Monitoria MAT1202 - Álgebra Linear 2
Apostila Notas de Aula

Matheus Nogueira

Resumo

Este documento consiste nas notas de aula da monitoria de MAT1202. Este material foi produzido com base em minhas anotações do curso de Álgebra Linear 2 do semestre 20.2 e do livro *Álgebra Linear e suas aplicações*, de *Gilbert Strang*. Qualquer dúvida, favor entrar em contato *matnogueira@gmail.com*

Sumário

1 Sistemas Lineares e Eliminação Gaussiana

1.1 Sistemas Lineares e Notação Matricial

Nosso foco é estudar sistemas de equações da forma $Ax = b$, onde A é a matriz com os termos que acompanham as variáveis (incógnitas), x é o vetor coluna com as incógnitas e b é o vetor coluna com os termos independentes.

Exemplo: Seja o seguinte sistema de equações...

$$\begin{aligned} x + 2y + 3 = 2 \\ -x + y - z = -3 \\ 2x + y - z = 0 \end{aligned}$$

Escrevê-lo em forma matricial é definir as seguinte matriz e vetores:

$$A = \begin{bmatrix} 1 & 2 & 3 \\ -1 & 1 & -1 \\ 2 & 1 & -1 \end{bmatrix}, \; x = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \;\text{ e } b = \begin{bmatrix} 2 \\ -3 \\ 0 \end{bmatrix}$$

Não é difícil perceber que a multiplicação representada por Ax resulta exatamente no sistema linear inicial.

1.2 Solução de um Sistema Linear

Nossa estratégia para calcular a solução de um sistema de equações lineares será a **Eliminação Gaussiana**.

Este método consiste em realizar operações na matriz do sistema $Ax = b$, chamadas *operações elementares*, para chegar a um *sistema triangular*. Ao ser obtido este sistema, basta realizar uma série de substituições retroativas para chegar à solução.

Definição: Matrizes Triangulares

Uma matriz é triangular - superior ou inferior - se todas as entradas abaixo ou acima, respectivamente, da diagonal principal são nulas. A matriz A abaixo é triangular superior, enquanto que B é triangular inferior.

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & -1 \\ 0 & 0 & -1 \end{bmatrix}, \; B = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 2 & 1 & -1 \end{bmatrix}$$

São 3 os possíveis tipos de solução de um sistema linear:

1. Exatamente 1 solução
2. Infinitas Soluções
3. Não há solução

Observação: lembrem-se que, para verificar qual das opções acima é a o caso da matriz a ser estudada, podemos olhar para o *determinante* da matriz. Se seu valor for zero, o sistema possui infinitas soluções ou nenhuma solução. Se for diferente de zero, uma solução.

1.2.1 Operações Elementares

Definição: dado um sistema linear $Ax = b$, são 3 as operações elementares que não alteram a solução do sistema.

1. Permutação de linhas ($L_i \leftrightarrow L_j$)
2. Multiplicação de linha por escalar ($L_i \rightarrow L_i \cdot k, k \neq 0$)
3. Somar um múltiplo de uma linha a outra linha ($L_i \rightarrow L_i + k \cdot L_j$)

1.2.2 Matrizes das operações elementares

Veremos que cada uma das 3 operações elementares descritas pode ser representada por meio de matrizes da seguinte forma:

Se queremos realizar a operação elementar e sobre a matriz A, devemos realizar a multiplicação $E \cdot A$, onde E é a matriz que representa a operação elementar e.

Vejamos as como montar as matrizes para as mesmas 3 operações já apresentadas. Por facilidade, usaremos matrizes $3x3$, pois o raciocínio para outras dimensões é o mesmo. Começamos sempre com a matriz identidade e:

1. Permutação de linhas ($L_i \leftrightarrow L_j$):
   basta permutar as linhas da matriz identidade de acordo com as linhas a serem permutadas na matriz A
2. Multiplicação de linha por escalar ($L_i \rightarrow L_i \cdot k, k \neq 0$):
   multiplicamos a linha correspondente da matriz identidade pelo escalar em questão.
3. Somar um múltiplo de uma linha a outra linha ($L_i \rightarrow L_i + k \cdot L_j$):
   colocamos na entrada i, j da matriz identidade o valor de k com o devido sinal.

$$L_2 \leftrightarrow L_3 \implies \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \leftrightarrow \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} = E \tag{1}$$

$$L_2 \rightarrow L_2 \cdot k \implies \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \leftrightarrow \begin{bmatrix} 1 & 0 & 0 \\ 0 & k \cdot 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = E \tag{2}$$

$$L_3 \rightarrow L_3 - 2 \cdot L_1 \implies \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \leftrightarrow \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix} = E \tag{3}$$

Ao final da *Eliminação Gaussiana*, depois de serem realizadas todas as devidas *operações elementares*, a matriz obtida estará na forma **escalonada**, isto é:

1. Se existem linhas nulas elas devem ser as últimas da matriz.
2. Em quaisquer duas linhas sucessivas não nulas, o pivô (primeiro elemento não nulo) da linha inferior deve estar mais à direita que o da linha superior.
3. Abaixo do pivô todas as entradas são nulas.

