|} \geq \frac{\|\tilde{X}_k y\|}{\|y\|} = (1+\beta)^{k-s} \frac{\|\tilde{X}_s y\|}{\|y\|} = C(1+\beta)^{k-s}, \tag{4.20}$$

onde $C = \frac{\|\tilde{X}_s y\|}{\|y\|} > 0$ depende somente de s. Além disso

$$\|\tilde{X}_k - X\| \geq \|\tilde{X}_k\| - \|X\|,$$

e de (4.20) segue que

$$\lim_{k \to \infty} (\|\tilde{X}_k\| - \|X\|) = \infty,$$

mostrando que a sequência $\{\tilde{X}_k\}$ diverge. ■

O teorema acima mostra que no caso em que os erros de arredondamento alteram o posto da aproximação que está sendo calculada, então o método pode divergir.

4.4 Exemplos numéricos

Nesta seção apresentamos exemplos e comparações do método (4.1). Quando não informado o critério de parada e a tolerância, será considerado o critério (4.10) com $tol = 10^{-8}$. Definimos o resíduo na iteração k por

$$\max\{\|AX_kA - A\|_2, \|X_kAX_k - X_k\|_2, \|(AX_k)^* - AX_k\|_2, \|(X_kA)^* - X_kA\|_2\}.$$

Exemplo 4.1. Encontrar a inversa de Moore-Penrose da matriz A do exemplo 3.3 via método (4.1) com $\beta = \frac{1}{3}$.

Seguindo o critério de convergência do método, considere $\alpha = 0,018$ e $X_0 = \alpha A^*$, aplicando o processo iterativo (4.1), depois de 22 iterações obtemos

$$A^\dagger = \begin{pmatrix} 0,0300 & 0,0109 & 0,0069 & 0,0191 & 0,0369 & 0,0198 \\ -0,0251 & -0,0237 & 0,0455 & -0,0013 & 0,0204 & -0,0129 \\ 0,0049 & -0,0129 & 0,0524 & 0,0178 & 0,0574 & 0,0069 \\ 0,0551 & 0,0346 & -0,0386 & 0,0204 & 0,0165 & 0,0326 \end{pmatrix},$$

com resíduo $3,45 \times 10^{-7}$.

No capítulo 2 vimos métodos da forma

$$X_{k+1} = X_k + C_k(I - AX_k), \tag{4.21}$$

onde o mais relevante de tais métodos é o de ordem 2. Além disso vimos um método proposto por Petković e Stanimirović em [23] e [24], que no caso em que $\beta = 1$ recai no método de ordem 2 citado acima. Na demonstração do Teorema 2.34, assim como descrito em [24], não há necessidade de que o parâmetro da aproximação inicial seja o mesmo da iteração, o que torna o método (2.9) uma generalização do método de ordem 2 estudado.

O método proposto neste capítulo possui um funcionamento parecido com o método (2.9), fazendo uso de um parâmetro β, que neste caso possui maior restrição para garantia de convergência, pois precisamos tomar $\beta \in (0, 1/3]$. Além disso, pelo termo de primeira ordem de erro mostrado no Teorema 4.4 podemos perceber que a convergência é mais rápida para valores de β maiores, sendo assim faremos uma comparação entre o método proposto e o método (2.9).

Exemplo 4.2. Neste exemplo comparamos o método iterativo proposto (4.1), e o método iterativo (2.9). Em ambos usamos o parâmetro $\beta = \frac{1}{3}$. Para tanto consideramos matrizes de tamanhos variados geradas aleatoriamente pelo comando *rand* do MATLAB. Para a aproximação inicial consideramos $X_0 = \alpha A^*$ com $\alpha = \frac{2}{\rho(A^*A)+1}$. A tabela 4.1 mostra o tempo em segundos, o número de iterações e o erro cometido em cada método.

