Álgebra Linear

Um Livro Colaborativo

19 de Dezembro de 2017

Organizadores

Diego Marcon Farias - UFRGS

Rafael Rigão Souza - UFRGS

Pedro Henrique de Almeida Konzen - UFRGS

Licença

Este trabalho está licenciado sob a Licença Creative Commons Atribuição-CompartilhaIgual 3.0 Não Adaptada. Para ver uma cópia desta licença, visite http://creativecommons.org/licenses/by-sa/3.0/ ou envie uma carta para Creative Commons, PO Box 1866, Mountain View, CA 94042, USA.

Nota dos organizadores

Este livro é parte de nosso projeto de escrita colaborativa de recursos educacionais abertos. Nosso objetivo é produzir um material didático no nível de graduação de excelente qualidade e de acesso livre pela colaboração entre professores e alunos de universidades, institutos de educação e demais interessados em tópicos de álgebra linear.

O sucesso deste projeto depende da colaboração! Edite você mesmo o livro, dê sugestões ou nos avise de erros e imprecisões. Toda a colaboração é bem vinda. Saiba mais visitando o repositório oficial do livro:

https://github.com/livroscolaborativos/AlgebraLinear

Desejamos-lhe ótimas colaborações!

Prefácio

Em construção ... Caso deseja colaborar com a escrita deste livro, veja como em: https://github.com/livroscolaborativos/AlgebraLinear

Conteúdo

1

Semana 1

1.1 Sistemas Lineares

Começamos nosso estudo diretamente com um exemplo:

Exemplo 1. Considere o sistema linear que consiste de duas equações e duas variáveis

$$\left\{\begin{array}{rcl} x+3y & = & 1 \\ 2x-y & = & -2 \end{array}\right. . \tag{1.1}$$

Nosso objetivo é descobrir o valor (ou todos os valores) das variáveis x e y que satisfazem a ambas equações.

Solução 1. Ao multiplicar toda a primeira equação por -2, não alteramos a solução do sistema, pois, caso quiséssemos voltar ao sistema original, bastaria dividir a primeira equação do sistema abaixo por -2. Ficamos assim com o sistema

$$\left\{\begin{array}{rcl} -2x-6y & = & -2 \\ 2x-y & = & -2 \end{array}\right. .$$

Em seguida, somamos as duas equações e mantemos uma delas.

$$\left\{\begin{array}{rcl} -2x-6y & = & -2 \\ -7y & = & -4 \end{array}\right. .$$

Este sistema também tem as mesmas soluções do anterior, pois podemos a qualquer hora retornar ao sistema original fazendo (linha 2) – (linha 1).

Em verdade, manter a equação da primeira linha multiplicada por -2 não ajuda em nada na resolução do sistema e voltaremos à equação original. Além disso, na linha dois já podemos isolar a variável y e descobrir o seu valor. Ficamos então com o sistema:

$$\left\{\begin{array}{rcl} x+3y & = & 1 \\ -7y & = & -4 \end{array}\right. . \tag{1.2}$$

Observamos, finalmente, que a solução do último sistema (que é a mesma do sistema original) já está muito fácil de ser obtida. Sabemos o valor de y pela segunda equação e, para descobrir o valor de x, basta usar a primeira equação:

$$y=\frac{4}{7} \implies x+3\cdot\frac{4}{7}=1 \implies x=1-\frac{12}{7}=-\frac{5}{7}.$$

Podemos escrever a solução do sistema como

$$\left\{\begin{array}{rcl} x & = & -5/7 \\ y & = & 4/7 \end{array}\right. \quad \triangleleft$$

O símbolo $\triangleleft$ acima nas nossas notas indica o fim de uma solução, exemplo ou observação.
Em seguida, fazemos algumas observações.

Observação 2. O sistema linear como o acima possui duas equações e duas variáveis e por isto é conhecido como um sistema linear 2×2 (lê-se dois por dois). Mais geralmente, um sistema linear com m equações e n variáveis é conhecido como um sistema linear $m \times n$ (lê-se m por n).

Observação 3. Sistemas 2×2 são dos mais simples de resolver e por isso o método acima pode parecer desnecessariamente complicado. Mas, como veremos, este método pode ser aplicado para qualquer sistema. Logo, é essencial que se entenda completamente todos os seus passos para um bom acompanhamento do que está por vir.

Vamos analisar a solução apresentada acima mais de perto. Para chegar do sistema linear em (1.1) para o sistema em (1.2), fizemos uma sequência de operações que:

- *não alteram a solução do sistema linear original e que;*
- resumidamente, podem ser descritas como *adicionar um múltiplo da linha um na linha dois.*

Que a solução permanece a mesma pode ser justificado tal qual está na "Solução 1" acima.

Operações nas linhas de um sistema que não alteram o conjunto de soluções são chamadas de **operações elementares**. São as seguintes:

1. Multiplicar uma linha por uma constante;
2. Trocar duas linhas de lugar;
3. Somar um múltiplo de uma linha a outra linha.

Um artifício que torna todo o processo mais "automático" é o de formar uma tabela de números com os coeficientes do sistema:

$$\begin{cases} x + 3y &= 1 \\ 2x - y &= -2 \end{cases} \quad \rightsquigarrow \quad \left[\begin{array}{cc|c} 1 & 3 & 1 \\ 2 & -1 & -2 \end{array}\right].$$

A tabela de números da direita é conhecida como a **matriz aumentada associada ao sistema linear**. Vamos nos referir às linhas da matriz associada como ℓ_1 e ℓ_2. Deste modo, "adicionar um múltiplo da linha 1 na linha dois" corresponde a fazer:

$$\left[\begin{array}{cc|c} 1 & 3 & 1 \\ 2 & -1 & -2 \end{array}\right] \xrightarrow{-2\ell_1+\ell_2 \text{ em } \ell_2} \left[\begin{array}{cc|c} 1 & 3 & 1 \\ 0 & -7 & -4 \end{array}\right].$$

Para este tipo de operação, olhamos para as colunas da matriz associada, fizemos mentalmente (ou em uma folha de rascunho) os cálculos envolvidos e preenchemos os valores na respectiva linha da nova matriz que está à direita:

$$(-2) \cdot 1 + 2 = 0, \qquad (-2) \cdot 3 + (-1) = -7, \qquad (-2) \cdot 1 + (-2) = -4.$$

Exemplo 4. Consideramos agora um sistema com três equações e três variáveis;

$$\begin{cases} x + 2y + z = 12 \\ x - 3y + 5z = 1 \\ 2x - y + 3z = 10 \end{cases}.$$

Vamos diretamente escrever a matriz aumentada associada ao sistema, que é a matriz com os coeficientes da variável x na primeira coluna, os coeficientes da variável y na segunda coluna, os coeficientes de z na terceira coluna e os coeficientes independentes na última coluna:

$$\left[\begin{array}{ccc|c} 1 & 2 & 1 & 12 \\ 1 & -3 & 5 & 1 \\ 2 & -1 & 3 & 10 \end{array}\right].$$

Em seguida, utilizamos operações elementares nas linhas como no exemplo anterior. Ao substituir $-\ell_1 + \ell_2$ em ℓ_2, obtemos

$$\left[\begin{array}{ccc|c} 1 & 2 & 1 & 12 \\ 0 & -5 & 4 & -11 \\ 2 & -1 & 3 & 10 \end{array}\right].$$

Licença CC-BY-SA-3.0. Contato: livroscolaborativos@gmail.com

Ou seja, eliminamos o valor 1 da primeira posição da segunda linha (em outras palavras, eliminamos a variável x da segunda equação). Prosseguindo com a eliminação dos valores abaixo da diagonal principal da matriz, obtemos:

$$\left[\begin{array}{ccc|c} 1 & 2 & 1 & 12 \\ 0 & -5 & 4 & -11 \\ 2 & -1 & 3 & 10 \end{array}\right] \xrightarrow{-2\ell_1+\ell_3 \text{ em } \ell_3} \left[\begin{array}{ccc|c} 1 & 2 & 1 & 12 \\ 0 & -5 & 4 & -11 \\ 0 & -5 & 1 & -14 \end{array}\right] \xrightarrow{-\ell_2+\ell_3 \text{ em } \ell_3} \left[\begin{array}{ccc|c} 1 & 2 & 1 & 12 \\ 0 & -5 & 4 & -11 \\ 0 & 0 & -3 & -3 \end{array}\right].$$

Voltando à notação original do sistema, obtemos

$$\left\{\begin{array}{llll} x & +2y & +z & = 12 \\ & -5y & +4z & = -11 \\ & & -3z & = -3 \end{array}\right.$$

Por causa do seu formato, este sistema é dito estar em **forma triangular** ou ainda que o sistema é **triangular**.

Como já tínhamos percebido no nosso primeiro exemplo, sistemas triangulares são fáceis de resolver. Pela última equação concluimos que $z = 1$. Subsituindo este valor na segunda equação, obtemos

$$-5y + 4 = -11 \implies y = 3.$$

Finalmente, substituindo estes dois valores na primeira equação, obtemos

$$x + 2 \cdot 3 + 1 = 12 \implies x = 5.$$

Portanto, a solução para o sistema é $(x, y, z) = (5, 3, 1)$.$\triangleleft$

Esta técnica de resolver sistemas lineares é importante por ser aplicável a sistemas de qualquer ordem e é conhecida como **escalonamento** ou **eliminação Gaussiana**.

1.2 Formas escalonadas e formas escalonadas reduzidas

Vimos dois exemplos de como resolver sistemas lineares por escalonamento. Vamos agora introduzir uma terminologia, esperando tornar o método mais sistemático.

Dizemos que uma matriz está em **forma escalonada** quando

1. As linhas que contém apenas zeros estão abaixo das demais.

2. O primeiro elemento não nulo de uma linha, conhecido como **elemento líder**, está em uma coluna à direita do elemento líder da linha acima.

Estas duas condições *implicam* que os elementos que estão abaixo de elementos líder são todos iguais a zero.

Por exemplo,

$$\begin{bmatrix} 1 & a_0 & a_1 & a_2 & a_3 \\ 0 & 0 & 2 & a_4 & a_5 \\ 0 & 0 & 0 & -1 & a_6 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{1.3}$$

é uma matriz em forma escalonada. Notamos que os coeficientes a_1, a_2, a_3, a_4, a_5 e a_6 podem ser quaisquer números sem alterar o fato de a matriz estar em forma escalonada.

Outros exemplos, retirados livro do David C. Lay:

$$\begin{bmatrix} \blacksquare & a_0 & a_1 & a_2 \\ 0 & \blacksquare & a_3 & a_4 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & \blacksquare & a_1 & a_2 & a_3 & a_4 & a_5 & a_6 & a_7 & a_8 \\ 0 & 0 & 0 & \blacksquare & a_9 & a_{10} & a_{11} & a_{12} & a_{13} & a_{14} \\ 0 & 0 & 0 & 0 & \blacksquare & a_{15} & a_{16} & a_{17} & a_{18} & a_{19} \\ 0 & 0 & 0 & 0 & 0 & \blacksquare & a_{20} & a_{21} & a_{22} & a_{23} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \blacksquare & a_{24} \end{bmatrix}.$$

Quando uma matriz está em forma escalonada, as posições marcadas com $\blacksquare$ são chamadas de **posições de pivô**. Observe que, caso a matriz não esteja em forma escalonada, o elemento líder de uma linha pode não estar na posição de pivô. Dizemos também que uma coluna é

Licença CC-BY-SA-3.0. Contato: livroscolaborativos@gmail.com

uma **coluna pivô** quando a coluna possui uma posição de pivô. Por exemplo, na primeira matriz acima, as duas primeiras colunas são colunas pivô enquanto a terceira e a quarta não são.

Uma matriz estará em **forma escalonada reduzida** quando:

1. Está na forma escalonada;

2. Todos os elementos líder são iguais a 1 e são os únicos elementos não nulos das suas colunas.

Por exemplo,

$$\begin{bmatrix} 1 & 0 & b_1 & 0 & b_2 \\ 0 & 1 & b_3 & 0 & b_3 \\ 0 & 0 & 0 & 1 & b_4 \end{bmatrix}$$

é uma forma escalonada reduzida da matriz da fórmula (1.3) acima.

Quanto aos outros exemplos, temos

$$\begin{bmatrix} 1 & 0 & b_1 & b_2 \\ 0 & 1 & b_3 & b_4 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 & b_1 & 0 & 0 & 0 & b_5 & b_6 & 0 & b_8 \\ 0 & 0 & 0 & 1 & 0 & 0 & b_{11} & b_{12} & 0 & b_{14} \\ 0 & 0 & 0 & 0 & 1 & 0 & b_{16} & b_{17} & 0 & b_{19} \\ 0 & 0 & 0 & 0 & 0 & 1 & b_{20} & b_{21} & 0 & b_{23} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & b_{24} \end{bmatrix}.$$

Observação 5. A forma escalonada reduzida ainda não apareceu na seção anterior quando estávamos resolvendo sistemas, mas notamos que poderia ter sido utilizada para automatizar ainda mais o processo de resolução do sistema. Vamos retomar o Exemplo 4. Fazendo mais algumas operações elementares, é possível transformar a matriz aumentada do sistema em uma matriz em forma escalonada reduzida:

$$\left[\begin{array}{ccc|c} 1 & 2 & 1 & 12 \\ 0 & -5 & 4 & -11 \\ 0 & 0 & -3 & -3 \end{array}\right] \xrightarrow{\ell_2 \div (-5) \text{ e } \ell_3 \div (-3)} \left[\begin{array}{ccc|c} 1 & 2 & 1 & 12 \\ 0 & 1 & -4/5 & 11/5 \\ 0 & 0 & 1 & 1 \end{array}\right]$$

$$\xrightarrow{\frac{4}{5}\ell_3 + \ell_2 \text{ em } \ell_2} \left[\begin{array}{ccc|c} 1 & 2 & 1 & 12 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & 1 \end{array}\right] \xrightarrow{-\ell_3 + \ell_1 \text{ em } \ell_1} \left[\begin{array}{ccc|c} 1 & 2 & 0 & 11 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & 1 \end{array}\right] \xrightarrow{-2\ell_2 + \ell_1 \text{ em } \ell_1} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 5 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & 1 \end{array}\right].$$

A última matriz acima está em forma escalonada reduzida e está associada ao sistema

$$\begin{cases} x & & & = 5 \\ & y & & = 3 \\ & & z & = 1 \end{cases}$$

que, na verdade, já é a solução explícita do sistema original.◁

1.3 Algoritmo de escalonamento

Sistematicamente, encontramos a forma escalonada de uma matriz aplicando os seguintes passos:

1. Analisar a primeira coluna pivô – da esquerda para a direita, esta é a primeira coluna que possui algum elemento diferente de zero – e, se necessário, aplicar operações elementares para que o elemento da primeira linha (esta é a posição de pivô!) passe a ser diferente de zero;

2. A partir de operações elementares, eliminar todos elementos que estão abaixo do elemento na posição de pivô que obtivemos no Passo 1;

3. Desconsiderando por um momento a primeira linha, procuramos pela próxima coluna pivô – aquela que tem algum elemento não nulo *abaixo da primeira linha*. Se necessário, aplicar operações elementares para que, na segunda linha, o elemento passe a ser diferente de zero;

Licença CC-BY-SA-3.0. Contato: livroscolaborativos@gmail.com

4. A partir de operações elementares, eliminar todos elementos que estão abaixo do elemento na posição de pivô que obtivemos no Passo 3;

5. Desconsiderando por um momento a primeira e a segunda linhas, procuramos pela próxima coluna pivô – aquela que tem algum elemento não nulo *abaixo da primeira linha e da segunda linha*. Se necessário, aplicar operações elementares para que, na segunda linha, o elemento passe a ser diferente de zero;

6. A partir de operações elementares, eliminar todos elementos que estão abaixo do elemento na posição de pivô que obtivemos no Passo 5 e ssim por diante.