1.3 Exemplo

Calculemos a solução do seguinte sistema, mostrando as matrizes das operações elementares.

$$\begin{aligned} 2x + y + z &= 5 \\ 4x - 6y &= 2 \\ -2x + 7y + 2z &= 9 \end{aligned}$$

Em forma matricial o sistema é:

$$\begin{bmatrix} 2 & 1 & 1 \\ 4 & -6 & 0 \\ -2 & 7 & 2 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 5 \\ -2 \\ 9 \end{bmatrix}$$

Seja a matriz aumentada a ser escalonada a seguir:

$$\begin{bmatrix} 2 & 1 & 1 & 5 \\ 4 & -6 & 0 & -2 \\ -2 & 7 & 2 & 9 \end{bmatrix}$$

Comecemos as operações elementares para chegar à matriz escalonada. A cada operação, indicaremos a matriz E correspondente.

$$L_2 \to L_2 - 2L_1 \text{ sendo } E_1 = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Nosso sistema fica...

$$\begin{bmatrix} 2 & 1 & 1 & 5 \\ 0 & -8 & -2 & -12 \\ -2 & 7 & 2 & 9 \end{bmatrix}$$

$$L_3 \to L_3 + L_1 \text{ sendo } E_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

Nosso sistema fica...

$$\begin{bmatrix} 2 & 1 & 1 & 5 \\ 0 & -8 & -2 & -12 \\ 0 & 8 & 3 & 14 \end{bmatrix}$$

$$L_3 \to L_3 + L_2 \text{ sendo } E_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

Nosso sistema fica...

$$\begin{bmatrix} 2 & 1 & 1 & 5 \\ 0 & -8 & -2 & -12 \\ 0 & 0 & 1 & 2 \end{bmatrix}$$

Chegamos à matriz escalonada. Agora basta realizar algumas substituições retroativas para calcularmos a solução.

Lendo e substituindo o sistema de baixo para cima temos:

$$\begin{aligned} z &= 2 \\ -8y - 2(2) &= -12 \rightarrow y = 1 \\ 2x + 1 + 2 &= 5 \rightarrow x = 1 \end{aligned}$$

Note que chegamos a uma solução única, o que faz sentido pois $\det(A) = -16 \neq 0$

Utilizando as matrizes das operações elementares, chegaríamos na mesma matriz escalonada:

$$E_3 \cdot E_2 \cdot E_1 \cdot A \text{ , onde A é a matriz aumentada}$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 2 & 1 & 1 & 5 \\ 4 & -6 & 0 & -2 \\ -2 & 7 & 2 & 9 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 1 & 5 \\ 0 & -8 & -2 & -12 \\ 0 & 0 & 1 & 2 \end{bmatrix}$$

1.4 Conclusão

Com este material sabemos como encontrar a solução de um sistema linear utilizando a Eliminação Gaussiana e as operações Elementares, com suas respectivas matrizes. O próximo assunto a ser abordado será **Fatoração LU**.

2 Fatoração A=LU

2.1 Sem permutação de linhas

No capítulo anterior vimos, ou relembramos, como resolver um sistema linear utilizando o processo da Eliminação Gaussiana por meio, principalmente, de operações elementares e suas matrizes. Neste capítulo continuaremos estudando sistemas lineares do tipo $Ax = b$ e apresentaremos uma maneira de fatorar a matriz A, escrevendo-a como $A = LU$.

Dito isso, já podemos definir a matriz A como a matriz de coeficientes do nosso sistema linear, ou seja, exatamente a mesma matriz A do capítulo anterior. Nosso sistema linear é:

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} b_{11} \\ b_{21} \\ b_{31} \end{bmatrix} \text{ logo, } A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

A matriz U é a matriz triangular superior que aparece ao final do processo de escalonamento da matriz A, obtida por meio das operações elementares. Você deve se lembrar que, em nossa aula 2 de *MATLAB*, aprendemos a função $[\ , U] = lu(A)$, sendo U o nome dado à variável que guarda o output da função *lu()*, isto é, a matriz escalonada resultante da eliminação gaussiana. Com U em mãos, tudo que nos restava fazer era uma substituição retroativa para descobrir a solução do sistema.

A última matriz que falta ser descoberta é L. Para isso, precisamos lembrar das matrizes E_i que representam as operações elementares. Se nos recordarmos, para escalonar A até U fazíamos:

$$U = E_n \cdot E_{n-1} \cdot ... \cdot E_2 \cdot E_1 \cdot A$$

sendo n o número de operações elementares a serem feitas.