Tamanho	α	Tempo		Iterações		Erro	
		(4.1)	(2.9)	(4.1)	(2.9)	(4.1)	(2.9)
400×350	$1,42 \times 10^{-3}$	$1,84$	$2,88$	59	90	$3,84 \times 10^{-9}$	$1,51 \times 10^{-8}$
500×400	$1,00 \times 10^{-3}$	$2,39$	$3,54$	57	87	$2,59 \times 10^{-9}$	$1,47 \times 10^{-8}$
800×700	$9,13 \times 10^{-4}$	$9,51$	$12,93$	63	93	$1,71 \times 10^{-9}$	$1,93 \times 10^{-8}$
1500×1000	$5,33 \times 10^{-4}$	$26,27$	$37,57$	56	85	$2,77 \times 10^{-9}$	$1,76 \times 10^{-8}$
2500×1800	$1,77 \times 10^{-4}$	$170,97$	$245,96$	59	89	$4,91 \times 10^{-9}$	$1,41 \times 10^{-8}$
4000×2000	$9,99 \times 10^{-5}$	$253,95$	$363,11$	55	82	$1,67 \times 10^{-9}$	$1,83 \times 10^{-8}$
3000×2500	$1,06 \times 10^{-6}$	$463,26$	$620,34$	63	89	$3,82 \times 10^{-8}$	$3,13 \times 10^{-8}$

Tabela 4.1: Comparação entre os métodos (4.1) e (2.9) para $\beta = \frac{1}{3}$.

Como podemos notar, nosso método apresenta uma quantidade de iterações consideravelmente menor que o método (2.9) para obter uma aproximação da inversa de Moore-Penrose, e consequentemente, o tempo gasto para obter tal aproximação é menor.

Exemplo 4.3. Vejamos como os erros de arredondamento influenciam na solução. Considere as matrizes

$$A_1 = \begin{pmatrix} 1 & 2 & -1 & 3 & 4 \\ 0 & -3 & 5 & 1 & 2 \\ 1 & 3 & 4 & 2 & -2 \\ 4 & 3 & -1 & 3 & -2 \\ 0 & -2 & 4 & -1 & 0 \\ 4 & 3 & 5 & -3 & 1 \\ 1 & 3 & -3 & 2 & 5 \\ 7 & -1 & -4 & 5 & 2 \end{pmatrix} \quad \text{e} \quad A_2 = \begin{pmatrix} 1 & 2 & -1 & 1 & 2 \\ 0 & -3 & 5 & 0 & -3 \\ 1 & 3 & 4 & 1 & 3 \\ 4 & 3 & -1 & 4 & 3 \\ 0 & -2 & 4 & 0 & -2 \\ 4 & 3 & 5 & 4 & 3 \\ 1 & 3 & -3 & 1 & 3 \\ 7 & -1 & -4 & 7 & -1 \end{pmatrix}.$$

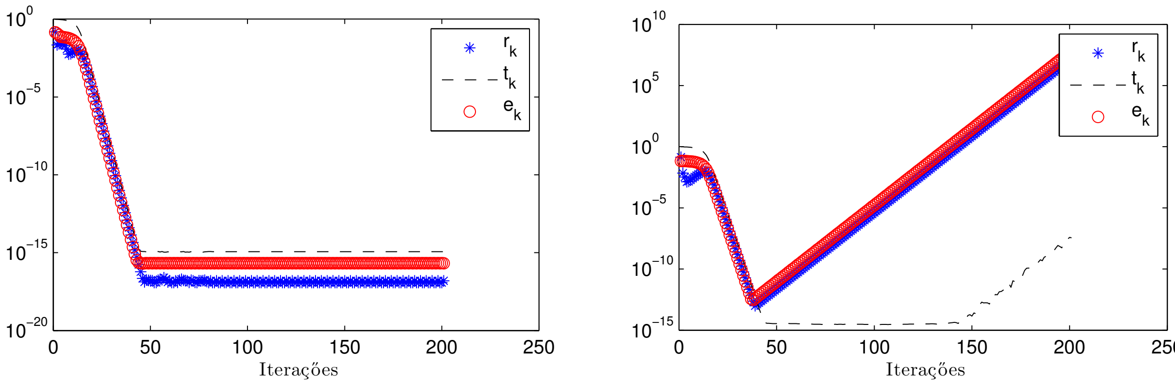
A_1 é uma matriz de posto completo, enquando A_2 foi obtida de A_1 trocando as duas últimas colunas pelas duas primeiras, portanto posto$(A_2) = 3$. Executamos 200 iterações do método (4.1) com $\beta = \frac{1}{3}$ para as matrizes A_1 e A_2 para analizar o comportamento dos erros de arredondamento. Em cada caso analisamos o comportamento dos valores