Essencialmente, são apenas dois passos que se repetem até que se esgotem as colunas que possuem posição de pivô. Vejamos um exemplo.

Exemplo 6. Considere a matriz

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 1 & -1 & 8 \\ 0 & 0 & -3 & -1 & 2 & -11 \\ 0 & 0 & 6 & 2 & -4 & 22 \\ 0 & 0 & -2 & 1 & 2 & -3 \end{bmatrix}.$$

Passo 1. A primeira coluna pivô é a terceira. Escolhemos um elemento não nulo da terceira coluna para ocupar a posição de pivô. Por exemplo, a segunda linha. Trocando a primeira e a segunda linhas de lugar (esta é uma operação elementar), obtemos:

$$\begin{bmatrix} 0 & 0 & 2 & 1 & -1 & 8 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -3 & -1 & 2 & -11 \\ 0 & 0 & 6 & 2 & -4 & 22 \\ 0 & 0 & -2 & 1 & 2 & -3 \end{bmatrix}.$$

Passo 2. Eliminar os elementos abaixo do 2 que está na posição de pivô da terceira coluna:

$$\begin{bmatrix} 0 & 0 & 2 & 1 & -1 & 8 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -3 & -1 & 2 & -11 \\ 0 & 0 & 6 & 2 & -4 & 22 \\ 0 & 0 & -2 & 1 & 2 & -3 \end{bmatrix} \xrightarrow{\frac{3}{2}\ell_1+\ell_3 \text{ em } \ell_3} \begin{bmatrix} 0 & 0 & 2 & 1 & -1 & 8 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1/2 & 1/2 & 1 \\ 0 & 0 & 6 & 2 & -4 & 22 \\ 0 & 0 & -2 & 1 & 2 & -3 \end{bmatrix}$$

$$\xrightarrow{-3\ell_1+\ell_4 \text{ em } \ell_4} \begin{bmatrix} 0 & 0 & 2 & 1 & -1 & 8 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1/2 & 1/2 & 1 \\ 0 & 0 & 0 & -1 & -1 & -2 \\ 0 & 0 & -2 & 1 & 2 & -3 \end{bmatrix} \xrightarrow{\ell_1+\ell_5 \text{ em } \ell_5} \begin{bmatrix} 0 & 0 & 2 & 1 & -1 & 8 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1/2 & 1/2 & 1 \\ 0 & 0 & 0 & -1 & -1 & -2 \\ 0 & 0 & 0 & 2 & 1 & 5 \end{bmatrix}.$$

Passo 3. A partir de agora, vamos essencialmente repetir o processo já realizado até agora. Notamos que a próxima coluna que contém elementos não nulos (sem olhar para a primeira linha!) é a quarta coluna. Portanto, esta é uma coluna pivô e vamos posicionar um elemento não nulo na segunda linha. Por exemplo, podemos trocar a segunda e a terceira linhas.

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

Antes de prosseguir, podemos ainda simplificar alguns cálculos multiplicando linhas por escalares (fazer isso é realizar um operação elementar!):

$$\begin{bmatrix} 0 & 0 & 2 & 1 & -1 & 8 \\ 0 & 0 & 0 & 1/2 & 1/2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & -1 & -2 \\ 0 & 0 & 0 & 2 & 1 & 5 \end{bmatrix} \xrightarrow{2\cdot\ell_3 \text{ e } (-1)\cdot\ell_4} \begin{bmatrix} 0 & 0 & 2 & 1 & -1 & 8 \\ 0 & 0 & 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 2 & 1 & 5 \end{bmatrix}$$

Licença CC-BY-SA-3.0. Contato: livroscolaborativos@gmail.com

Passo 4. Prosseguimos como no Passo 2 para eliminar todos os elementos que estão abaixo da posição de pivô da quarta coluna:

$$\begin{bmatrix} 0 & 0 & 2 & 1 & -1 & 8 \\ 0 & 0 & 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 2 & 1 & 5 \end{bmatrix} \xrightarrow{-\ell_2+\ell_4 \text{ em } \ell_4} \begin{bmatrix} 0 & 0 & 2 & 1 & -1 & 8 \\ 0 & 0 & 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 1 & 5 \end{bmatrix}$$

$$\xrightarrow{-2\ell_2+\ell_5 \text{ em } \ell_5} \begin{bmatrix} 0 & 0 & 2 & 1 & -1 & 8 \\ 0 & 0 & 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 1 \end{bmatrix}$$

Passo 5. Finalmente, identificamos a coluna 5 como coluna pivô e obtemos uma matriz em forma escalonada ao deixar um elemento não nulo na posição de pivô:

$$\begin{bmatrix} 0 & 0 & 2 & 1 & -1 & 8 \\ 0 & 0 & 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} . \triangleleft$$

Agora, para obter a forma escalonada reduzida *a partir da forma escalonada*, primeiramente aplicamos os passos acima para obter uma matriz em forma escalonada. Em seguida, aplicamos outros passos iterativos:

1. Começar pela posição de pivô mais à *direita* e eliminar os elementos não nulos *acima* da posição de pivô;
2. Se necessário, dividir a linha pelo valor do elemento líder (que está na posição de pivô) para que o elemento líder fique igual a 1;
3. Repetir os primeiros passos para o elemento líder na próxima (da direita para a esquerda) coluna pivô.

Observamos que poderíamos ter primeiro realizado o Passo 2 e depois o Passo 1 se julgássemos que seria mais simples para efetuar os cálculos.

Exemplo 7. Voltamos ao Exemplo 6. Identificando novamente as posições de pivô:

$$\begin{bmatrix} 0 & 0 & \mathbf{2} & 1 & -1 & 8 \\ 0 & 0 & 0 & \mathbf{1} & 1 & 2 \\ 0 & 0 & 0 & 0 & \mathbf{-1} & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} .$$

O mais da direita é o -1 da quinta coluna. Eliminando os termos não nulos acima, obtemos:

$$\begin{bmatrix} 0 & 0 & 2 & 1 & -1 & 8 \\ 0 & 0 & 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \xrightarrow{\ell_3+\ell_2 \text{ em } \ell_2} \begin{bmatrix} 0 & 0 & 2 & 1 & -1 & 8 \\ 0 & 0 & 0 & 1 & 0 & 3 \\ 0 & 0 & 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \xrightarrow{-\ell_3+\ell_1 \text{ em } \ell_1} \begin{bmatrix} 0 & 0 & 2 & 1 & 0 & 7 \\ 0 & 0 & 0 & 1 & 0 & 3 \\ 0 & 0 & 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} .$$

A próxima posição de pivô mais à direita é o 1 na linha 2, coluna 4. Eliminando o termo não nulo acima, obtemos:

$$\xrightarrow{-\ell_2+\ell_1 \text{ em } \ell_1} \begin{bmatrix} 0 & 0 & 2 & 0 & 0 & 4 \\ 0 & 0 & 0 & 1 & 0 & 3 \\ 0 & 0 & 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} .$$

Licença CC-BY-SA-3.0. Contato: livroscolaborativos@gmail.com

Finalmente, dividimos a primeira linha por 2 e a terceira linha por -1 para chegar na forma escalonada reduzida da matriz inicial:

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

1.4 Existência e unicidade de solução

1.4.1 Sistemas que possuem apenas uma solução

Os sistemas que vimos até agora possuiam apenas uma solução. Esta propriedade, como veremos, nem sempre é válida. No entanto, é fácil de identificar quando um sistema possui solução única analisando a forma escalonada da matriz associada: **quando todas as colunas referentes às variáveis da equação possuirem posição de pivô**. Por exemplo,

$$\left[\begin{array}{cc|c} 1 & 3 & 1 \\ 0 & -7 & -4 \end{array}\right], \left[\begin{array}{ccc|c} 1 & 2 & 1 & 12 \\ 0 & -5 & 4 & -11 \\ 0 & 0 & -3 & -3 \end{array}\right], \left[\begin{array}{cccc|c} 2 & 0 & 0 & 0 & 10 \\ 0 & 1 & 0 & 0 & -12 \\ 0 & 0 & 2 & 0 & 6 \\ 0 & 0 & 0 & 1 & 1 \end{array}\right].$$

De fato, todas as matrizes com esta propriedade tem em sua forma escalonada reduzida a matriz identidade do lado esquerdo da barra que separa as variáveis e os valores do lado direito da igualdade do sistema:

$$\left[\begin{array}{cc|c} 1 & 0 & -5/7 \\ 0 & 1 & 4/7 \end{array}\right], \left[\begin{array}{ccc|c} 1 & 0 & 0 & 5 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & 1 \end{array}\right], \left[\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 5 \\ 0 & 1 & 0 & 0 & -12 \\ 0 & 0 & 1 & 0 & 3 \\ 0 & 0 & 0 & 1 & 1 \end{array}\right].$$

Escrevendo o sistema associado as estas últimas matrizes, é fácil de ver que a solução será única.

1.4.2 Sistemas impossíveis ou inconsistentes

Nem todo sistema linear possui solução. Um exemplo simples deste fato é

$$\begin{cases} x + y = 1 \\ x + y = 2 \end{cases}. \tag{1.4}$$

É claro que $x + y$ não pode ser ao mesmo tempo igual a 1 e igual a 2; portanto, o sistema não possui solução. O fato de o sistema não possuir solução – ser **inconsistente** – nem sempre é tão fácil de ser identificado apenas olhando para o sistema. Mas esta propriedade salta aos olhos quando analisamos a forma escalonada da matriz associada ao sistema. De fato, na matriz escalonada de um sistema inconsistente aparece uma linha da forma

$$\left[\begin{array}{ccccc|c} 0 & 0 & 0 & 0 & 0 & a \end{array}\right],$$

com uma constante $a \neq 0$. Esta linha representa no sistema uma equação do tipo $0 = a \neq 0$, que é impossível de ser satisfeita.

No sistema da fórmula (1.4) acima, temos a matriz aumentada

$$\left[\begin{array}{cc|c} 1 & 1 & 1 \\ 1 & 1 & 2 \end{array}\right] \xrightarrow{-\ell_1+\ell_2 \text{ em } \ell_2} \left[\begin{array}{cc|c} 1 & 1 & 1 \\ 0 & 0 & 1 \end{array}\right].$$

Logo, nosso sistema é equivalente ao sistema impossível

$$\begin{cases} x + y & = 1 \\ \quad 0 & = 1 \end{cases}.$$

Licença CC-BY-SA-3.0. Contato: livroscolaborativos@gmail.com

1.4.3 Sistemas com mais de uma solução

Com exceção da subseção anterior, todos os sistemas que resolvemos anteriormente possuiam apenas uma solução. Vejamos um exemplo onde este não é o caso.

Exemplo 8. Resolver

$$\left\{\begin{array}{rrrrrcl} & 3x_2 & -6x_3 & +6x_4 & +4x_5 & = & -5 \\ 3x_1 & -7x_2 & +8x_3 & -5x_4 & +8x_5 & = & 9 \\ 3x_1 & -9x_2 & +12x_3 & -9x_4 & +6x_5 & = & 15 \end{array}\right. .$$

A matriz aumentada associada a este sistema é

$$\left[\begin{array}{ccccc|c} 0 & 3 & -6 & 6 & 4 & -5 \\ 3 & -7 & 8 & -5 & 8 & 9 \\ 3 & -9 & 12 & -9 & 6 & 15 \end{array}\right].$$

Transformando esta matriz em sua forma escalonada reduzida (faça todos os cálculos!), obtemos

$$\left[\begin{array}{ccccc|c} 1 & 0 & -2 & 3 & 0 & -24 \\ 0 & 1 & -2 & 2 & 0 & -7 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{array}\right],$$

que é equivalente ao sistema linear

$$\left\{\begin{array}{rcl} x_1 - 2x_3 + 3x_4 & = & -24 \\ x_2 - 2x_3 + 2x_4 & = & -7 \\ x_5 & = & 4 \end{array}\right.$$

Neste sistema x_5 possui um valor fixo igual a 4, mas as outras variáveis tem dois graus de liberdade. De fato, atribuindo valores quaisquer para x_3 e x_4, obtemos os valores de x_1 e x_2 pelas duas primeiras equações acima e temos uma solução do sistema. Diferentes escolhas dos parâmetros x_3 e x_4 geram soluções diferentes.

Uma forma sistemática de analisar a situação é a seguinte: As variáveis correspondentes a colunas que possuem posição de pivô são chamadas de **variáveis dependentes** e as demais são chamadas de **variáveis livres**. Assim, podemos escrever

$$\left\{\begin{array}{l} x_1 = 2x_3 - 3x_4 - 24 \\ x_2 = 2x_3 - 2x_4 - 7 \\ x_3, x_4 \text{ livres} \\ x_5 = 4 \end{array}\right.$$

e o sistema possui infinitas soluções.$\triangleleft$

1.4.4 Conclusões

Ao analisar a matriz aumentada associada a um sistema linear, concluiremos *inevitavelmente* que uma das seguintes situações é válida:

- Todas as colunas referentes às variáveis do sistema possuem posição de pivô, *e.g*,

  $$\left[\begin{array}{ccc|c} 1 & 3 & -6 & 6 \\ 0 & -8 & 11 & \pi \\ 0 & 0 & 12 & -9 \end{array}\right].$$

  Neste caso, vimos que o sistema possui **apenas uma solução**. Nota que *esta possibilidade apenas é possível se tivermos um sistema com o número de equações igual ao número de variáveis.*

- Caso a coluna referentes à alguma variável do sistema não possua posição de pivô, teremos duas possibilidades

(i) Alguma coluna não é coluna pivô mas não existem linhas inconsistentes, *e.g*,

$$\left[\begin{array}{ccccc|c} 1 & 3 & 2 & -6 & 2 & 6 \\ 0 & -8 & 4 & 11 & -1 & -3 \\ 0 & 0 & 0 & 12 & 1 & -9 \end{array}\right].$$

Neste caso, as variáveis referentes a colunas que não são pivô (na matriz acima, colunas 3 e 5) podem ser consideradas variáveis livres e o sistema **possui infinitas soluções**.