Chamemos de E a matriz resultante de todas as multiplicações de E_i. Podemos reescrever a equação acima como $U = E \cdot A$. Queremos chegar na faturação $A = LU$, logo, não é difícil perceber que basta multiplicar ambos os lados de $U = E \cdot A$ por E^{-1} à esquerda que obteremos algo similar à fatoração desejada.

$$E^{-1}U = E^{-1}E \cdot A \to E^{-1} \cdot U = A$$

De fato, a matriz L da fatoração $A = LU$ é, justamente, a multiplicação de todas as inversas das matrizes elementares utilizada. Sendo assim, definimos

$$L = E_1^{-1} \cdot E_2^{-1} \cdot ... \cdot E_n^{-1}$$

Convença-se de que L está corretamente definida!

O único empecilho para esta definição é garantir que todas as matrizes elementares são inversíveis. Para isso, seus determinantes devem ser diferentes de 0. Como estamos estudando, nesta seção, apenas o caso sem trocas de linha, é trivial notar que todas as matrizes E_i possuem 1 em sua diagonal principal e são triangulares inferiores. Sendo assim, seus determinantes são sempre 1. Convença-se deste fato.

Agora podemos apresentar a versão completa da função do *MATLAB*, $[L, U] = lu(A)$. Esta função retorna, não somente a matriz escalonada U, como a matriz L, o que faz todo o sentido dado o nome da função... A análise da matriz L é importante para ver se houve trocas de linha na execução interna do algoritmo da função.

Com todas estas definições em mãos, podemos partir para um exemplo.

2.1.1 Exemplo

Dada a seguinte matriz A, calculemos cada uma das matrizes envolvidas na fatoração $A = LU$ e mostremos que essa igualdade vale.

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 3 & 5 \\ 4 & 6 & 8 \end{bmatrix}$$

Realizando seu escalonamento, chegamos às seguintes matrizes elementares:

$$E_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -4 & 0 & 1 \end{bmatrix}; E_2 = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}; E_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{bmatrix}$$

Podemos verificar (deixo por conta de você, caro aluno) que:

$$E_3 \cdot E_2 \cdot E_1 \cdot A = U \text{ onde, } U = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 3 \\ 0 & 0 & -2 \end{bmatrix}$$

Note que U é uma matriz triangular superior assim como prevê a teoria! Calculemos agora a matriz L.

$$L = E_1^{-1} \cdot E_2^{-1} \cdot E_3^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 4 & 2 & 1 \end{bmatrix} \tag{4}$$

Como previsto, a matriz L é triangular inferior com todas as entradas da diagonal principal igual a 1.

Dica: para inverter uma matriz elementar basta trocar o sinal da entrada não nula fora da diagonal principal.

Podemos, por fim, verificar que:

$$L \cdot U = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 4 & 2 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 3 \\ 0 & 0 & -2 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 3 & 5 \\ 4 & 6 & 8 \end{bmatrix} = A$$

2.2 Fatoração PA=LU (com permutação de linhas)

Caso seja necessário realizar alguma permutação de linhas a fim de garantir que U será uma matriz escalonada, precisamos corrigir a matriz A, introduzindo as permutações necessárias para, então, realizar a fatoração LU. Uma vez detectadas as permutações realizadas, podemos carregar essa informação em uma matriz P e multiplicá-la por A de modo que

$$PA = LU$$

Vejamos um exemplo.

2.2.1 Exemplo

Dada a seguinte matriz A, calculemos cada uma das matrizes envolvidas na fatoração $A = LU$, mostremos que serão necessárias permutações, montemos a matriz P e verifiquemos a validade da igualdade $PA = LU$.

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

Para simplificar as contas, usemos a função do $MATLAB$ $[L, U] = lu(A)$. O retorno desta função é:

$$L = \begin{bmatrix} 0.5 & 0 & 1 \\ 1 & 0 & 0 \\ 0.5 & 1 & 0 \end{bmatrix} ; U = \begin{bmatrix} 2 & 4 & 5 \\ 0 & 1 & 1.5 \\ 0 & 0 & 0.5 \end{bmatrix}$$

A matriz L, neste caso, não é triangular inferior, o que indica que o algoritmo interno da função realizou permutações na matriz A. Precisamos, então, montar a matriz P de permutações. Podemos começar trocando as linhas 2 e 3. Para isso, chamemos de P_1 a seguinte matriz de modo que...

$$P_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \text{ de modo que } P_1 \cdot L = \begin{bmatrix} 0.5 & 0 & 1 \\ 0.5 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

Agora definimos P_2 a partir da troca das linhas 1 e 3.

$$P_2 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \text{ de modo que } P_2 \cdot P_1 \cdot L = \begin{bmatrix} 1 & 0 & 0 \\ 0.5 & 1 & 0 \\ 0.5 & 0 & 1 \end{bmatrix}$$

Definimos, então, $P = P_2 \cdot P_1$. O valor da matriz P está exibido logo abaixo.