$$t_k = \|(X_k - X)A_i\|_2, \quad e_k = \|X_k - X\|_2 \quad \text{e} \quad r_k = \|X_{k+1} - X_k\|_2, \quad i = 1, 2.$$

A Figura 4.1 mostra o comportamento dos valores acima para as matrizes A_1 e A_2. Como podemos notar para a matriz A_1, que possui posto completo, mesmo depois de muitas iterações o erro se mantém estável, enquanto que no caso da matriz A_2, que possui posto incompleto, o erro atinge um valor mínimo e depois começa a crescer.

Note também que no caso da matriz A_2 o valor de t_k se mantém estável por mais iterações que os valores e_k e r_k, e ao verificar as equações de Penrose para alguns valores de k, podemos notar que o crescimento do erro começa na equação $XA_2X = X$. A tabela 4.2 mostra como se comportam as equações de Penrose para $k = 40$, $k = 125$ e $k = 200$.

O que ocorre neste exemplo é que em determinada iteração a matriz X_k passa a ter posto maior que a matriz X_{k-1} devido a erros de arredondamento, e conforme o Teorema 4.7 o método diverge quando k tende a infinito. Esse fato mostra que para matrizes de posto incompleto não devemos exigir uma tolerância excessivamente baixa, porém ainda assim podemos obter uma boa aproximação da inversa de Moore-Penrose, por exemplo para essa matriz A_2 a tabela 4.2 mostra que uma tolerância na casa de 10^{-13} foi atingida.



(a) Matriz A_1.

(b) Matriz A_2.

Figura 4.1: Valores de $t_k = \|(X_k - X)A_i\|_2$, $e_k = \|X_k - X\|_2$ e $r_k = \|X_{k+1} - X_k\|_2$, para $i = 1, 2$. em escala linear-log para as matrizes do Exemplo 4.3.

k	$\|BX_kB - X\|_2$	$\|X_kBX_k - X_k\|_2$	$\|(BX_k)^* - BX_k\|_2$	$\|(X_kB)^* - X_kB\|_2$
40	$8,86 \times 10^{-13}$	$6,60 \times 10^{-13}$	$4,80 \times 10^{-15}$	$3,26 \times 10^{-15}$
125	$2,11 \times 10^{-15}$	$2,75 \times 10^{-2}$	$5,28 \times 10^{-15}$	$2,92 \times 10^{-15}$
200	$2,70 \times 10^{-7}$	$6,45 \times 10^{+7}$	$2,80 \times 10^{-8}$	$2,74 \times 10^{-8}$

Tabela 4.2: Resíduo em cada equação de Penrose durante a execução do método (4.1) para a matriz A_2 do Exemplo 4.3.

Exemplo 4.4. Neste exemplo, aplicamos o método iterativo proposto para resolver o Problema Teste 3.1. A, b e x foram obtidos pela rotina shaw.m de [14] para $n = 128$. A matriz A obtida é muito mal condicionada (número de condição de A: $2,46 \times 10^{20}$). Primeiramente aplicamos o método iterativo, e o critério de padara não foi satisfeito. Então analisamos o resíduo em cada equação de Penrose a cada 25 iterações, os valores são mostrados na Tabela 4.3.