(ii) Alguma coluna não é coluna pivô mas existem linhas que são inconsistentes, *e.g*,

$$\left[\begin{array}{ccccc|c} 1 & 3 & 2 & -6 & 2 & 6 \\ 0 & -8 & 4 & 11 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 & -9 \end{array}\right].$$

Neste caso, **não existem soluções** para o sistema, pois uma das equações é impossível de ser satisfeita.

Concluimos assim que *um sistema linear ou não possui solução, ou tem apenas uma solução ou então possui infinitas soluções*. São estes três todos os casos possíveis e podemos decidir em qual estamos apenas olhando para a forma escalonada da matriz aumentada associada ao sistema.

Licença CC-BY-SA-3.0. Contato: livroscolaborativos@gmail.com

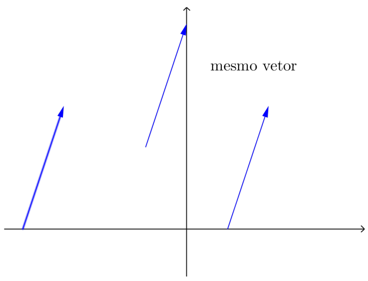
2

Semana 2

Existem maneiras diferentes de representar e de interpretar elementos de Álgebra Linear. Nesta parte, vemos algumas das principais.

2.1 Vetores

Um **vetor** no espaço tridimensional é um objeto com magnitude, direção e sentido bem definidos. Assim sendo, ao representar um vetor como na figura abaixo, não importa onde está o seu ponto inicial. Todas as setas representam o mesmo vetor.



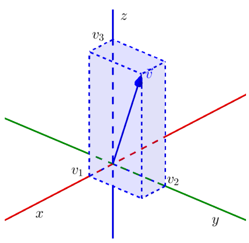
Fixando o sistema de coordenadas Cartesiano usual e usando a origem como referência, podemos introduzir coordenadas:

$$\vec{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = v_1\vec{e_1} + v_2\vec{e_2} + v_3\vec{e_3},$$

onde

$$\vec{e_1} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \vec{e_2} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \quad \text{e} \quad \vec{e_3} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

Intuitivamente, podemos pensar que o sistema de coordenadas Cartesiano fixa a origem como sendo o ponto inicial do vetor e o ponto (v_1, v_2, v_3) do espaço tridimensional como o ponto final.



No livro “Álgebra Linear e suas Aplicações”, de David Lay, vetores são representados por letras em negrito. Nestas notas, utilizamos uma setinha em cima da letra para indicar um vetor.

Raramente vamos confundir o ponto (v_1, v_2, v_3) com o vetor

$$\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}.$$

De qualquer maneira, é bom estar atento, pois outros textos podem fazê-lo.

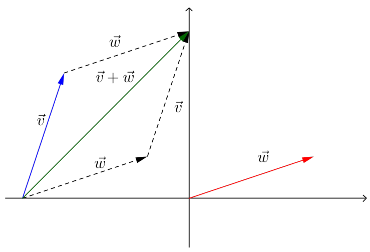
O conjunto de todos os vetores com três componentes como acima é chamado de $\mathbb{R}^3$ (leia-se “erre três”). Em coordenadas, se tivermos

$$\vec{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \quad \text{e} \quad \vec{w} = \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix},$$

podemos definir a **soma de vetores** componente a componente:

$$\vec{v} + \vec{w} := \begin{bmatrix} v_1 + w_1 \\ v_2 + w_2 \\ v_3 + w_3 \end{bmatrix}.$$

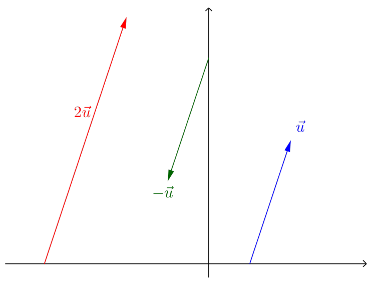
Geometricamente, isto pode ser interpretado como segue:



A **multiplicação** do vetor $\vec{v}$ **por escalar** $k \in \mathbb{R}$ é definida como:

$$k\vec{v} := \begin{bmatrix} kv_1 \\ kv_2 \\ kv_3 \end{bmatrix}.$$

Geometricamente,



Estas considerações que fizemos também são válidas para outro número de componentes, com a possível perda de visualização geométrica, no caso de quatro ou mais componentes.

Mais geralmente, um vetor $\vec{v}$ do conjunto $\mathbb{R}^n$ pode ser pensado como um objeto com n componentes:

$$\vec{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_{n-1} \\ v_n \end{bmatrix} = v_1\vec{e}_1 + v_2\vec{e}_2 + \cdots + v_{n-1}\vec{e}_{n-1} + v_n\vec{e}_n,$$

Licença CC-BY-SA-3.0. Contato: livroscolaborativos@gmail.com

onde

$$\vec{e}_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}, \quad \vec{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \cdots, \quad \vec{e}_{n-1} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ 0 \end{bmatrix}, \quad \vec{e}_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}.$$

A soma de vetores é dada em coordenadas por:

$$\vec{v} + \vec{w} := \begin{bmatrix} v_1 + w_1 \\ v_2 + w_2 \\ \vdots \\ v_{n-1} + w_{n-1} \\ v_n + w_n \end{bmatrix}.$$

A multiplicação por escalar por:

$$k\vec{v} := \begin{bmatrix} kv_1 \\ kv_2 \\ \vdots \\ kv_{n-1} \\ kv_n \end{bmatrix}.$$

2.1.1 Representação matricial para sistemas Lineares

Na linguagem da seção anterior, nós dizemos que dois vetores são iguais quando todas as suas componentes são iguais. Desta forma, podemos interpretar as equações de um sistema linear

$$\begin{cases} x + 3y &= 1 \\ 2x - y &= -2 \end{cases}$$

como uma igualdade entre vetores de $\mathbb{R}^2$, isto é, de vetores com duas componentes:

$$\begin{bmatrix} x + 3y \\ 2x - y \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}.$$

Além disso, é desta forma que surge naturalmente o **produto de uma matriz por um vetor**:

$$\begin{bmatrix} 1 & 3 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \end{bmatrix},$$

definido de modo que as duas últimas igualdades acima signifiquem a mesma coisa. Esta última notação é conhecida como a **forma matricial** do sistema linear.

Exemplo 9. O sistema, que já apareceu nas notas da semana anterior

$$\begin{cases} x_1 + 2x_2 + x_3 = 12 \\ x_1 - 3x_2 + 5x_3 = 1 \\ 2x_1 - x_2 + 3x_3 = 10 \end{cases}$$

pode ser representado matricialmente por

$$\begin{bmatrix} 1 & 2 & 1 \\ 1 & -3 & 5 \\ 2 & -1 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 12 \\ 1 \\ 10 \end{bmatrix}.$$

De forma mais sucinta,

$$\boxed{A\vec{x} = \vec{b}}$$

onde

$$A = \begin{bmatrix} 1 & 2 & 1 \\ 1 & -3 & 5 \\ 2 & -1 & 3 \end{bmatrix}, \quad \vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \quad \text{e} \quad \vec{b} = \begin{bmatrix} 12 \\ 1 \\ 10 \end{bmatrix}.$$

Licença CC-BY-SA-3.0. Contato: livroscolaborativos@gmail.com

Chegamos a esta conclusão colocando os coeficientes da primeira variável x_1 na primeira coluna, os coeficientes da segunda variável x_2 na segunda coluna e os coeficientes da terceira variável x_3 na terceira coluna. $\triangleleft$

Mais geralmente, uma matriz do tipo $m \times n$ (leia-se "m por n") é uma matriz com m linhas e n colunas:

$$A = (a_{ij}) = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

e pode estar associada a um sistema com m equações e n variáveis, já que o produto

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix},$$

quando visto componente a componente, representa:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \qquad\qquad\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases}$$

2.2 Combinações lineares de vetores

Uma **combinação linear** dos vetores $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k$ é um vetor da forma:

$$\vec{v} = x_1\vec{v}_1 + x_2\vec{v}_2 + \cdots + x_k\vec{v}_k = \sum_{i=1}^{k} x_i\vec{v}_i,$$

onde $x_1, x_2, \ldots, x_k$ são números reais. Em outras palavras, uma combinação linear é uma soma de múltiplos dos vetores $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k$.

Exemplo 10. Para os vetores

$$\vec{v}_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \quad \text{e} \quad \vec{v}_2 = \begin{bmatrix} 1 \\ 3 \end{bmatrix},$$

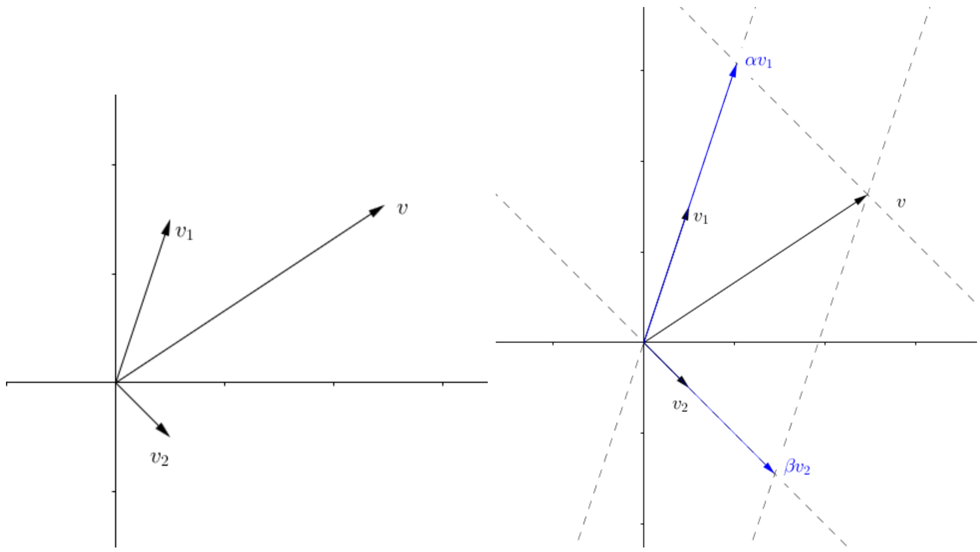
algumas combinações lineares são:

- $\vec{v}_1 + \vec{v}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix} + \begin{bmatrix} 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$;
- $4\vec{v}_1 = 4\vec{v}_1 + 0\vec{v}_2 = 4\begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 4 \\ -4 \end{bmatrix}$;
- $\vec{v}_2 = 0\vec{v}_1 + 1\vec{v}_2 = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$;
- $0\vec{v}_1 + 0\vec{v}_2 = \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \vec{0}$.

Geometricamente, qualquer vetor do plano pode ser representado como combinação linear de vetores que não são colineares.

Vamos ilustrar este fato na figura abaixo. Por exemplo, para representar o vetor $\vec{v}$ como combinação linear de $\vec{v}_1$ e $\vec{v}_2$, podemos traçar retas paralelas aos vetores, passando pela origem e pela extremidade do vetor $\vec{v}$, como na figura da direita. Assim, pela interpretação geométrica da soma de vetores (ver início deste capítulo), sabemos que $\vec{v}$ é a soma dos vetores em azul. Mas

Licença CC-BY-SA-3.0. Contato: livroscolaborativos@gmail.com



estes, por construção, são colineares aos vetores iniciais e logo múltiplos destes. Concluimos que

$$\vec{v} = \alpha\vec{v}_1 + \beta\vec{v}_2,$$

isto é, $\vec{v}$ é uma combinação linear de $\vec{v}_1$ e $\vec{v}_2$. $\lhd$

De forma mais geral, nós dizemos que um vetor $\vec{v} \in \mathbb{R}^m$ é combinação linear dos k vetores $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k \in \mathbb{R}^m$ quando conseguirmos encontrar números reais $x_1, x_2, \ldots, x_k$ tais que

$$x_1\vec{v}_1 + x_2\vec{v}_2 + \cdots + x_k\vec{v}_k = \vec{v}.$$

Nós vamos nos referir a este tipo de equação por **equação vetorial**. Para decidir se um vetor é combinação linear de outros, devemos decidir se existem estes números reais $x_1, x_2, \ldots, x_k$. Escrevendo em coordenadas:

$$\vec{v}_1 = \begin{bmatrix} v_{11} \\ v_{21} \\ v_{31} \\ \vdots \\ v_{m1} \end{bmatrix}, \ \vec{v}_2 = \begin{bmatrix} v_{12} \\ v_{22} \\ v_{32} \\ \vdots \\ v_{m2} \end{bmatrix}, \ \vec{v}_3 = \begin{bmatrix} v_{13} \\ v_{23} \\ v_{33} \\ \vdots \\ v_{m3} \end{bmatrix}, \ \cdots, \ \vec{v}_k = \begin{bmatrix} v_{1k} \\ v_{2k} \\ v_{3k} \\ \vdots \\ v_{mk} \end{bmatrix}, \ \vec{v} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_m \end{bmatrix},$$

vemos que encontrar os coeficientes da combinação linear, caso estes existam, equivale a resolver a equação vetorial

$$x_1 \begin{bmatrix} v_{11} \\ v_{21} \\ v_{31} \\ \vdots \\ v_{m1} \end{bmatrix} + x_2 \begin{bmatrix} v_{12} \\ v_{22} \\ v_{32} \\ \vdots \\ v_{m2} \end{bmatrix} + x_3 \begin{bmatrix} v_{13} \\ v_{23} \\ v_{33} \\ \vdots \\ v_{m3} \end{bmatrix} + \cdots + x_k \begin{bmatrix} v_{1k} \\ v_{2k} \\ v_{3k} \\ \vdots \\ v_{mk} \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_m \end{bmatrix},$$

que, por sua vez, é equivalente a

$$\begin{bmatrix} v_{11}x_1 + v_{12}x_2 + v_{13}x_3 + \cdots + v_{1k}x_k \\ v_{21}x_1 + v_{22}x_2 + v_{23}x_3 + \cdots + v_{2k}x_k \\ v_{31}x_1 + v_{32}x_2 + v_{33}x_3 + \cdots + v_{3k}x_k \\ \vdots \\ v_{m1}x_1 + v_{m2}x_2 + v_{m3}x_3 + \cdots + v_{mk}x_k \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_m \end{bmatrix}.$$

Mas isto é resolver um sistema linear! Portanto este conceito (aparentemente) novo está intimamente relacionado com tudo o que já tínhamos visto antes.