Agora a matriz L possui as características necessárias segundo a teoria, isto é, ser triangular inferior e possuir todas as entradas da diagonal principal igual a 1.

Podemos, finalmente, verificar que:

$$PA = LU \leftrightarrow \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 1 & 3 & 4 \end{bmatrix} = \begin{bmatrix} 0.5 & 0 & 1 \\ 1 & 0 & 0 \\ 0.5 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} 2 & 4 & 5 \\ 0 & 1 & 1.5 \\ 0 & 0 & 0.5 \end{bmatrix}$$

2.3 Conclusão

Neste capítulo foram apresentados os conceitos de fatoração $A = LU$ em seu caso sem permutação de linhas e $PA = LU$ quando são necessárias essas permutações. Este é um algoritmo importante para a compreensão dos métodos de resolução de sistemas lineares e seu entendimento desmistifica o funcionamento da função $lu()$ utilizada no *MATLAB*.

3 Espaços Fundamentais de uma Matriz

3.1 Definições:

Estudaremos os 4 subespaços fundamentais de uma matriz. Para todo este estudo, considere A uma matriz $m \times n$ São eles:

1. Espaço Coluna, ou Imagem
2. Espaço Linha
3. Espaço Nulo, ou Núcleo
4. Espaço Numo da transposta

3.1.1 Espaço Coluna - $Im(A)$

O espaço coluna, ou imagem de uma matriz A é o subespaço vetorial gerado pelas colunas da matriz A.

Def:

$$Im(A) = \{v \in R^m \text{ tal que } A \cdot u = v \text{ para algum } u \in R^n\}$$

É importante lembrar de alguns conceitos como *Espaços e Subespaços Vetoriais* e *Independência Linear*, uma vez que nada garante que as m colunas sejam LI e gerem um espaço de dimensão m.

Base para $Im(A)$: podemos fazer uma Eliminação Gaussiana de A e observar quais colunas da matriz U resultante deste processo possuem pivôs. Se as colunas c_i de U possuem pivô, então as colunas c_i de A serão base da *Imagem* de A. Consegue perceber o por quê?

Observação Importante: $Im(A) \neq Im(U)$

Posto de uma matriz A é a dimensão da *Imagem* dessa matriz $A \rightarrow posto(A) = dim(Im(A))$

3.1.2 Espaço Nulo - $N(A)$

O espaço nulo de A é o espaço vetorial gerado pelos vetores x tal que $A \cdot x = 0$.

Def:

$$N(A) = \{x \in R^n \text{ tal que } A \cdot x = 0\}$$

3.1.3 Espaço Linha - $Im(A^T)$

O espaço linha de A é o espaço vetorial gerado pelos vetores linha de A. De modo análogo, é o *espaço coluna da matriz transposta* de A.

Def:

$$Im(A^T) = \{v \in R^m \text{ tal que } A \cdot u = v \text{ para algum } u \in R^n\}$$

Outra maneira de encontrar o espaço linha de A é, novamente, por meio da Eliminação Gaussiana. Note que, se U for a matriz escalonada da Eliminação Gaussiana, então o espaço linha de U é igual ao espaço linha de U. Isso quer dizer que uma base de $Im(U^T)$ é também base de $Im(A^t)$.

3.1.4 Espaço Nulo da Transposta- $N(A^T)$

O espaço nulo da transposta de A é o espaço vetorial gerado pelos vetores x tal que $A^T \cdot x = 0$.

Def:

$$N(A^T) = \{x \in R^m \text{ tal que } A^T \cdot x = 0\}$$

3.2 Exemplo:

Encontre os 4 espaços fundamentais da matriz abaixo.

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \\ 3 & 3 & 3 \end{bmatrix}$$

Para encontrar o espaço coluna de A, vamos escalonar esta matriz. Podemos usar o comando já aprendido $[, U] = lu(A)$, que nos retorna:

$$U = \begin{bmatrix} 3 & 3 & 3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Já podemos perceber a existência de apenas 1 pivô, logo **posto(A)=dim(Im(A))=1**. Com além disso, como o pivô está na primeira coluna de U, a base da imagem de A será formada pela primeira coluna de A. Também podemos usar a função *colspace(sym((A))* do MATLAB.

$$\beta_{Im(A)} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

Para o núcleo de A devemos resolver o sistema linear $A \cdot x = 0$ e os vetores x que satisfizerem esta igualdade serão nosso núcleo. Analogamente, e para facilitar nossa vida, podemos usar o comando *B=null(sym(A))*, que retorna:

$$B = \begin{pmatrix} -1 & -1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Podemos confirmar esta resposta multiplicando $A * B$ e verificando que esta conta dá **zero**.

Base: para verificar que estes vetores nas colunas de B são base do núcleo, devemos verificar que eles são *LI*. Uma vez confirmado, temos que:

$$\beta_{N(A)} = \{ \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \}$$

Observação: note que $dim(N(A)) = 2$. Isso faz sentido pois, lembrando de Álgebra 1, $dim(Im(A)) + dim(N(A)) = n$.