k	$\|AX_kA - A\|_2$	$\|X_kAX_k - X_k\|_2$	$\|(AX_k)^* - AX_k\|_2$	$\|(X_kA)^* - X_kA\|_2$
25	$2,39 \times 10^{-2}$	$112,59$	$8,59 \times 10^{-16}$	$6,63 \times 10^{-16}$
50	$5,91 \times 10^{-4}$	$9,65 \times 10^{+4}$	$6,04 \times 10^{-14}$	$1,78 \times 10^{-14}$
75	$9,67 \times 10^{-6}$	$1,58 \times 10^{+8}$	$5,26 \times 10^{-11}$	$5,49 \times 10^{-13}$
100	$4,29 \times 10^{-7}$	$1,67 \times 10^{+11}$	$5,42 \times 10^{-8}$	$2,09 \times 10^{-11}$
125	$5,78 \times 10^{-9}$	$2,31 \times 10^{+14}$	$4,79 \times 10^{-5}$	$6,91 \times 10^{-10}$
150	$3,20 \times 10^{-9}$	$4,11 \times 10^{+16}$	$1,67 \times 10^{-2}$	$2,17 \times 10^{-8}$
175	$2,07 \times 10^{-7}$	$3,16 \times 10^{+20}$	$103,47$	$8,62 \times 10^{-7}$
200	$1,26 \times 10^{-5}$	$9,07 \times 10^{+23}$	$5,21 \times 10^{+5}$	$5,32 \times 10^{-5}$

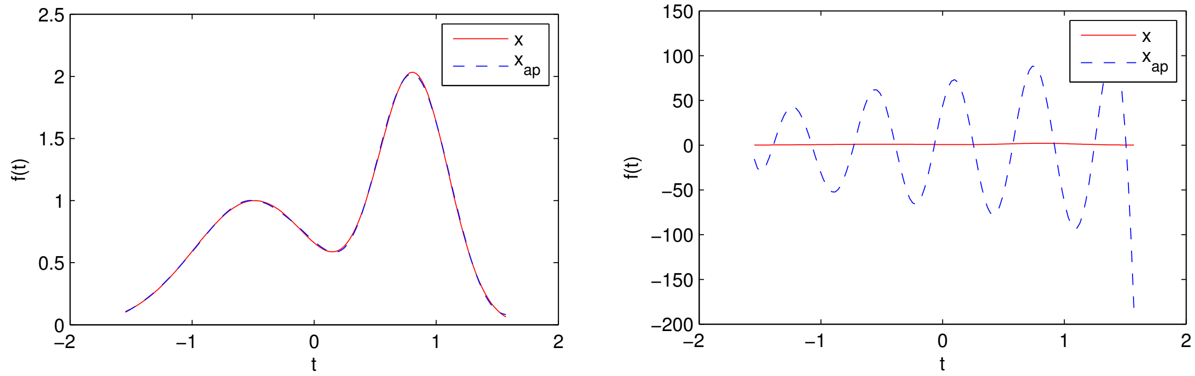
Tabela 4.3: Análise do resíduo em cada equação de Penrose para a sequência X_k, produzida pelo método iterativo 4.1.

Note que o resíduo da equação $AX_kA = A$ diminui conforme iteramos, porém o resíduo da equação $X_kAX_k = X_k$ cresce a cada iteração, fazendo inclusive com que o resíduo nas equações $(AX_k)^* = AX_k$ e $(X_kA)^* = X_kA$ cresça, as quais de acordo com o Lema 4.1 deveriam ser satisfeita para todo k.

Vimos no capítulo 1 que dado um sistema $Ax = b$, caso o sistema seja consistente podemos calcular a solução por $x = Xb$, com X uma inversa-$\{1,4\}$ de A. Caso o sistema seja inconsistente a solução é dada por $x = Xb$ com X sendo uma inversa-$\{1,3\}$ de A.