O **espaço gerado** por todas as combinações lineares dos vetores $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k$ é denotado por $\mathrm{Span}\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k\}$.

Licença CC-BY-SA-3.0. Contato: livroscolaborativos@gmail.com

- Vimos que o conjunto gerado pelos dois vetores do exemplo anterior é o plano inteiro, isto é, $\mathrm{Span}\{\vec{v}_1, \vec{v}_2\} = \mathbb{R}^2$.
- Nota que o vetor nulo sempre pertence ao espaço gerado, assim como todos os vetores $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k$.
- O conjunto gerado por apenas um vetor não nulo representa uma reta.

2.3 Sistemas lineares e combinações lineares das colunas

A associação feita acima pode ser resumida como:

”resolver o sistema linear $A\vec{x} = \vec{b}$ equivale a decidir se o vetor $\vec{b}$ é uma combinação linear das colunas de A”

ou ainda

”resolver o sistema linear $A\vec{x} = \vec{b}$ equivale a decidir se o vetor $\vec{b}$ pertence ao espaço gerado pelas colunas de A”

Desta forma, tudo o que já vimos sobre existência de soluções para sistemas lineares pode ser “traduzido” para o contexto de combinações lineares e espaços gerados (faça as traduções!).

Exemplo 11. Considere o sistema linear do Exemplo 9:

$$\begin{cases} x_1 + 2x_2 + x_3 = 12 \\ x_1 - 3x_2 + 5x_3 = 1 \\ 2x_1 - x_2 + 3x_3 = 10 \end{cases}$$

(i) Vimos ques este sistema pode ser representado em **forma matricial** por

$$\begin{bmatrix} 1 & 2 & 1 \\ 1 & -3 & 5 \\ 2 & -1 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 12 \\ 1 \\ 10 \end{bmatrix}.$$

Este formato é interessante principalmente para uma análise mais teórica. Podemos escalonar a matriz para decidir sobre a existência de soluções. Podemos também encarar o sistema como uma equação *aparentemente* escalar $A\vec{x} = \vec{b}$; por exemplo, mais adiante no nosso curso, vamos tentar entender quando que pode fazer sentido escrever a solução como $\vec{x} = A^{-1}\vec{b}$; isto é o que se faria em uma equação linear escalar (adiantando: nem sempre se pode fazer isso!).

(ii) Também podemos representar o sistema como uma **matriz aumentada associada**:

$$\left[\begin{array}{ccc|c} 1 & 2 & 1 & 12 \\ 1 & -3 & 5 & 1 \\ 2 & -1 & 3 & 10 \end{array}\right].$$

Esta notação é boa para a resolução do sistema. Escalonando a matriz aumentada associada, conseguiremos dizer se o sistema possui solução (no caso em que uma linha da forma escalonada for contraditória), se o sistema possui apenas uma solução (quando todas as colunas referentes às variáveis do sistema possuem posição de pivô) ou então se existem infinitas soluções (caso sem linhas contraditórias em que existem colunas sem posição de pivô – assocaidas com variáveis livres).

(iii) Outra forma possível é a **forma vetorial** (equivalentemente **combinação linear das colunas** da matriz associada)

$$x_1 \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} + x_2 \begin{bmatrix} 2 \\ -3 \\ -1 \end{bmatrix} + x_3 \begin{bmatrix} 1 \\ 5 \\ 3 \end{bmatrix} = \begin{bmatrix} 12 \\ 1 \\ 10 \end{bmatrix}.$$

Esta maneira de representar o sistema é favorável para descrições geométricas, pois já vimos como interpretar geometricamente combinações lineares de vetores. Também será interessante futuramente no curso quando a imagem de transformações lineares.

Licença CC-BY-SA-3.0. Contato: livroscolaborativos@gmail.com

O entendimento dessas interpretações diferentes facilita (e muito!) o entendimento de vários conceitos que veremos mais adiante no nosso curso de Álgebra Linear.

Licença CC-BY-SA-3.0. Contato: livroscolaborativos@gmail.com

3

Semana 3

3.1 Dependência e independência linear

Nós vimos, por definição, que um vetor $\vec{v} \in \mathbb{R}^m$ é combinação linear dos k vetores $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k \in \mathbb{R}^m$ quando conseguirmos encontrar números reais $x_1, x_2, \ldots, x_k$ tais que

$$x_1\vec{v}_1 + x_2\vec{v}_2 + \cdots + x_k\vec{v}_k = \vec{v}.$$

Vimos também que para decidir se um vetor é combinação linear de outros, devemos verificar se existem estes números reais $x_1, x_2, \ldots, x_k$. Em coordenadas, isto é equivalente a resolver o sistema linear:

$$\begin{bmatrix} v_{11} & v_{12} & v_{13} & \cdots & v_{1k} \\ v_{21} & v_{22} & v_{23} & \cdots & v_{2k} \\ v_{31} & v_{32} & v_{33} & \cdots & v_{3k} \\ \vdots & \vdots & \vdots & & \vdots \\ v_{m1} & v_{m2} & v_{m3} & \cdots & v_{mk} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_k \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_k \end{bmatrix}.$$

Agora, dizemos que os vetores $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k \in \mathbb{R}^m$ são **linearmente independentes (LI)** se nenhum dos vetores puder ser escrito combinação linear dos demais. Uma forma de escrever isto matematicamente é a seguinte:

$$\boxed{\text{Se } x_1\vec{v}_1 + x_2\vec{v}_2 + \cdots + x_k\vec{v}_k = \vec{0}, \text{ então } x_1 = x_2 = \cdots = x_k = 0.}$$

De fato, imaginem que pelo menos um dos coeficientes acima fosse diferente de zero, digamos $x_2 \neq 0$. Daí podemos dividir por x_2 e conseguiríamos escrever o vetor $\vec{v}_2$ como combinação linear dos demais:

$$\vec{v}_2 = -\frac{x_1}{x_2}\,\vec{v}_1 - \frac{x_3}{x_2}\,\vec{v}_3 - \cdots - \frac{x_k}{x_2}\,\vec{v}_k.$$

Se os vetores $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k \in \mathbb{R}^m$ não forem linearmente independentes, então nós dizemos que eles são **linearmente dependentes (LD)**.

Exemplo 12. Vamos verificar se os vetores

$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

são LI ou LD.

Para verificar se são LI, nós devemos considerar a equação vetorial

$$x_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \vec{0} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Como a equação é homogênea, temos pelo menos a solução trivial: $x_1 = 0$, $x_2 = 0$ e $x_3 = 0$. Se esta for a única solução, então os vetores são LI. Se existir alguma outra solução que não seja a trivial, então os vetores são LD.

Resolver a equação vetorial equivale a resolver o sistema linear

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

É fácil de ver que a única solução deste sistema é a trivial e, portanto, os vetores são LI. Se este sistema não fosse fácil de resolver, deveríamos começar a resolvê-lo por escalonamento, como de costume. $\lhd$

Exemplo 13. Analisamos agora os vetores

$$\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \quad \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}.$$

Como no exemplo anterior, para verificar se são LI, nós devemos considerar a equação vetorial

$$x_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + x_4 \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Como a equação é homogênea, temos pelo menos a solução trivial: $x_1 = 0$, $x_2 = 0$ e $x_3 = 0$. Se esta for a única solução, então os vetores são LI. Se existir alguma outra solução que não seja a trivial, então os vetores são LD.

Resolver a equação vetorial equivale a resolver o sistema linear

$$\begin{bmatrix} 1 & 1 & 1 & -1 \\ 0 & 1 & 1 & 2 \\ 1 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Em um sistema com mais variáveis do que equações, podemos não ter soluções (no caso de alguma linha inconsistente) ou podemos ter infinitas soluções (no caso de haver variáveis livres). Não é possível ter unicidade. Como já sabemos que em um sistema homogêneo, sempre temos pelo menos a solução trivial $x_1 = 0$, $x_2 = 0$, $x_3 = 0$ e $x_4 = 0$, podemos concluir que existem infinitas soluções. Logo, o conjunto de vetores é LD.

O argumento que utilizamos no Exemplo 13 acima é geral e se aplica em várias situações. Temos

$$\boxed{\text{Se } k > m, \text{ então os vetores } \vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k \in \mathbb{R}^m \text{ são necessariamente LD.}}$$

Desta forma, não existem quatro vetores LI em $\mathbb{R}^3$. Não conseguiríamos encontrar sete vetores LI em $\mathbb{R}^5$ e assim por diante.

Façamos agora algumas observações gerais sobre independência linear:

- O vetor nulo, mesmo que sozinho, é LD. Se $\vec{v} = \vec{0}$ estiver em um conjunto de vetores, este conjunto será automaticamente LD.
- Ao considerarmos apenas dois vetores $\vec{u}$ e $\vec{v}$, dizer que estes vetores são LD significa que eles são múltiplos um do outro e, portanto, colineares.
- Veremos mais adiante no curso que a dimensão do espaço gerado por vetores linearmente independentes é exatamente o número de vetores. Desta maneira, concluiremos, por exemplo, que o conjunto gerado por dois vetores do espaço tridimensional é um espaço linear de dimensão dois: um plano.

3.2 Independência linear e sistemas lineares

Na última semana (capítulo anterior), vimos que resolver o sistema linear $A\vec{x} = \vec{b}$ equivale a decidir se o vetor $\vec{b}$ é uma combinação linear das colunas de A. Na seção anterior, introduzimos a questão de decidir se os vetores $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k$ são LI ou LD. Isto, consequentemente[1], equivale a decidir se existe solução não trivial para o sistema linear homogêneo

$$A\vec{x} = \vec{0},$$

onde a matriz A é formada com os vetores $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k$ como colunas.

Exemplo 14. Naturalmente, podemos traduzir os exemplos da seção anterior desta forma. Ou seja no Exemplo 12, temos que a matriz

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}$$

tem as colunas linearmente independentes (LI). Por outro lado, as colunas da matriz

$$B = \begin{bmatrix} 1 & 1 & 1 & -1 \\ 0 & 1 & 1 & 2 \\ 1 & 0 & 1 & 1 \end{bmatrix}$$

do Exemplo 13 são linearmente dependentes (LD). $\triangleleft$

Podemos pensar ainda de outra forma. Considere uma matriz A de ordem $m \times n$. Para decidir se as colunas de A são LI, devemos procurar por soluções não triviais do sistema linear cuja matriz associada é A. Ao resolver um **sistema homogêneo** por escalonamento, sabemos que

- Se existirem mais colunas do que linhas (*i.e.* $m < n$), então as colunas de A são necessariamente LD.
- Se existirem mais linhas do que colunas (*i.e.* $m > n$), procedemos por escalonamento. Se todas as colunas da matriz A possuirem posição de pivô, então as colunas são LI (pois daí a única solução do sistema homogêneo é a trivial).
- No caso de alguma coluna não possuir posição de pivô, o sistema homogêneo possui pelo menos uma variável livre; logo, as colunas de A são LD.

Exemplo 15. Decidir se as colunas de A são LI:

$$A = \begin{bmatrix} 2 & 1 & -1 & 8 \\ -3 & -1 & 2 & -11 \\ 6 & 2 & -4 & 22 \\ -2 & 1 & 2 & -3 \end{bmatrix}.$$

Como as colunas de A são quatro vetores de $\mathbb{R}^4$, pode ser que sejam LI (caso tivesse uma coluna a mais, saberíamos automaticamente serem LD). Aplicando o algoritmo de escalonamento, sabemos que

$$A \sim \begin{bmatrix} 2 & 1 & -1 & 8 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Notando que a última coluna não possui posição de pivô, concluimos que as colunas de A são LD. Isto pois, neste caso, o sistema linear homogêneo associado, $A\vec{x} = \vec{0}$, cuja matriz aumentada associada é

$$\left[\begin{array}{cccc|c} 2 & 1 & -1 & 8 & 0 \\ 0 & 1 & 1 & 2 & 0 \\ 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right],$$

possui uma variável livre (e portanto outras soluções que não a trivial). $\triangleleft$

[1] Lembra que uma equação vetorial pode ser reescrita como uma equação matricial!

Licença CC-BY-SA-3.0. Contato: livroscolaborativos@gmail.com

3.3 Transformações lineares

Uma **transformação linear** é um tipo especial de função que associa um vetor a outro vetor

$$T : \{\text{vetores}\} \to \{\text{vetores}\}$$

e com a propriedade adicional:

$$T\big(a\vec{u} + b\vec{v}\big) = aT(\vec{u}) + bT(\vec{v}), \quad \text{para todos } a, b \in \mathbb{R} \quad \text{e todos } \vec{u}, \vec{v}.$$

Em geral, pensamos em uma transformação linear como "transformando um vetor em outro" de forma linear, isto é, satisfazendo a propriedade acima.