Para calcularmos o espaço linha, temos duas opções. Primeiro, transpor a matriz A e calcular a imagem desta nova matriz da maneira já explicada. Por exemplo: *colspace(sym(transpose(A)))*. Outra maneira é realizar a fatoração LU e olhar para as linhas de U, uma vez que $Im(U^T) = Im(A^T)$. Como já temos o resultado da função $lu(A)$, podemos notar que a primeira linha de U é base do espaço linha de U. Logo,

$$\beta_{Im(A^T)} = \begin{pmatrix} 3 \\ 3 \\ 3 \end{pmatrix}$$

Finalmente, para o espaço nulo da transposta, podemos utilizar um processo similar ao cálculo do espaço nulo. Será que *null(sym(transpose(A)))*, que retorna:

$$\begin{pmatrix} -2 & -3 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Verificamos se estes vetores são, de fato, a base do núcleo da transposta ao verificar que eles satisfazem $transpose(A) \cdot null(symtranspose(A))) = 0$ e que eles são LI. Por fim, temos

$$\beta_{N(A)} = \{ \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -3 \\ 0 \\ 1 \end{pmatrix} \}$$

3.3 Conclusão

Neste capítulo foram apresentados os 4 espaços fundamentais de uma matriz qualquer, bem como os procedimentos necessários para calcular estes subespaços vetoriais. Na próxima aula veremos relações de ortogonalidade entre estes subespaços.

4 Ortogonalidade

4.1 Produto Interno

O conceito de produto interno já é comum a nós. Sendo assim, vamos apenas defini-lo brevemente:

DEF: o produto interno entre dois vetores u e v, sendo $u, v \in R^n$, representado por $u \cdot v$ ou $\langle u, v \rangle$, é definido por

$$\sum_{k=1}^{n} x_k y_k = x_1 y_1 + x_2 y_2 + ... + x_n y_n$$

Algumas propriedades importantes do produto interno são:

1. O produto interno é linear para qualquer argumento $\rightarrow$ $\langle ku, v \rangle = k\langle u, v \rangle, k \in R$ e $\langle u + w, v \rangle = \langle u, v \rangle + \langle w, v \rangle, w \in R^n$

2. $\langle u, u \rangle \geq 0$

3. $\langle u, v \rangle = \langle v, u \rangle$

Nós utilizamos o conceito de produto interno para definir *ortogonalidade* da seguinte maneira.

DEF: dois vetores u e v são ditos **ortogonais** se e somente se $\langle u, v \rangle = 0$

4.2 Norma

Podemos pensar na norma de um vetor $u \in R^n$ como o seu "tamanho"ou "comprimento". Para calcular a norma de um vetor, representada por $\|u\|$ sabemos que vale:

$$\|u\| = \sqrt{x_1^2 + x_2^2 + ... + x_n^2}$$

Note, portanto, que podemos expressar a norma de um vetor utilizando o produto interno!

$$\|u\| = \sqrt{\langle u, u \rangle}$$

DEF: Normalização de um vetor
Chamamos de normalização de um vetor o processo de dividi-lo pela sua norma com o intuito do vetor resultante possuir norma igual a 1.

$$\|\frac{u}{\|u\|}\| = 1$$

DEF: dois vetores u e v são ditos **ortonormais** se e somente se $\langle u, v \rangle = 0$ e $\|u\| = \|v\| = 1$

4.3 Ortogonalidade e Espaços Vetoriais

4.3.1 Espaços Ortogonais

DEF: Dizer que V e W são dois espaços vetoriais ortogonais, ou seja $V \perp W$ é

$$V \perp W \iff \langle v, w \rangle = 0, \forall v \in V, \forall w \in W$$

A discussão de espaços ortogonais é interessante para avaliarmos os espaços fundamentais de uma matriz A, estudados nas últimas aulas. D maneira direta, podemos averiguar que as seguintes duplas de espaços são ortogonais:

- Espaço Coluna $Im(A)$ e Espaço Nulo da Transposta $N(A^t)$
- Espaço Linha $Im(A^t)$ e Núcleo $N(A)$

4.3.2 Complemento Ortogonal

DEF: Seja V um subespaço de R^n. O conjunto

$$W = \{w \in R^n : \langle v, w \rangle = 0, \forall v \in V\}$$

forma um subespaço de R^n, chamado de *complemento ortogonal* de V e denotado por $V^\perp$

Propriedades Importantes:

1. $dim(V) + dim(V^\perp) = n$, sendo n a dimensão de R^n
2. $V \cap V^\perp = \emptyset$
3. $V \cup V^\perp = R^n$

4.4 Conclusão

Neste capítulo foram apresentados conceitos de ortogonalidade, produto interno, norma e complementos ortogonais.