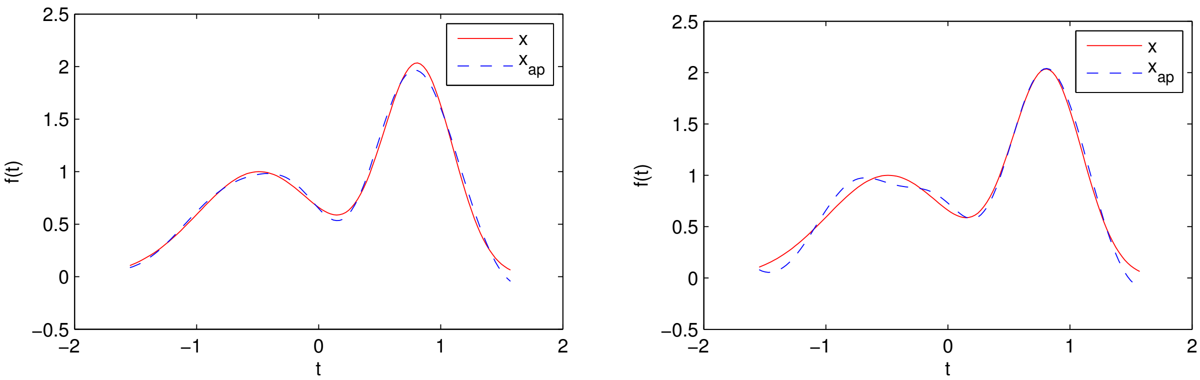
Assim, considerando o comportamento do resíduo nas equações de Penrose para matrizes mal condicionadas apresentado acima, caso o método iterativo proposto seja interrompido adequadamente, podemos ter X_k uma boa aproximação de uma inversa-$\{1,3,4\}$, a qual pode ser usada para computar a solução de sistemas $Ax = b$. Sendo assim, usaremos como critério de parada do método a expressão $\|AX_kA - A\|_2 < tol$, lembrando que exigir uma tolerância muito pequena pode fazer o método divergir.

Aplicamos o método iterativo (4.1) para resolver os sistemas $Ax = b$ e $Ax = \tilde{b}$, sendo $\tilde{b} = b + e$, com as tolerâncias $tol = 10^{-6}$ e $tol = 10^{-3}$. As figuras 4.2 e 4.3 comparam a solução exata x com a solução aproximada x_{ap} para os sistemas $Ax = b$ e $Ax = \tilde{b}$ com $tol = 10^{-6}$ e $tol = 10^{-3}$ respectivamente.



(a) $Ax = b$. (b) $Ax = \tilde{b}$ com NL$= 0,0012$.

Figura 4.2: Solução exata do Problema Teste 3.1 e solução aproximada construída por Xb e $X\tilde{b}$ respectivamente, sendo X calculada pelo método iterativo proposto com critério de parada $\|AX_kA - A\|_2 < 10^{-6}$.



(a) $Ax = b$. (b) $Ax = \tilde{b}$ com NL$= 0,0012$.

Figura 4.3: Solução exata do Problema Teste 3.1 e solução aproximada construída por Xb e $X\tilde{b}$ respectivamente, sendo X calculada pelo método iterativo proposto com critério de parada $\|AX_kA - A\|_2 < 10^{-3}$.

Como podemos observar, o método iterativo funciona bem para matrizes bem condicionadas. Quando aplicados em matrizes mal condicionadas, o método não conseguiu computar uma aproximação da inversa de Moore-Penrose, porém se um bom critério de parada for usado, podemos obter uma aproximação de uma inversa-$\{1,3,4\}$, a qual possui importantes aplicações, como por exemplo para resolver sistemas da forma $Ax = b$, tanto para $b \in \mathcal{R}(A)$ quanto para $b \notin \mathcal{R}(A)$.

Conclusão

Nesta tese, apresentamos os principais métodos computacionais para calcular inversas generalizadas existentes na literatura, a partir dos quais propomos três novos métodos: um método direto baseado em decomposição conjugada para calcular a inversa de Moore-Penrose; um método direto baseado em decomposição conjugada para calcular a inversa de Drazin; e um método iterativo baseado nas equações de Penrose para aproximar a inversa de Moore-Penrose.