Exemplo 16. A aplicação $T : \mathbb{R}^3 \to \mathbb{R}^3$ dada por $T(\vec{u}) = 3\vec{u}$ é a transformação linear que transforma um vetor de $\mathbb{R}^3$ no vetor que tem o triplo do comprimento. Utilizamos a própria fórmula da transformada para verificar que a tranformação T é de fato linear:

$$T\big(a\vec{u} + b\vec{v}\big) = 3(a\vec{u} + b\vec{v}) = a \cdot 3\vec{u} + b \cdot 3\vec{v} = aT(\vec{u}) + bT(\vec{v}).\triangleleft$$

Exemplo 17. Também é linear a transformação que faz uma rotação de um ângulo $\pi/2$. Tente se convencer geometricamente que esta transformação é uma transformação linear. Voltaremos a este exmeplo com mais detalhes na seção seguinte. $\triangleleft$

Exemplo 18. Consideramos a transformação $T : \mathbb{R}^2 \to \mathbb{R}^3$, que transforma vetores de $\mathbb{R}^2$ em vetores de $\mathbb{R}^3$, dada pela fórmula

$$T(x_1, x_2) = (x_1 - 3x_2, 3x_1 + 5x_2, -x_1 + x_2).$$

Nota que, excepcionalmente, estamos representando os vetores como linhas. É comum quando tratamos de transformações lineares. Esta fórmula nos diz, por exemplo, que os vetores

$$\vec{u} = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \quad \text{e} \quad \vec{v} = \begin{bmatrix} 0 \\ 3 \end{bmatrix}$$

são transformados nos vetores

$$T\big(\vec{u}\big) = \begin{bmatrix} 1 - 3 \cdot (-1) \\ 3 \cdot 1 + 5 \cdot (-1) \\ -1 + (-1) \end{bmatrix} = \begin{bmatrix} 4 \\ -2 \\ -2 \end{bmatrix} \quad \text{e} \quad T\big(\vec{v}\big) = \begin{bmatrix} 0 - 3 \cdot 3 \\ 3 \cdot 0 + 5 \cdot 3 \\ -0 + 3 \end{bmatrix} = \begin{bmatrix} -9 \\ 15 \\ 3 \end{bmatrix}.\triangleleft$$

3.4 Matriz de uma transformação linear

Voltemos à transformação linear

$$T(x_1, x_2) = (x_1 - 3x_2, 3x_1 + 5x_2, -x_1 + x_2)$$

do Exemplo 18. Lembramos que um vetor qualquer de $\mathbb{R}^2$ pode ser escrito como combinação linear dos vetores

$$\vec{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{e} \quad \vec{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

De fato,

$$\vec{u} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} = x_1\vec{e}_1 + x_2\vec{e}_2.$$

Logo, utilizando a propriedade de a transformação ser linear, temos que

$$T(\vec{u}) = T\big(x_1\vec{e}_1 + x_2\vec{e}_2\big) = x_1T(\vec{e}_1) + x_2T(\vec{e}_2).$$

Calculamos $T(\vec{e}_1)$ e $T(\vec{e}_2)$ pela fórmula dada da transformação linear:

$$T(\vec{e}_1) = \begin{bmatrix} 1 \\ 3 \\ -1 \end{bmatrix} \quad \text{e} \quad T(\vec{e}_2) = \begin{bmatrix} -3 \\ 5 \\ 1 \end{bmatrix}.$$

Licença CC-BY-SA-3.0. Contato: livroscolaborativos@gmail.com

Concluimos que

$$T(\vec{u}) = x_1 \begin{bmatrix} 1 \\ 3 \\ -1 \end{bmatrix} + x_2 \begin{bmatrix} -3 \\ 5 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & -3 \\ 3 & 5 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

Desta forma, associamos uma matriz de ordem 3×2 à transformação linear $T : \mathbb{R}^2 \to \mathbb{R}^3$.

O procedimento que aplicamos acima não é particular do exemplo que analisamos, de modo que é sempre possível associar a uma transformação linear $T : \mathbb{R}^n \to \mathbb{R}^m$ uma matriz A de ordem $m \times n$, chamada **matriz canônica associada à transformação linear** T ou apenas **matriz associada** a T, cujas colunas são os vetores $T(\vec{e}_1), T(\vec{e}_2), T(\vec{e}_3), \ldots, T(\vec{e}_n) \in \mathbb{R}^m$ (e portanto n colunas com m componentes cada).

Exemplo 19. A transformação linear $T : \mathbb{R}^3 \to \mathbb{R}^3$, $T(\vec{u}) = 5\vec{u}$, satisfaz

$$T(\vec{e}_1) = \begin{bmatrix} 5 \\ 0 \\ 0 \end{bmatrix}, \quad T(\vec{e}_2) = \begin{bmatrix} 0 \\ 5 \\ 0 \end{bmatrix} \quad \text{e} \quad T(\vec{e}_3) = \begin{bmatrix} 0 \\ 0 \\ 5 \end{bmatrix}.$$

Assim, podemos escrever

$$T(\vec{x}) = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} . \triangleleft$$

Exemplo 20. O método acima pode também ser aplicado para conseguirmos uma fórmula para transformações lineares como as do Exemplo 17. Vamos considerar a aplicação $T : \mathbb{R}^2 \to \mathbb{R}^2$ definida por

$$T(\vec{x}) = \text{rotação no sentido anti-horário de } \vec{x} \text{ por um ângulo } \theta \in (0, 2\pi).$$

A matriz da transformação linear tem por colunas a imagem por T dos vetores $\vec{e}_1$ e $\vec{e}_2$. Observamos que (ver figura)

$$T(\vec{e}_1) = \begin{bmatrix} \cos\theta \\ \operatorname{sen}\theta \end{bmatrix}, \quad T(\vec{e}_2) = \begin{bmatrix} -\operatorname{sen}\theta \\ \cos\theta \end{bmatrix}.$$

Logo, concluimos que

$$T(\vec{x}) = \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

3.5 Transformações injetoras, sobrejetoras e invertíveis

Como de costume, dada uma função $f : A \to B$, diz-se que A é o domínio de f enquanto B é o contradomínio. A imagem de f é o subconjunto de B que consiste de todos os elementos $y \in B$ tais que $f(x) = y$ ou, intuitivamente, que consiste de "todos os elementos de B que são atingidos pela função f".

Na hora de decidir se uma função é invertível ou não, duas propriedades são essenciais:

1) cada elemento de B ser a imagem de no máximo um elemento de A, caso em que f é dita **injetora** ou **injetiva**;

2) a imagem de f ser igual ao contradomínio, caso em que f diz-se **sobrejetora** ou **sobrejetiva**.

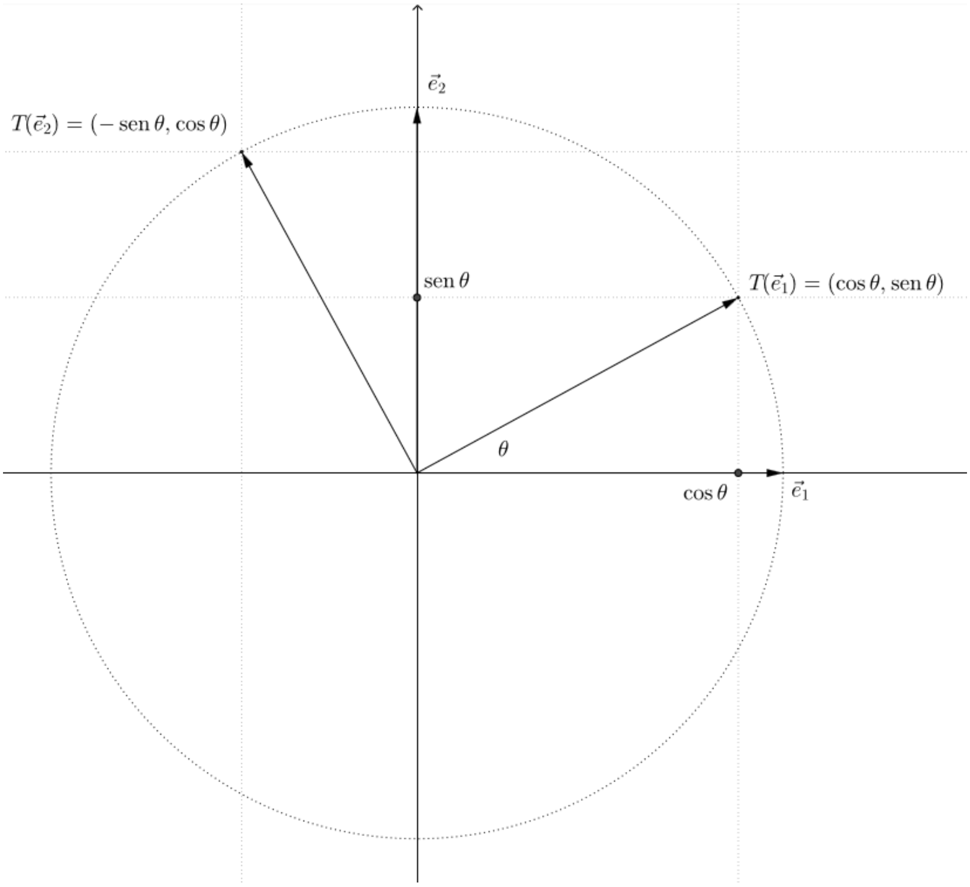
Quando estas duas propriedades são satisfeitas, isto é, quando a função f é injetora e sobrejetora, vale que f é **invertível**: podemos encontrar uma função $f^{-1} : B \to A$ que satisfaz

$$f^{-1}\big(f(x)\big) = x, \text{ para todo } x \in A \quad \text{e} \quad f\big(f^{-1}(y)\big) = y, \text{ para todo } y \in B.$$

Ao tentarmos encontrar uma **função inversa** f^{-1}, as propriedades de ser injetora e sobrejetora, aparecem naturalmente.

Estas definições são usuais para funções quaisquer. A partir de agora, vamos analisar quando que transformações lineares são injetoras, sobrejetoras e/ou invertíveis. Vamos ver, em particular, que é bem mais fácil estudar tais propriedades para transformações lineares.

Licença CC-BY-SA-3.0. Contato: livroscolaborativos@gmail.com



3.5.1 Transformações lineares injetoras

A transformação linear $T : \mathbb{R}^n \to \mathbb{R}^m$ é injetora quando, para $\vec{b} \in \mathbb{R}^m$, a equação

$$T(\vec{x}) = \vec{b}$$

possuir uma única solução ou nenhuma (no máximo uma – comparar com a definição acima). Como vimos, existe uma matriz A de ordem $m \times n$ associada à transformação linear T, de modo que temos que analisar as soluções de

$$A\vec{x} = \vec{b}.$$

Recaimos novamente em um sistema linear!

No caso particular em que $\vec{b} = \vec{0}$, o sistema homogêneo $A\vec{x} = \vec{0}$ sempre possui a solução trivial $\vec{x} = \vec{0}$. Neste caso, para que a transformação linear seja injetora devemos verificar que esta é a única solução de $A\vec{x} = \vec{0}$. Na verdade, é possível verificar que:

$$\boxed{T \text{ é injetora se, e somente se, } A\vec{x} = \vec{0} \text{ possui apenas a solução trivial.}}$$

Em particular, este caso somente é possível se $n \leq m$ (reler observações que seguem o Exemplo 3).

Justificativa da afirmação. Vamos provar matematicamente que, se soubermos que $A\vec{x} = \vec{0}$ possui apenas a solução trivial, então vale que $A\vec{x} = \vec{b}$ possui no máximo uma solução, para qualquer escolha de vetor $\vec{b}$. Isto, por sua vez, implica que T é injetora. Caso não existam soluções ou caso exista apenas uma, nada há para provar. Suponhamos que então que existem infinitas soluções (esta seria a única outra possibilidade). Em particular, existiriam duas distintas:

$$\vec{x}_1 \neq \vec{x}_2 \text{ tais que } A\vec{x}_1 = \vec{b} \text{ e também } A\vec{x}_2 = \vec{b}.$$

Por linearidade, temos então que $A(\vec{x}_1 - \vec{x}_2) = \vec{b} - \vec{b} = \vec{0}$. No entanto, nossa hipótese garante que somente existe a solução trivial para a equação $A\vec{x} = \vec{0}$, de modo que deveríamos ter $\vec{x}_1 = \vec{x}_2$. Em outras palavras, se $A\vec{x} = \vec{0}$ possuir apenas a solução trivial, então não existe mais do que uma solução para $A\vec{x} = \vec{b}$. Portanto, T é injetora. □

Notem que a afirmação acima é também equivalente a

$$\boxed{T \text{ é injetora se, e somente se, as colunas de } A \text{ são linearmente independentes.}}$$

Vamos deixar exemplos de injetividade para a subseção seguinte (ver Exemplos 21 e 22).

3.5.2 Transformações lineares sobrejetoras

A transformação linear $T : \mathbb{R}^n \to \mathbb{R}^m$ é sobrejetora quando, para todo $\vec{b} \in \mathbb{R}^n$, a equação

$$T(\vec{x}) = \vec{b}$$

possui alguma solução (comparar com a definição no início desta seção). Seja A a matriz de ordem $m \times n$ associada a T. Assim, para verificar que T é sobrejetora, devemos verificar que, para qualquer $\vec{b} \in \mathbb{R}^m$, o sistema linear

$$A\vec{x} = \vec{b}$$

possui ao menos uma solução (uma ou infinitas). Isto é equivalente a:

$$\boxed{T \text{ é sobrejetora se, e somente se, as colunas de } A \text{ geram todo o espaço } \mathbb{R}^m.}$$

Em particular, este caso somente é possível se $n \geq m$ (veremos mais adiante no curso – a imagem da transformação $T : \mathbb{R}^n \to \mathbb{R}^m$ será de dimensão no máximo igual a n).

Exemplo 21. Considere a transformação linear T cuja matriz associada é

$$A = \begin{bmatrix} 5 & 3 & 1 & 1 \\ 0 & -1 & 1 & -1 \\ 0 & 0 & 0 & 3 \end{bmatrix}.$$

Como são quatro colunas de $\mathbb{R}^3$, vemos que as colunas são LD e, portanto, T não é injetora.

Por outro lado, a matriz já está em forma escalonada. Vamos analisar se o sistema

$$A\vec{x} = \vec{b} \iff \left[\begin{array}{cccc|c} 5 & 3 & 1 & 1 & b_1 \\ 0 & -1 & 1 & -1 & b_2 \\ 0 & 0 & 0 & 3 & b_3 \end{array}\right]$$

possui solução para todo $\vec{b} \in \mathbb{R}^3$. De fato, o sistema possui solução (já que nenhuma linha da sua forma escalonada é inconsistente). Em verdade, o sistema possui uma variável livre. Logo, T é sobrejetora.

Exemplo 22. Considere a transformação linear T cuja matriz associada é

$$A = \begin{bmatrix} 3 & 1 \\ 5 & 7 \\ 0 & -4 \end{bmatrix}.$$

Como são somente duas colunas, é fácil ver que uma não é múltipla da outra: por causa das primeiras componentes, podemos pensar que a primeira coluna é três vezes a primeira, mas verificando a segunda linha já não dá certo $3 \cdot 7 \neq 5$. Logo, as colunas são LI e a transformação T é injetora.

Por outro lado, as colunas de A geram um espaço de dimensão no máximo 2; logo, não geram $\mathbb{R}^3$. Portanto, T não é sobrejetora.

3.5.3 Transformações lineares invertíveis

Pelas observações das últimas subseções, só é possível da transformação linear $T : \mathbb{R}^n \to \mathbb{R}^m$ ser invertível quando $m = n$, pois T deverá ser ao mesmo tempo injetora e sobrejetora. Veremos na semana seguinte um método para calcular a transformação inversa T^{-1}, que será também uma transformação linear.