5 Projeções Ortogonais

5.1 Conceitos importantes

Embora seja provável que estes conceitos já sejam bem conhecidos, vamos apenas relembrar o que são produto interno e sua relação com ortogonalidade, a Desigualdade de Cauchy-Schwarz e a Desigualdade Triangular.

Def: o produto interno entre dois vetores u e v pode ser definido como:

$$\langle u, v \rangle = \sum_{k=1}^{n} x_k y_k = x_1 y_1 + x_2 y_2 + ... + x_n y_n$$

Uma outra maneira de definir o mesmo produto interno é:

$$\langle u, v \rangle = ||u|| \cdot ||v|| \cdot cos(\theta)$$

,

onde $||u||$ é a norma do vetor u e θ é o ângulo entre os vetores u e v.

A partir desta definição, podemos derivar que dois vetores são ortogonais se e somente se o produto interno entre eles é zero. Isso é justificado por $cos(90) = 0$.

Def: Desigualdade de Cauchy-Schwarz em R^n

$$||u \cdot v|| \leq ||u|| \cdot ||v||$$

Def: Desigualdade Triangular

$$||u + v|| \leq ||u|| + ||v||$$

5.2 Projeção ortogonal sobre um vetor

Imagine que queiramos projetar um vetor v sobre um vetor u. Note que o vetor w da figura 1 abaixo é, justamente, o resultado dessa projeção.

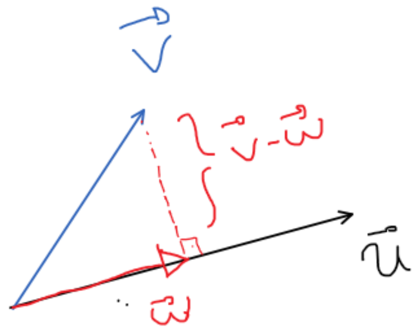

Figura 1: Projeção ortogonal de v em u

Precisamos de duas premissas, as quais são facilmente observadas na figura acima:

$$w = \alpha \cdot u$$
$$\langle u, v - w \rangle = 0$$

A maneira de encontrar, explicitamente, w em função de u e v, com base nessas premissas, está descrita abaixo:

$$\begin{aligned} \langle u, v - w \rangle &= 0 \\ \langle u, v - \alpha u \rangle &= 0 \\ u \cdot v - \alpha \cdot u \cdot u &= 0 \\ \alpha \cdot u \cdot u &= u \cdot v \\ \alpha &= \frac{u \cdot v}{u \cdot u} \end{aligned}$$

Sendo assim, temos:

$$w = \left(\frac{u \cdot v}{u \cdot u}\right) \cdot u$$

Note que o resultado dos produtos internos entre parênteses é um número, garantindo que w e u são paralelos.

5.2.1 Matriz de Projeção sobre o vetor u

Podemos encapsular a projeção ortogonal sobre um vetor u em uma matriz P, da seguinte forma:

$$\begin{aligned}
w &= \left(\frac{u \cdot v}{u \cdot u}\right) \cdot u \\
&= \left(\frac{u^T v}{u^T u}\right) \cdot u \\
&= u\left(\frac{u^T v}{u^T u}\right) \\
&= \left(\frac{u u^T}{u^T u}\right) v \\
&= P \cdot v
\end{aligned}$$

5.2.2 Exemplo:

Calcule a matriz de projeção sobre o vetor $u = (1, -1, 1)$

$$P = \left(\frac{u u^T}{u^T u}\right)$$

$$P = \frac{1}{3}\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}\begin{pmatrix} 1 & -1 & 1 \end{pmatrix}$$

$$P = \frac{1}{3}\begin{pmatrix} 1 & -1 & 1 \\ -1 & 1 & -1 \\ 1 & -1 & 1 \end{pmatrix}$$

5.2.3 Propriedades:

A matriz P definida possui as seguintes propriedades:

1. P é simétrica
2. $P^2 = P$
3. Posto(P)=1

5.3 Projeção Ortogonal sobre um subespaço qualquer

Considere as matriz A e o vetor b definidos abaixo. Além disso, seja w a projeção ortogonal de b sobre o espaço coluna de A.

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix}, b = \begin{pmatrix} b_{11} \\ b_{21} \\ b_{31} \end{pmatrix}$$

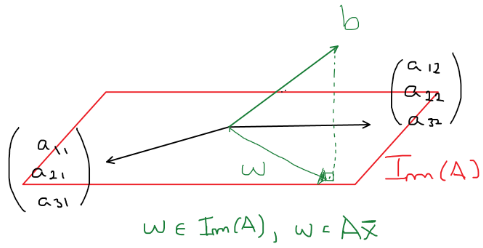

Figura 2: projeção ortogonal sobre imagem de A

Note que:

1. o vetor $b - w$ é ortogonal ao espaço coluna de A
2. o vetor $b - A\overline{x}$ é ortogonal ao espaço coluna de A, sendo $w = A\overline{x}$.
3. Logo, $b - A\overline{x} \in$ Núcleo da Transposta de A
4. $A^T(b - A\overline{x}) = 0$
5. $A^T b = A^T A\overline{x} \rightarrow$ Equação Normal

Obs: se as colunas de A são LI, então $A^T A$ é inversível.