Quanto aos métodos baseados em decomposição conjugada, o principal foco para que os métodos sejam competitivos é a execução da decomposição conjugada, que pode ser realizada de diferentes maneiras. Em sua proposição original a decomposição conjugada sugere que seja calculada via método de gradientes conjugados, o qual se mostrou pouco estável para matrizes grandes, como mostra o Exemplo 3.1. Na sua essência o método de gradientes conjugados consiste de um processso de ortonormalização de Gram-Schmidt, onde o conjunto de vetores é escolhido de maneira adequada para que o produto interno do vetor r_k com os componentes ortogonais q_i seja zero para $i = 1, 2, \ldots, k-2$, poupando assim o cálculo de $k-2$ produtos internos na construção do k-ésimo vetor ortonormal. Porém, para matrizes de tamanho elevado os erros de arredondamento fazem com que esses produtos internos não sejam exatamente iguais a zero, e nos últimos vetores o erro acumulado faz com que as colunas da matriz Q não sejam ortonormal. Enquanto isso, o processo de ortonormalização de Gram-Schmidt modificado mantém a ortonormalidade das colunas de Q, o que gera boas aproximações da inversa de Moore-Penrose, conforme mostra o Exemplo 3.2.

Com relação a matrizes mal condicionadas, vimos que o método baseado em decomposição conjuaga pode ser usado para obter uma aproximação da inversa de Moore-Penrose se considerarmos uma tolerância adequada, dessa forma o método aproxima a inversa de Moore-Penrose de uma matriz A por $B^\dagger$, onde a matriz B aproxima A, com posto(B) <posto(A). A questão principal é como estabelecer uma tolerância adequada, o que será alvo de estudos futuros.

Tendo em vista as considerações acima, os métodos baseados em decomposição conjugada são bastante promissores, e podem vir a ser muito competitivos, se a decomposição conjugada for computada por um algoritmo rápido e estável.

Quanto ao método iterativo para aproximar a inversa de Moore-Penrose, baseado em alguns métodos existentes na literatura que usam as equações de Penrose para sugerir a iteração, derivamos um novo método baseado nas equações

$$AXA = A \qquad \text{e} \qquad XAX = X.$$

Nosso método segue o estilo do método proposto por Petković e Stanimirović em [23] e [24], o qual faz uso de um parâmetro β na iteração. Dentro das condições de convergência do método, como podemos observar nos exemplos do capítulo 4, nosso método se mostrou mais rápido e preciso que o método (2.9), o qual é uma generalização do clássico método de ordem 2 de que trata o Teorema 2.31, mostrando que o método iterativo proposto é competitivo.

Para matrizes mal condicionadas, o método iterativo não é adequado para obter a inversa de Moore-Penrose. Ainda assim, o método iterativo pode ser usado para obter uma aproximação de uma inversa-$\{1, 3, 4\}$, que possui importantes aplicações. Para isso, é necessário um bom critério de parada, o que será alvo de estudos futuros.

De maneira geral, este trabalho apresenta um vasto estudo sobre inversas generalizadas, mostrando aspectos da teoria e aspectos computacionais, além de contribuir apresentando novos métodos computacionais promissores.

Como estudos futuros, pretendemos encontrar maneiras rápidas e estáveis de obter a decomposição conjugada de uma matriz, melhorando assim os métodos que são baseados em decomposição conjugada. Expandir o intervalo no qual o parâmetro β garante a convergência do método iterativo para alguns conjuntos de matrizes importantes. Buscar novos métodos para computar inversas generalizadas e principalmente novos métodos para resolver o sistema $Ax = b$, sem que seja necessário o cálculo explícito de $A^\dagger$.

Referências Bibliográficas

[1] BEN-ISRAEL, A. **A note on an iterative method for generalized inversion of matrices**, Math. Comp. 20 (1966), 439-440.

[2] BEN-ISRAEL, A. **An iterative method for computing the generalized inverse of an arbritary matrix**, Math. Comp. 19 (1965), 452-455.

[3] BEN-ISRAEL, A. CHARNES, A. **Contributions to the theory of generalized inverses**, SIAM J. 11 (1963), 667-669.

[4] BEN-ISRAEL, A. COHEN, D. **On iterative computation of generalized inverses and associated projections**, SIAM J. Numer. Anal. 3 (1966), 410-419.

[5] BEN-ISRAEL, A. GREVILLE, T. N. E. **Generalized inverses: Theory and Applications**, Springer Verlag , New York, 2003.

[6] BOULDIN, R. **The pseudo-inverse of a product**, SIAM Journal on Applied Mathematics 24(4) (1973), 489-495.