Licença CC-BY-SA-3.0. Contato: livroscolaborativos@gmail.com

4

Semana 4

A operação de soma de matrizes e de multiplicação por escalar são definidas da mesma forma que definimos estas operações para vetores. Na soma, basta somar componente a componente, enquanto que na multiplicação por escalar o que fazemos é multiplicar todas as componentes da matriz pelo escalar (número) dado. Desta forma, por exemplo,

$$\begin{bmatrix} 1 & -1 & \pi & 0 \\ 11 & 11 & -2 & 0 \\ 9 & 0 & 4 & 4 \end{bmatrix} + \begin{bmatrix} 1 & 4 & 1 & 9 \\ -5 & 10 & 2 & 11 \\ -3 & 8 & 0 & -6 \end{bmatrix} = \begin{bmatrix} 1+1 & -1+4 & \pi+1 & 0+9 \\ 11-5 & 11+10 & -2+2 & 0+11 \\ 9-3 & 0+8 & 4+0 & 4-6 \end{bmatrix}$$
$$= \begin{bmatrix} 2 & 3 & \pi+1 & 9 \\ 6 & 21 & 0 & 11 \\ 6 & 8 & 4 & -2 \end{bmatrix}$$

enquanto

$$5 \cdot \begin{bmatrix} 1 & 4 & 1 & 9 \\ -5 & 10 & 2 & 11 \\ -3 & 8 & 0 & -6 \end{bmatrix} = \begin{bmatrix} 5\cdot 1 & 5\cdot 4 & 5\cdot 1 & 5\cdot 9 \\ 5\cdot(-5) & 5\cdot 10 & 5\cdot 2 & 5\cdot 11 \\ 5\cdot(-3) & 5\cdot 8 & 5\cdot 0 & 5\cdot(-6) \end{bmatrix} = \begin{bmatrix} 5 & 20 & 5 & 45 \\ -25 & 50 & 10 & 55 \\ -15 & 40 & 0 & -30 \end{bmatrix}.$$

4.1 Produto de matrizes

É no produto de matrizes que aparece uma operação diferente do que estamos acostumados. Já vimos como a multiplicação de uma matriz de ordem $m \times n$ por um vetor de $\mathbb{R}^n$ (ou matriz de ordem $n \times 1$) aparece naturalmente quando resolvemos uma sistema linear:

$$A\vec{x} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \cdots a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots a_{mn}x_n \end{bmatrix}.$$

O resultado é um vetor de $\mathbb{R}^m$ (ou matriz de ordem $m \times 1$). Observamos que necessariamente o número de colunas da matriz A deve ser igual ao número de linhas do vetor $\vec{x}$, para que seja possível realizar o produto como indicado.

Vamos generalizar este procedimento para definir o produto entre duas matrizes de qualquer ordem. Por exemplo, se A é como acima e B é uma matriz B de ordem $n \times 2$, o que vamos fazer é pensar nas colunas de B como dois vetores $\vec{b}_1$ e $\vec{b}_2$ e realizar os produtos como já estamos acostumados. O resultado será uma matriz cujas colunas são $A\vec{b}_1$ e $A\vec{b}_2$:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 & y_1 \\ x_2 & y_2 \\ \vdots & \vdots \\ x_n & y_n \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n & a_{11}y_1 + a_{12}y_2 + \cdots + a_{1n}y_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n & a_{21}y_1 + a_{22}y_2 + \cdots + a_{2n}y_n \\ \vdots & \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n & a_{m1}y_1 + a_{m2}y_2 + \cdots + a_{mn}y_n \end{bmatrix}.$$

O resultado é portanto uma matriz de ordem $m \times 2$.

De forma mais geral, podemos multiplicar uma matriz A de ordem $m \times n$ por qualquer outra matriz de ordem $n \times p$ e o resultado obtido será uma matriz de ordem $m \times p$. É fácil de lembrar: o número de colunas da primeira matriz deve ser igual ao número de linhas da segunda matriz e a matriz resultante terá por ordem o que "sobrar":

$$[\cdots]_{m\times \cancel{n}}[\cdots]_{\cancel{n}\times p} = [\cdots]_{m\times p}.$$

As colunas da matriz resultante são obtidas ao fazer o produto da primeira matriz com os vetores que formam as colunas da segunda:

$$AB = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} | & | & & | \\ \vec{b}_1 & \vec{b}_2 & \cdots & \vec{b}_p \\ | & | & & | \end{bmatrix} = \begin{bmatrix} | & | & & | \\ A\vec{b}_1 & A\vec{b}_2 & \cdots & A\vec{b}_p \\ | & | & & | \end{bmatrix}.$$

Acima, $\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_p \in \mathbb{R}^n$ são as colunas da matriz B. Conclua que $A\vec{b}_1, A\vec{b}_2, \ldots, A\vec{b}_p \in \mathbb{R}^m$ e que AB é uma matriz de ordem $m \times p$.

4.1.1 Exemplos

Vamos mostrar algumas contas que exemplificam tanto o procedimento quanto algumas das propriedades essenciais do produto de matrizes.

Exemplo 23. Vamos calcular um produto de matrizes utilizando o procedimento explicado na subseção anterior:

$$\begin{bmatrix} 1 & -1 & \pi & 0 \\ 11 & 11 & -2 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 & 4 & 1 & 9 \\ -5 & 10 & 2 & 11 \\ -3 & 8 & 0 & -6 \\ 0 & 2 & -2 & 0 \end{bmatrix} = \begin{bmatrix} 6-3\pi & -6+8\pi & -1 & -2-6\pi \\ 50 & 138 & 33 & 232 \end{bmatrix}.$$

O elemento $(AB)_{ij}$ da linha i e coluna j da matriz AB, pode ser obtido ao "multiplicar"[1] a linha i de A pela coluna j de B:

$$\begin{array}{ccccccccccc} (AB)_{11} & = & 1\cdot 1 & + & (-1)\cdot(-5) & + & \pi\cdot(-3) & + & 0\cdot 0 & = & 6-3\pi \\ (AB)_{12} & = & 1\cdot 4 & + & (-1)\cdot 10 & + & \pi\cdot 8 & + & 0\cdot 2 & = & -6+8\pi \\ (AB)_{13} & = & 1\cdot 1 & + & (-1)\cdot 2 & + & \pi\cdot 0 & + & 0\cdot(-2) & = & -1 \\ (AB)_{14} & = & 1\cdot 9 & + & (-1)\cdot 11 & + & \pi\cdot(-6) & + & 0\cdot 0 & = & -2-6\pi \\ (AB)_{21} & = & 11\cdot 1 & + & 11\cdot(-5) & + & (-2)\cdot(-3) & + & 0\cdot 0 & = & 50 \\ (AB)_{22} & = & 11\cdot 4 & + & 11\cdot 10 & + & (-2)\cdot 8 & + & 0\cdot 2 & = & 138 \\ (AB)_{23} & = & 11\cdot 1 & + & 11\cdot 2 & + & (-2)\cdot 0 & + & 0\cdot(-2) & = & 33 \\ (AB)_{24} & = & 11\cdot 9 & + & 11\cdot 11 & + & (-2)\cdot(-6) & + & 0\cdot 0 & = & 232. \end{array}$$

Observamos, a partir deste exemplo, que **o produto de matrizes não é comutativo**. De fato, as matrizes AB e BA sequer tem a mesma ordem. Além disto, pode ser que seja possível calcular AB, mas não BA. Pense, por exemplo, em uma matriz A de ordem 3×2 e em outra matriz B de ordem 2×5.

Exemplo 24. Uma propriedade de **matrizes quadradas**, isto é, matrizes cujo número de linhas e de colunas é igual, é que o produto delas fornece uma terceira matriz com a mesma ordem:

$$[\cdots]_{n\times \cancel{n}}[\cdots]_{\cancel{n}\times n} = [\cdots]_{n\times n}.$$

Mesmo neste caso, o produto de matrizes **pode não ser comutativo**: é possível encontrar matrizes tais que $AB \neq BA$. Por exemplo,

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

enquanto

$$\begin{bmatrix} 1 & -1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & -1 \end{bmatrix}. \triangleleft$$

[1] O leitor já familiarizado com geometria analítica pode identificar o elemento ij da matriz resultante como o produto escalar entre a linha i de A e a coluna j de B.

Licença CC-BY-SA-3.0. Contato: livroscolaborativos@gmail.com

4.1.2 Uma interpretação para resolução de sistemas lineares

Nós podemos pensar em

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 & y_1 \\ x_2 & y_2 \\ \vdots & \vdots \\ x_n & y_n \end{bmatrix} = \begin{bmatrix} b_1 & c_1 \\ b_2 & c_2 \\ \vdots & \vdots \\ b_n & c_n \end{bmatrix}$$

como a forma matricial de escrever dois sistemas lineares que possuem a mesma matriz associada:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} : \quad A\vec{x} = \vec{b}, \quad A\vec{y} = \vec{c}.$$

(por quê?) É possível de resolver os dois sistemas concomitantemente ao escrever uma matriz aumentada generalizada:

$$\left[\begin{array}{cccc|cc} a_{11} & a_{12} & \cdots & a_{1n} & b_1 & c_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 & c_2 \\ \vdots & \vdots & & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_n & c_n \end{array}\right]$$

A análise das soluções é feita da mesma forma que aprendemos nas notas da primeira semana. Observem que esta é uma forma de economizar as contas e não precisar escalonar duas vezes a matriz A associada!

Exercício 1. Escrever dois sistemas lineares que possuem a mesma matriz associada. Resolver usando a matriz aumentada generalizada acima.

Evidentemente, este método funcionaria para resolver ao mesmo tempo 3 ou 4 ou 20 sistemas lineares ao mesmo tempo, desde que a matriz associada seja a mesma em todos eles.

4.2 Matriz transposta

A **matriz transposta** de uma matriz A, de ordem $m \times n$, é a matriz A^T que tem por colunas as linhas de A. Consequentemente, A^T é uma matriz de ordem $n \times m$.

Exemplo 25. As transpostas das matrizes

$$A = \begin{bmatrix} -1 \\ 10 \\ -9 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & -1 \\ 2 & 11 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & -1 & \pi & 0 \\ 11 & 11 & -2 & 0 \\ 9 & 4 & 4 & 4 \end{bmatrix}$$

são, respectivamente, as matrizes

$$A^T = \begin{bmatrix} -1 & 10 & 9 \end{bmatrix}, \quad B^T = \begin{bmatrix} 1 & 2 \\ -1 & 11 \end{bmatrix}, \quad C^T = \begin{bmatrix} 1 & 11 & 9 \\ -1 & 11 & 4 \\ \pi & -2 & 4 \\ 0 & 0 & 4 \end{bmatrix}. \triangleleft$$

Valem as seguintes propriedades:

- $(A^T)^T = A$
- Linearidade: $(xA + yB)^T = xA^T + yB^T$
- Produto: $(AB)^T = B^T A^T$

Exercício 2. Verifique que estas propriedades são válidas em alguns exemplos (escolha algumas matrizes para fazer as contas).

Licença CC-BY-SA-3.0. Contato: livroscolaborativos@gmail.com

4.3 Matriz inversa

Agora que temos definido um produto de matrizes, é natural de nos perguntarmos quais das propriedades usuais do produto entre números reais se mantém válidas.

Por exemplo, a **matriz identidade** I_n é a matriz quadrada de ordem $n \times n$ que tem 1 na diagonal principal e 0 nas demais posições. No caso 3×3, temos

$$I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Esta matriz satisfaz $AI_n = A$ para qualquer matriz A de ordem $m \times n$. Da mesma forma, temos $I_n B = B$ para qualquer matriz B de ordem $n \times m$ (observe atentamente os índices!). O nome identidade vem desta propriedade, que é parecida com a propriedade da unidade nos números reais: $1 \cdot x = x = x \cdot 1$.

Existindo a matriz identidade, outra pergunta natural é se existe um inverso multiplicativo. Para números reais, qualquer número não nulo possui:

$$x \in \mathbb{R}, x \neq 0 \implies x^{-1} = \frac{1}{x} \text{ é um inverso multiplicativo,}$$

isto é, satisfaz $x \cdot x^{-1} = 1$.

Vamos procurar por **matrizes inversas**: dada uma matriz A, queremos encontrar uma matriz A^{-1} de modo que

$$AA^{-1} = I_n = A^{-1}A.$$

Escrevemos de duas formas diferentes acima, pois o produto de matrizes não é comutativo. A matriz A^{-1} é chamada a **matriz inversa** de A. Observe que A **deve ser quadrada** (por quê?).

4.3.1 Caso 2×2

Vamos procurar pela inversa de

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

Escrevemos

$$A^{-1} = \begin{bmatrix} x_1 & y_1 \\ x_2 & y_2 \end{bmatrix}$$

e queremos descobrir os valores x_1, x_2, y_1, y_2 que satisfazem

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x_1 & y_1 \\ x_2 & y_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Pela interpretação que demos anteriormente ao produto de matrizes, encontrar

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad \text{e} \quad \vec{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

é equivalente a resolver ao mesmo tempo os dois sistemas

$$A\vec{x} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \vec{e}_1 \quad \text{e} \quad A\vec{y} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \vec{e}_2.$$

A ideia é então resolver por escalonamento os sistemas cuja matriz aumentada associada é:

$$\left[\begin{array}{cc|cc} a & b & 1 & 0 \\ c & d & 0 & 1 \end{array}\right].$$

Tem-se

$$\left[\begin{array}{cc|cc} a & b & 1 & 0 \\ c & d & 0 & 1 \end{array}\right] \xrightarrow{c\cdot\ell_1 \text{ e } a\cdot\ell_2} \left[\begin{array}{cc|cc} ac & bc & c & 0 \\ ac & ad & 0 & a \end{array}\right] \xrightarrow{-\ell_1+\ell_2 \text{ em } \ell_2} \left[\begin{array}{cc|cc} ac & bc & c & 0 \\ 0 & ad-bc & -c & a \end{array}\right]$$

Licença CC-BY-SA-3.0. Contato: livroscolaborativos@gmail.com

$$\xrightarrow{\ell_2 \div (ad-bc)} \left[\begin{array}{cc|cc} ac & bc & c & 0 \\ 0 & 1 & \dfrac{-c}{ad-bc} & \dfrac{a}{ad-bc} \end{array}\right] \xrightarrow{-bc\ell_2+\ell_1 \text{ em } \ell_1} \left[\begin{array}{cc|cc} ac & 0 & \dfrac{acd}{ad-bc} & \dfrac{-abc}{ad-bc} \\ 0 & 1 & \dfrac{-c}{ad-bc} & \dfrac{a}{ad-bc} \end{array}\right]$$

$$\xrightarrow{\ell_1 \div ac} \left[\begin{array}{cc|cc} 1 & 0 & \dfrac{d}{ad-bc} & \dfrac{-b}{ad-bc} \\ 0 & 1 & \dfrac{-c}{ad-bc} & \dfrac{a}{ad-bc} \end{array}\right].$$

Daí concluimos (depois colocar em evidência o fator $ad - bc$ que está dividindo) que:

$$\boxed{A^{-1} = \frac{1}{ad-bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} = \frac{1}{\det A} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.}$$

Nesta última expressão, definimos o **determinante** de uma matriz de ordem 2×2:

$$\det A = ad - bc$$

e, como foi necessário dividir por este valor, concluimos que:

$$\boxed{\text{só existe a matriz inversa de } A \text{ caso } \det A \neq 0.}$$

Observação 26. Veremos na seção seguinte que o processo que utilizamos para inverter a matriz A funciona para matrizes de qualquer ordem, mas é trabalhoso. No caso de uma matriz 2×2, talvez seja interessante memorizar a fórmula, já que não é tão complicada. Podemos pensar da seguinte maneira:

- Calculamos $\det A$. Se for diferente de zero, existe a inversa. Nós partimos de A e dividimos pelo determinante.
- Em seguida, trocamos de lugar os elementos da diagonal principal (a e d).
- Finalmente, trocamos o sinal dos elementos da outra diagonal.