Com isso, chegamos na matriz de projeção sobre o espaço coluna de A:

$$P = A\left(A^T A\right)^{-1} A^T$$

5.4 Conclusão

Neste capítulo foram apresentados conceitos projeções ortogonais sobre u vetor e sobre um subespaço vetorial qualquer.

6 Mínimos Quadrados

6.1 Motivação

Como já sabemos, estamos quase sempre interessados em resolver sistemas do tipo $Ax = b$. No entanto, há casos em que a solução deste sistema é impossível. Por exemplo, se o nosso sistema linear for um sistema linear de equações incompatíveis.

Def: Sistemas lineares incompatíveis são sistemas de m equações e n incógnitas onde $m > n$. Outra maneira de olhar para essa definição é pensar em matrizes A_{mxn}, as quais possuem m linhas (equações) e n colunas (incógnitas). Sabemos que sistemas desse tipo não possuem solução.

O que fazer, pode ser feito, além de abandonar o problema e ir dormir em paz, é procurar uma solução aproximada de tal modo que o **erro** desta solução seja o menor possível.

Def: o erro de uma solução aproximada de um sistema $Ax = b$, é definido como:

$$||Ax - b|| \text{ , onde b é a solução real e Ax é a solução aproximada.}$$

Perceba que, para casos em que podemos encontrar a solução real do sistema, esse erro é zero, uma vez que $Ax = b$, logo $||Ax - b|| = 0$.

Mas e quando não é possível encontrar a solução?

x	y
-1	1
1	1
2	3

6.2 Relembrando Projeções Ortogonais

Lembre-se da capítulo passado sobre projeções ortogonais e acrescente um detalhe.

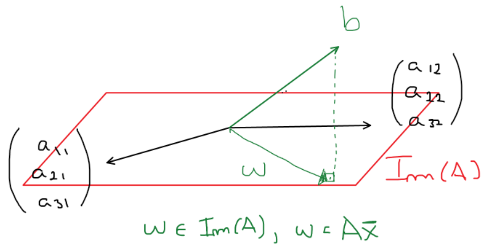

Figura 3: projeção ortogonal sobre imagem de A

No âmbito de minimização de erro, podemos pensar que queremos a menor distância entre a solução real e a aproximada. Sendo assim, a *Projeção Ortogonal* aprendida de mostra muito adequada para esse problema de aproximação.

A própria equação normal $A^T A\overline{x} = A^T b$ relaciona esses valores. Tenha isso em mente.

6.3 Mínimos Quadrados

Uma vez que o sistema linear $Ax = b$ é incompatível, sabemos que $b \notin Col(A)$. Sendo assim, já sabemos que a melhor solução, a que minimiza o erro, será a solução $\overline{x}$ tal que:

$$A\overline{x} = Proj_{Col(A)}b$$

Como vimos na aula anterior, podemos definir uma matriz P de projeção sobre o espaço coluna de A. Utilizando essa informação, chegamos no seguinte sistema:

$$A\overline{x} = Pb$$

Este sistema, agora, é compatível e possui uma solução.

Vejamos, por fim, um exemplo prático de como calcular a melhor reta para um dado conjunto de pontos (regressão linear).

6.3.1 Exemplo:

Encontre a reta que melhor interpola os seguintes pontos:

Dados esses pontos, podemos montar o seguinte sistema linear $Ax = y$:

$$\begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ 1 & x_3 \end{pmatrix} \cdot \begin{pmatrix} b \\ a \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$$

$$\begin{pmatrix} 1 & -1 \\ 1 & 1 \\ 1 & 2 \end{pmatrix} \cdot \begin{pmatrix} b \\ a \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix}$$

Rescrevemos como da seguinte maneira:

$$Ax = y$$
$$A^T Ax = A^T y$$

Colocando valores:

$$\begin{pmatrix} 1 & 1 & 1 \\ -1 & 1 & 2 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 1 & 1 \\ 1 & 2 \end{pmatrix} \cdot \begin{pmatrix} b \\ a \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ -1 & 1 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix}$$

Resolvendo o sistema, encontramos $a = \frac{4}{7}$ e $b = \frac{9}{7}$.

Logo, a reta desejada é: $r(x) = \frac{4}{7} \cdot x + \frac{9}{7}$

6.4 Conclusão

Neste capítulo foi relembrado o conceito de projeções ortogonais sobre um subespaço vetorial qualquer e apresentado o algoritmo de Mínimos Quadrados.

7 Matrizes Ortogonais, Ortogonalização de Grand-Schimdt e Fatoração QR

7.1 Matrizes Ortogonais

Para definir uma matriz ortogonal, precisamos lembrar o que é um conjunto de vetores, ou uma base, se quisermos, ortonormal.