[7] CAMPBELL, S. L. MEYER JR., C. D. **Generalized Inverse of Linear Transformation**, Pitman, London, 1979.

[8] CLINE, R. E. **Representation of the generalized inverse of a partitioned matrix**, J. Soc. Indust. Appl. Math., 12 (1964), 588-600.

[9] CLINE, R. E. GREVILLE, T. N. E. **A Drazin inverse for rectangular matrices**, Linear Algebra Appl. 29 (1980), 54-62

[10] COURRIEU, P. **Fast Computation of Moore-Penrose Inverse Matrices**, Neural Information Processing-Letters and Reviews 8(2) (2005), 25-29.

[11] DEMMEL, J. W. **Applied Numerical Linear Algebra**, SIAM, Philadelphia, 1997.

[12] GOLUB, G. H. VAN LOAN, C. **Matrix Computation**, John Hopkins University Press, 3rd. Edition, 1996.

[13] GREVILLE, T.N.E. **Note on the generalized inverse of a matrix product**, SIAM Review 8 (1966), 518–521.

[14] HANSEN, P. C. **Regularization Tools: A MATLAB package for analysis and solution of discrete ill-posed problems**, Numer. Algorithms, 6 (1994), 1-35.

[15] HIGHAM, N. J. **Accuracy and stability of numerical algorithms**, SIAM J. Sci. Philadelphia, 1996.

[16] HORN, R. A. JOHNSON, C. R. **Matrix Analysis**, Cambridge University Press, Cambridge, New York, New Rochelle, Melbourne, Sydney, 1986.

[17] HOUSEHOLDER, A. S. **The theory of matrices in numerical analysis**, Blaisdell, New York, 1964.

[18] MOORE, E. H. **On the reciprocal of the general algebraic matrix**, Bull, Amer. Math. Soc., 26 (1920), 394-395.

[19] MOROZOV, V. A. **Regularization methods for ill-posed problems**, CRC press, Florida,1993.

[20] NAJAFI, H. S. SOLARY, M. S. **Computational algorithms for computing the inverse of a square matrix, quasi-inverse of a non square matrix and block matrices**, Appl. Math. Comput., 183 (2006), 539-550.

[21] PAN, V. SEHREIBER, R. **An improved Newton iteration for the generalized inverse of a matrix, with applications**, SIAM J. Sci. Stat. Comput. 12 (1991), 1109-1130.

[22] PENROSE, R. **A generalized inverse for matrices**, Proc. Cambridge Philos. Soc., 51 (1955), 406-413.

[23] PETKOVIĆ, M. D. STANIMIROVIĆ, P. S. **Iterative method for computing the Moore-Penrose inverse based on Penrose equations**, J. Comput. Appl. Math. 235 (2011), 1604-1613.

[24] PETKOVIĆ, M. D. STANIMIROVIĆ, P. S. **Two improvements of the iterative method for computing Moore-Penrose inverse based on Penrose equations**, J. Comput. Appl. Math. 267 (2014), 61-71.

[25] PIERCE, W. H. **A self-correcting matrix iteration for the Moore-Penrose generalized inverse**, Linear Algebra Appl. 244 (1996), 357-363.

[26] QUARTERONI, A. SACCO, R. SALERI, F., **Numerical Mathematics**, Springer, Nova York, 2000.

[27] SCHULZ, G. **Iterative berechmmg der reziproken matrix**, Z. Angew. Math. Mech 13 (1933), 57-59.

[28] TOUTOUNIAN, F. ATAEI, A. **A new method for computing Moore-Penrose inverse matrices**, J. Comput. Appl. Math. 228 (2009), 412-417

[29] WANG, G. WEI, Y. QIAO, S. **Generalized Inverses: Theory and Computations**, Science Press, Beijing/New York, 2004.

[30] WANG, L. P. YUAN, J. Y. **Conjugate Decomposition and its Applications**, Journal of the Operation Research society of China, 1 (2013), 199–215.

[31] ZHONG, J. LIU, X. ZHOU, G. YU, Y. **A new iterative method for computing the Drazin inverse**, Filomat 26:3 (2012), 597–606.