Atenção! Este método apenas funciona para matrizes quadradas de ordem 2×2!

Exemplo 27. Sejam as matrizes

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

Calculamos

$$\det A = 1 \cdot 3 - (-1) \cdot 1 = 4 \neq 0.$$

Logo, A possui inversa e temos

$$A^{-1} = \frac{1}{4} \begin{bmatrix} 3 & -1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 3/4 & -1/4 \\ 1/4 & 1/4 \end{bmatrix}.$$

(trocamos de lugar os elementos da diagonal principal e de sinal os elementos da outra diagonal).

Façamos o mesmo para a matriz B:

$$\det B = 1 \cdot 1 - 1 \cdot 0 = 1 \neq 0.$$

Logo,

$$B^{-1} = \frac{1}{1} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 3/4 & -1/4 \\ 1/4 & 1/4 \end{bmatrix}.$$

(trocamos de lugar os elementos da diagonal principal – neste caso eram iguais – e de sinal dos elementos da outra diagonal).

Já para a matriz C, temos

$$\det C = (-1) \cdot 2 - 2 \cdot (-1) = 0$$

e portanto C não é invertível (isto é, não existe a matriz inversa C^{-1}). $\triangleleft$

Licença CC-BY-SA-3.0. Contato: livroscolaborativos@gmail.com

4.3.2 Algoritmo para ordem maior

Considere uma matriz de ordem $n \times n$:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}.$$

Para obter a matriz inversa, devemos ter:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nn} \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}.$$

Da mesma forma que na seção anterior, isto é equivalente a resolver simultaneamente os sistemas:

$$A\vec{x}_1 = \vec{e}_1, \quad A\vec{x}_2 = \vec{e}_2, \ldots, A\vec{x}_n = \vec{e}_n,$$

onde $\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_n$ são as colunas da matriz inversa A^{-1}. Assim, devemos escrever a matriz aumentada associada a estes sistemas lineares:

$$\left[\begin{array}{cccc|cccc} a_{11} & a_{12} & \cdots & a_{1n} & 1 & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & a_{2n} & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & 0 & 0 & \cdots & 1 \end{array}\right]$$

ou, de forma mais sucinta,

$$\left[\, A \mid I \,\right].$$

E este é o algoritmo de resolução: Deixar a matriz A em forma escalonada reduzida (até chegar na matriz identidade) de modo que as soluções obtidas já serão a matriz inversa de A:

$$\left[\, I \mid A^{-1} \,\right].$$

Observação 28. Para que seja possível encontrar as soluções, como indicado acima, a forma escalonada da matriz A deve possuir n posições de pivô (caso contrário, algum dos sistemas acima seria impossível), de modo que todas as colunas de A são colunas pivô.

Se A possuir alguma coluna sem posição de pivô, podemos parar imediatamente o algoritmo, pois isto significa que a matriz não é invertível.

Exemplo 29. Vamos decidir se a matriz abaixo é invertível e, em caso afirmativo, determinar a inversa:

$$A = \begin{bmatrix} 2 & 1 & 1 & 0 \\ 4 & 3 & 3 & 1 \\ 8 & 7 & 9 & 5 \\ 6 & 7 & 9 & 8 \end{bmatrix}.$$

De acordo com o algoritmo, devemos escalonar

$$\left[\begin{array}{cccc|cccc} 2 & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 4 & 3 & 3 & 1 & 0 & 1 & 0 & 0 \\ 8 & 7 & 9 & 5 & 0 & 0 & 1 & 0 \\ 6 & 7 & 9 & 8 & 0 & 0 & 0 & 1 \end{array}\right].$$

Caso a matriz seja invertível, conseguiremos deixar a matriz identidade do lado esquerdo desta matriz aumentada. Começamos eliminando os termos da primeira coluna (abaixo do 2 que está na posição de pivô):

$$\left[\begin{array}{cccc|cccc} 2 & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & -2 & 1 & 0 & 0 \\ 0 & 3 & 5 & 5 & -4 & 0 & 1 & 0 \\ 0 & 4 & 6 & 8 & -3 & 0 & 0 & 1 \end{array}\right] \xrightarrow{\text{segunda coluna}} \left[\begin{array}{cccc|cccc} 2 & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & -2 & 1 & 0 & 0 \\ 0 & 0 & 2 & 2 & 2 & -3 & 1 & 0 \\ 0 & 0 & 2 & 4 & 5 & -4 & 0 & 1 \end{array}\right]$$

Licença CC-BY-SA-3.0. Contato: livroscolaborativos@gmail.com

$$\xrightarrow{\text{terceira}} \left[\begin{array}{cccc|cccc} 2 & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & -2 & 1 & 0 & 0 \\ 0 & 0 & 2 & 2 & 2 & -3 & 1 & 0 \\ 0 & 0 & 0 & 2 & 3 & -1 & -1 & 1 \end{array}\right] \xrightarrow{\text{reduzida - } 4^a} \left[\begin{array}{cccc|cccc} 2 & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & -7/2 & 3/2 & 1/2 & -1/2 \\ 0 & 0 & 1 & 0 & -1/2 & -1 & 1 & -1/2 \\ 0 & 0 & 0 & 2 & 3 & -1 & -1 & 1 \end{array}\right]$$

$$\to \left[\begin{array}{cccc|cccc} 2 & 1 & 0 & 0 & 3/2 & 1 & -1 & 1/2 \\ 0 & 1 & 0 & 0 & -3 & 5/2 & -1/2 & 0 \\ 0 & 0 & 1 & 0 & -1/2 & -1 & 1 & -1/2 \\ 0 & 0 & 0 & 1 & 3/2 & -1/2 & -1/2 & 1/2 \end{array}\right] \to \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 9/4 & -3/4 & -1/4 & 1/4 \\ 0 & 1 & 0 & 0 & -3 & 5/2 & -1/2 & 0 \\ 0 & 0 & 1 & 0 & -1/2 & -1 & 1 & -1/2 \\ 0 & 0 & 0 & 1 & 3/2 & -1/2 & -1/2 & 1/2 \end{array}\right].$$

Concluimos que

$$A^{-1} = \begin{bmatrix} 9/4 & -3/4 & -1/4 & 1/4 \\ -3 & 5/2 & -1/2 & 0 \\ -1/2 & -1 & 1 & -1/2 \\ 3/2 & -1/2 & -1/2 & 1/2 \end{bmatrix}.$$

Verifique como exercício que esta matriz é de fato a matriz inversa de A, isto é, calcule o produto $A \cdot A^{-1}$ e verifique que resulta em I_4.

Notamos que, caso nosso único interesse fosse decidir se A é invertível, sem necessariamente calcular sua inversa, poderíamos ter feito o escalonamento de A (nem precisa ser da matriz aumentada $\left[\, I \mid A^{-1} \,\right]$) e parado o processo quando chegamos em

$$A \sim \begin{bmatrix} 2 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 2 & 2 \\ 0 & 0 & 0 & 2 \end{bmatrix},$$

pois já está claro que todas as colunas possuem posição de pivô. $\triangleleft$

Exemplo 30. Vamos decidir se a matriz abaixo é invertível e, em caso afirmativo, determinar a inversa:

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 3 & -3 \\ -3 & -9 & 9 \end{bmatrix}.$$

De acordo com o algoritmo, devemos escalonar

$$\left[\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 3 & -3 & 0 & 1 & 0 \\ -3 & -9 & 9 & 0 & 0 & 1 \end{array}\right].$$

Temos

$$\left[\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 2 & -4 & -1 & 1 & 0 \\ 0 & -6 & 12 & 3 & 0 & 1 \end{array}\right] \to \left[\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 2 & -4 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 3 & 1 \end{array}\right].$$

Portanto, a terceira coluna não possui posição de pivô e a matriz A não possui inversa.

4.3.3 Aplicação na resolução de sistemas lineares

Sistemas lineares de ordem $n \times n$

$$A\vec{x} = \vec{b},$$

cuja matriz associada A é invertível, são sistemas que possuem exatamente uma solução, para qualquer vetor $\vec{b} \in \mathbb{R}^n$ (ver Subseção 4.3.4).

A existência da matriz inversa A^{-1} permite-nos multiplicar ambos os lados do sistema por A^{-1} para obter $\vec{x} = A^{-1} \cdot A = A^{-1}\vec{b}$. Logo,

$$\boxed{\vec{x} = A^{-1}\vec{b}.}$$

Licença CC-BY-SA-3.0. Contato: livroscolaborativos@gmail.com

Exemplo 31. Resolver o sistema

$$A = \begin{bmatrix} 1 & 1 \\ -1 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}.$$

Já calculamos a matriz inversa no Exemplo 27:

$$A^{-1} = \begin{bmatrix} 3/4 & -1/4 \\ 1/4 & 1/4 \end{bmatrix}.$$

Segue que a solução é

$$\vec{x} = A^{-1}\vec{b} = \begin{bmatrix} 3/4 & -1/4 \\ 1/4 & 1/4 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 1/4 \\ 3/4 \end{bmatrix}. \lhd$$

Embora nossa solução do sistema possa parecer "elegante" ao utilizar a matriz inversa, este método é muito ineficiente. Na verdade, escalonar a matriz aumentada associada ao sistema exige um custo computacional muito menor do que calcular a inversa da matriz e, portanto, em geral se usa o escalonamento.

Matrizes inversas têm uma importância mais teórica no nosso curso, como veremos na subseção abaixo.

4.3.4 Uma primeira caracterização de matrizes invertíveis

Vamos, nesta subseção, verificar (provar de forma rigorosa) que

"a matriz A é invertível se, e somente se, o sistema linear $A\vec{x} = \vec{b}$ possui exatamente uma solução, para qualquer vetor $\vec{b} \in \mathbb{R}^n$."

($\Longrightarrow$) Se A é invertível, então conseguimos resolver todos os sistemas

$$A\vec{x}_1 = \vec{e}_1, \quad A\vec{x}_2 = \vec{e}_2, \cdots, A\vec{x}_n = \vec{e}_n$$

concomitantemente. De fato, qualquer vetor $\vec{b}$ pode ser escrito como

$$\vec{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = b_1\vec{e}_1 + b_2\vec{e}_2 + \cdots + b_n\vec{e}_n,$$

e daí verificamos que se pode construir uma solução $\vec{x}$ pela fórmula

$$\vec{x} = b_1\vec{x}_1 + b_2\vec{x}_2 + \cdots + b_n\vec{x}_n,$$

já que pela linearidade do produto da matriz A por vetores, temos

$$\begin{aligned} A\vec{x} &= b_1A\vec{x}_1 + b_2A\vec{x}_2 + \cdots + b_nA\vec{x}_n, \\ &= b_1\vec{e}_1 + b_2\vec{e}_2 + \cdots + b_n\vec{e}_n \\ &= \vec{b}. \end{aligned}$$

($\Longleftarrow$) Reciprocamente, suponhamos que o sistema possua exatamente uma solução, para qualquer vetor $\vec{b} \in \mathbb{R}^n$. Em particular, podemos resolver os sitemas

$$A\vec{x}_1 = \vec{e}_1, \quad A\vec{x}_2 = \vec{e}_2, \cdots, A\vec{x}_n = \vec{e}_n$$

e escrever a matriz inversa de acordo com o nosso algoritmo:

$$A^{-1} = \begin{bmatrix} | & | & & | \\ \vec{x}_1 & \vec{x}_2 & \cdots & \vec{x}_n \\ | & | & & | \end{bmatrix}.$$

Licença CC-BY-SA-3.0. Contato: livroscolaborativos@gmail.com

5

Semana 5

5.1 Interpretações de sistemas lineares e de matrizes invertíveis

Uma das belezas da Álgebra Linear é que diversos conceitos podem ser interpretados de maneiras diferentes. Assim, propriedades intuitivas de uma formulação de um certo problema podem ser traduzidas para outras formulações onde estas propriedades nem sempre são tão aparentes.

Um sistema linear pode ser reescrito de diversas maneiras.

1. **Sistema linear**:
$$\begin{cases} 7x_1 - 3x_2 &= 2 \\ 3x_1 - x_2 + x_3 &= 0 \\ x_2 + 2x_3 &= -2 \end{cases}.$$
Aparece geralmente na formulação ou na modelagem do problema. As variáveis x_1, x_2, x_3 indicam as variáveis que estamos interessados em analisar e o problema específico no fornece os coeficientes do sistema.

2. **Matriz aumentada associada**:
$$\left[\begin{array}{ccc|c} 7 & -3 & 0 & 2 \\ 3 & -1 & 1 & 0 \\ 0 & 1 & 2 & -2 \end{array}\right]$$
Esta forma é adequada para resolver o sistema por escalonamento (também conhecido como método de eliminação gaussiana). Poderíamos fazer eliminação gaussiana sem recorrer à matriz aumentada associada, mas isto implicaria em ter que ficar reescrevendo as variáveis independentes a todo momento, além de tomar certo cuidado para não se confundir com os coeficientes nulos. O escalonamento de uma matriz torna o procedimento mais automatizado e minimiza as chances de erro.