Def: uma base β de vetores de R^n é dita ortonormal se todos os vetores que a formam possuem, simultaneamente, norma 1 e são ortogonais entre si.

Com essa definição, podemos definir tranquilamente uma Matriz Ortogonal.

Def: Uma Matriz Ortogonal A é uma matriz cujas colunas formam uma base ortonormal.

Algumas propriedades das matrizes ortogonais são importantes para nosso curso.

7.1.1 Propriedades

(1) Toda matriz ortogonal Q satisfaz $Q^T Q = I$. Dessa propriedade, nota-se que $Q^T = Q^{-1}$

Prova: note que uma multiplicação matricial é, de certo modo, um produto interno entre as colunas e linhas das matrizes. Sendo assim, como as colunas de Q são ortonormais, $\langle v_i, v_j^T \rangle = 0 \forall i \neq j$, sendo v_i os vetores da coluna de Q e v_j os vetores linhas de Q^T. A partir disso mostramos que $Q^T Q = I$.

(2) A multiplicação de uma matriz ortogonal por um vetor preserva o comprimento do vetor $\rightarrow ||Qx|| = ||x||$.

Dem: $||Qx||^2 = \langle Qx, Qx \rangle = (Qx)^T Qx = x^T Q^T Qx = x^T x = ||x||^2$

(3)) O ângulo entre vetores se preserva por matriz ortogonal.

Dem: $cos\theta = \dfrac{\langle Qu, Qv \rangle}{||Qu|||Qv|||} = \dfrac{uQ^TQv}{||u|||v|||} = \dfrac{u^Tv}{||u|||v|||}$

7.2 Ortogonalização de Grand Schimdt

O processo de ortogonalização de Grand-Schimdt consiste em uma estratégia para, a partir de uma base qualquer $\{a_1, ..., a_n\}$, obter uma base ortonormal $\{q_1, ...q^n\}$ para o espaço gerado pela base a. Temos a seguinte construção:

$$a_1^{'} = a_1 \rightarrow q_1 = \frac{a_1^{'}}{||a_1^{'}||}$$
$$a_2^{'} = a2 - (q_1^T a_2)q_1 \rightarrow q_2 = \frac{a_2^{'}}{||a_2^{'}||}$$
$$a_3^{'} = a3 - (q_1^T a_3)q_1 - (q_2^T a_3)q_2 \rightarrow q_3 = \frac{a_3^{'}}{||a_3^{'}||}$$
$$a_n^{'} = an - (q_1^T a_n)q_1 - ... - (q_{n-1}^T a_n)q_{n-1} \rightarrow q_n = \frac{a_n^{'}}{||a_n^{'}||}$$

7.3 Fatoração QR

O processo de ortogonalização de Grand Schimdt nos entrega uma fatoração conhecida como **fatoração** $A = QR$. Não é difícil imaginar que a matriz Q desta fatoração será a matriz cujas colunas são a base $\{q_1, ...q^n\}$ obtida no processo de ortogonalização e A é a matriz cujas colunas são a base $\{a_1, ..., a_n\}$.

Antes de definirmos a matriz R, lembremos:

Se $\{q_1, ...q^n\}$ é base ortonormal de R^N e b pertence a esse espaço gerado, então...

$$b = c_1 q_1 + c_2 + q_2 + ... + c_n q_n$$

onde

$$c_1 = q_1^T b, c_2 = q_2^T b ... c_n = q_n^T b$$

Vejamos como construir a matriz R para o caso de uma base do R^3. Esse procedimento é facilmente generalizado para R^N.

Exemplo: Seja $\{a, b, c\}$ base para o R^3 e $\{q_1, q_2, q^3\}$ base ortonormal obtida via ortogonalização de Grand Schimdt a partir da base $\{a, b, c\}$.

Temos que:

$a = (q_1^T a)q_1 + (q_2^T a)q_2 + (q_3^T a)q_3 = (q_1^T a)q_1$

A primeira coluna de R será, então, $(q_1^T a, 0, 0)^T$

$b = (q_1^T b)q_1 + (q_2^T b)q_2 + (q_3^T b)q_3 = (q_1^T b)q_1 + (q_2^T b)q_2$

A segunda coluna de R será, então, $(q_1^T a, q_2^T b, 0)^T$

$a = (q_1^T c)q_1 + (q_2^T c)q_2 + (q_3^T c)q_3$

A terceira coluna de R será, então, $(q_1^T a, q_2^T b, q_3^T c)^T$

Logo, $R = \begin{pmatrix} q_1^T a & q_1^T a & q_1^T a \\ 0 & q_2^T b & , q_2^T b \\ 0 & 0 & q_3^T c \end{pmatrix}$

7.4 Conclusão

Neste capítulo apresentados os conceitos de Matriz Ortogonal, Processo de Ortogonalização de Grand-Schimdt e Fatoração $A = QR$.