3. **Forma matricial**:
$$\begin{bmatrix} 7 & -3 & 0 \\ 3 & -1 & 1 \\ 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ -2 \end{bmatrix} \quad \text{ou} \quad A\vec{x} = \vec{b}.$$
Nesta forma, aparece o produto de uma matriz por um vetor. Depois podemos considerar produto de matrizes e a resolução de sistemas lineares concomitantemente (ver também capítulo da Semana 04). Caso a matriz A acima seja invertível, sabemos que o sistema possui apenas uma solução e que
$$\vec{x} = A^{-1}\vec{b}.$$

4. **Forma vetorial** ou **combinação linear das colunas**:
$$x_1 \begin{bmatrix} 7 \\ 3 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} -3 \\ -1 \\ 1 \end{bmatrix} + x_3 \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ -2 \end{bmatrix}$$

Aqui, há uma interpretação mais geométrica. Estudar a existência de soluções do sistema linear é equivalente a se perguntar se o vetor

$$\vec{b} = \begin{bmatrix} 2 \\ 0 \\ -2 \end{bmatrix}$$

pode ser escrito como combinação linear dos vetores

$$\vec{v}_1 = \begin{bmatrix} 7 \\ 3 \\ 0 \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} -3 \\ -1 \\ 1 \end{bmatrix} \quad \text{e} \quad \vec{v}_3 = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix},$$

que são as colunas da matriz A. Logo, resolver o sistema linear é equivalente a perguntar se $\vec{b}$ está no espaço gerado por $\vec{v}_1, \vec{v}_2$ e $\vec{v}_3$, isto é, se $\vec{b} \in \text{Span}\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$.

5. **Transformações lineares**: Decidir se $\vec{b}$ pertence à imagem da transformação linear $T : \mathbb{R}^3 \to \mathbb{R}^3$, definida por:

$$T(x_1, x_2, x_3) = (7x_1 - 3x_2, 3x_1 - x_2 + x_3, x_2 + 2x_3).$$

Apesar de esta ser apenas uma outra maneira de escrever a forma matricial, ela tem uma interpretação diferente, já que traz a linguagem de teoria de funções para o contexto de Álgebra Linear. De fato, vemos que a matriz canônica associada a T é A, de modo que

$$T(\vec{x}) = A\vec{x}$$

e os conceitos introduzidos anteriormente podem ser traduzidos de forma natural. Por exemplo, combinando as interpretações das formas matricial e vetorial, podemos dizer que são equivalentes:

- existir solução do sistema linear $A\vec{x} = \vec{b}$;
- vetor $\vec{b}$ pertencer ao conjunto gerado pelas colunas da matriz A;
- vetor $\vec{b}$ pertencer à imagem da transformação linear $T(\vec{x}) = A\vec{x}$.

Vamos analisar dois sistemas lineares, com todas as ferramentas que estudamos até agora e com dois objetivos principais: familiarizar ainda mais os leitores com os conceitos até então introduzidos e preparar o caminho para o resultado da subseção seguinte sobre matrizes invertíveis.

Os exemplos abaixo são de sistemas lineares cuja matriz associada é **quadrada**, já que estamos estudando matrizes invertíveis, que devem ser quadradas. De qualquer maneira, alguns raciocínios podem ser adaptados para quaisquer sistemas (quais?).

Exemplo 32. Seja o sistema linear dado acima, cuja matriz aumentada associada é

$$\left[\begin{array}{ccc|c} 7 & -3 & 0 & 2 \\ 3 & -1 & 1 & 0 \\ 0 & 1 & 2 & -2 \end{array}\right].$$

Nossa técnica básica de escalonamento permite fazer uma análise de todos os pontos de vista acima. Os primeiros passos do escalonamento podem ser:

$$\xrightarrow{\ell_2 \leftrightarrow \ell_3} \left[\begin{array}{ccc|c} 7 & -3 & 0 & 2 \\ 0 & 1 & 2 & -2 \\ 3 & -1 & 1 & 0 \end{array}\right] \xrightarrow{-(3/7)\ell_1 + \ell_3 \text{ em } \ell_3} \left[\begin{array}{ccc|c} 7 & -3 & 0 & 2 \\ 0 & 1 & 2 & -2 \\ 0 & 2/7 & 1 & -6/7 \end{array}\right]$$

$$\xrightarrow{7 \cdot \ell_3} \left[\begin{array}{ccc|c} 7 & -3 & 0 & 2 \\ 0 & 1 & 2 & -2 \\ 0 & 2 & 7 & -6 \end{array}\right] \xrightarrow{-2\ell_2 + \ell_3 \text{ em } \ell_3} \left[\begin{array}{ccc|c} 7 & -3 & 0 & 2 \\ 0 & 1 & 2 & -2 \\ 0 & 0 & 3 & -2 \end{array}\right]$$

Estas continhas iniciais já revelam muito! Por exemplo:

Licença CC-BY-SA-3.0. Contato: livroscolaborativos@gmail.com

- Chegamos à forma escalonada e não há linha do tipo $[0\ 0\ 0 \mid 1]$; logo, o sistema possui solução;
- De fato, podemos dizer mais: já que todas as colunas da matriz A possuem posição de pivô, existe apenas uma solução;
- Podemos dizer mais ainda: já que todas as colunas da matriz A possuem posição de pivô, o sistema possuiria solução, quaisquer que fossem os números do lado direito da igualdade, isto é, as coordenadas do vetor $\vec{b}$;
- Todas as colunas da matriz A possuem posição de pivô e portanto a forma escalonada reduzida da matriz A será a matriz identidade I_3. Em particular, seria possível obter a matriz inversa a A, desde que começássemos escalonando a matriz $[A \mid I_3]$.

Agora, vejamos como poderíamos reescrever estas conclusões utilizando os conceitos estudados nas semanas anteriores:

- Para todo $\vec{b} \in \mathbb{R}^3$, o sistema $A\vec{x} = \vec{b}$ possui exatamente uma solução;
- Todo vetor $\vec{b} \in \mathbb{R}^3$ pode ser escrito como combinação linear das colunas $\vec{v}_1, \vec{v}_2, \vec{v}_3$ de A;
- Todo vetor $\vec{b} \in \mathbb{R}^3$ pertence a $\mathrm{Span}\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$, espaço gerado pelas colunas de A;
- $\mathrm{Span}\{\vec{v}_1, \vec{v}_2, \vec{v}_3\} = \mathbb{R}^3$;
- A é equivalente por linhas à matriz identidade I_3;
- A é invertível;
- A imagem da transformação linear $T(\vec{x}) = A\vec{x}$ é todo o espaço $\mathbb{R}^3$, isto é, T é sobrejetora;
- A equação $T(\vec{x}) = \vec{b}$ possui no máximo uma solução para cada $\vec{b} \in \mathbb{R}^3$ (de fato, possui exatamente uma!), de modo que T é injetora;
- T é invertível, já que é ao mesmo tempo injetora e sobrejetora.

Todas estas conclusões vieram de alguns passos que levaram a matriz A à sua forma escalonada. Como veremos adiante, ainda mais coisas poderiam ser ditas! $\triangleleft$

Exemplo 33. Analisamos agora o sistema linear cuja matriz aumentada assciada é

$$\left[\begin{array}{ccc|c} 1 & -3 & 0 & 2 \\ 0 & 4 & 1 & -3 \\ 3 & -1 & 2 & 0 \end{array}\right].$$

Nossa técnica básica de escalonamento permite fazer uma análise de todos os pontos de vista acima. Os primeiros passos do escalonamento podem ser:

$$\xrightarrow{-3\ell_1+\ell_3 \text{ em } \ell_3} \left[\begin{array}{ccc|c} 1 & -3 & 0 & 2 \\ 0 & 4 & 1 & -3 \\ 0 & 8 & 2 & -6 \end{array}\right] \xrightarrow{-2\ell_2+\ell_3 \text{ em } \ell_3} \left[\begin{array}{ccc|c} 1 & -3 & 0 & 2 \\ 0 & 4 & 1 & -3 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

Assim como no exemplo anterior, estas contas também revelam toda a estrutura do sistema linear e de sua matriz associada.

- Este sistema possui infinitas soluções, pois possui uma variável livre;
- Nem todas as colunas da matriz A possuem posição de pivô. Consequentemente, escolhas diferentes do vetor $\vec{b}$, poderiam resultar em um sistema impossível. De fato, foi uma coincidência que nossa última conta foi $(-2)\cdot(-3)+(-6)$, que resultou em um zero. Neste exemplo, qualquer variação na segunda componente de $\vec{b}$ nos daria um sistema impossível:

$$\left[\begin{array}{ccc|c} 1 & -3 & 0 & 2 \\ 0 & 4 & 1 & 1 \\ 3 & -1 & 2 & 0 \end{array}\right] \xrightarrow{-3\ell_1+\ell_3 \text{ em } \ell_3} \left[\begin{array}{ccc|c} 1 & -3 & 0 & 2 \\ 0 & 4 & 1 & 1 \\ 0 & 8 & 2 & -6 \end{array}\right] \xrightarrow{-2\ell_2+\ell_3 \text{ em } \ell_3} \left[\begin{array}{ccc|c} 1 & -3 & 0 & 2 \\ 0 & 4 & 1 & 1 \\ 0 & 0 & 0 & -8 \end{array}\right].$$

Licença CC-BY-SA-3.0. Contato: livroscolaborativos@gmail.com

- É um fato geral: se *nem todas* as colunas de A possuem posição de pivô, podemos encontrar vetores $\vec{b}$ de modo que o sistema seja impossível e vetores $\vec{b}$ cujo sistema possua infinitas soluções;
- A forma escalonada reduzida de A não é a identidade, isto é, A não é equivalente por linhas a I_3. Em particular, A não é invertível.

Vamos reescrever estas conclusões como no exemplo anterior:

- Existe $\vec{b}_1 \in \mathbb{R}^3$ tal que o sistema $A\vec{x} = \vec{b}_1$ possui infinitas soluções e existe $\vec{b}_2 \in \mathbb{R}^3$ tal que o sistema $A\vec{x} = \vec{b}_2$ possui infinitas soluções;
- Nem todo vetor $\vec{b} \in \mathbb{R}^3$ pode ser escrito como combinação linear das colunas $\vec{v}_1, \vec{v}_2, \vec{v}_3$ de A;
- Existe $\vec{b} \in \mathbb{R}^3$ não pertencente a $\mathrm{Span}\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$;
- $\mathrm{Span}\{\vec{v}_1, \vec{v}_2, \vec{v}_3\} \neq \mathbb{R}^3$;
- A não é equivalente por linhas à matriz identidade I_3;
- A não é invertível;
- A imagem da transformação linear $T(\vec{x}) = A\vec{x}$ não é todo o espaço $\mathbb{R}^3$, isto é, T não é sobrejetora;
- A equação $T(\vec{x}) = \vec{b}$ pode possuir mais de uma solução, de modo que T não é injetora;
- T não é invertível, já que não é injetora e nem sobrejetora. $\lhd$

5.1.1 Caracterizações de matrizes invertíveis

A discussão feita na seção anterior é válida para matrizes quadradas de forma geral e vamos enunciar isto em um grande Teorema. A prova deste resultado abaixo segue as mesmas ideias dos exemplos da seção anterior. Não faremos a prova, mas é importante que o leitor esteja familiarizado com estas ideias.

Teorema 34. *Seja A uma matriz quadrada de ordem $n \times n$. As afirmações abaixo são equivalentes (ou todas falsas ou todas verdadeiras):*

1. *A é invertível;*
2. *Existe uma matriz B tal que $AB = I_n$;*
3. *Existe uma matriz C tal que $CA = I_n$;*
4. *A é equivalente por linhas à matriz identidade I_n;*
5. *A tem n posições de pivô;*
6. *Para todo $\vec{b} \in \mathbb{R}^3$, o sistema $A\vec{x} = \vec{b}$ possui exatamente uma solução;*
7. *O sistema $A\vec{x} = \vec{0}$ possui somente a solução trivial;*
8. *As colunas $\vec{v}_1, \vec{v}_2, \vec{v}_3$ de A são linearmente independentes;*
9. *Todo vetor $\vec{b} \in \mathbb{R}^3$ pode ser escrito como combinação linear das colunas $\vec{v}_1, \vec{v}_2, \vec{v}_3$ de A;*
10. *Todo vetor $\vec{b} \in \mathbb{R}^3$ pertence a $\mathrm{Span}\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$, espaço gerado pelas colunas de A;*
11. *As colunas de A geram $\mathbb{R}^n$: $\mathrm{Span}\{\vec{v}_1, \vec{v}_2, \vec{v}_3\} = \mathbb{R}^3$;*
12. *$T(\vec{x}) = A\vec{x}$ é sobrejetora;*
13. *$T(\vec{x}) = A\vec{x}$ é injetora;*
14. *$T(\vec{x}) = A\vec{x}$ é invertível,*

15. *A^T é invertível.*

Exercício 3. Justificar, baseando-se nos exemplos da seção anterior, o porquê das afirmações acima serem equivalentes.

O teorema é utilizado na prática da mesma maneira que fizemos na seção anterior.

5.2 Espaços vetoriais

Um espaço vetorial (sobre o conjunto $\mathbb{R}$ de escalares) é um conjunto V equipado com as operações de soma de vetores e de multiplicação por escalar e que satisfazem as propriedades usuais dos espaços $\mathbb{R}^n$. A ideia é que vários conjuntos mais abstratos possuem a estrutura parecida com a dos espaços $\mathbb{R}^n$ e esta abordagem permite que façamos uma análise sistemática de todos estes casos.

De forma mais precisa, um **espaço vetorial sobre** $\mathbb{R}$ é um conjunto V, cujos elementos são chamados vetores, equipado com duas operações:

- Soma entre vetores, que satisfaz
  1. Associatividade: $(u + v) + w = u + (v + w)$, para quaisquer $u, v, w \in V$;
  2. Elemento neutro: existe o vetor $0 \in V$ que satisfaz $v + 0 = 0 + v = v$, para qualquer $v \in V$;
  3. Inverso aditivo: para cada $v \in V$, existe o inverso $u = -v \in V$, que satisfaz $v + u = 0$;
  4. Comutatividade: $u + v = v + u$, para quaisquer $u, v \in V$;
- Multiplicação de vetor por escalar, que satisfaz
  5. Associatividade da multiplicação por escalar: $a \cdot (b \cdot v) = (a \cdot b) \cdot v$, para quaisquer $a, b \in \mathbb{R}$ e qualquer $v \in V$;
  6. Vale que $1 \cdot v = v$, ou seja, a unidade dos números reais não altera os vetores de V;
  7. Distributiva de um escalar em relação à soma de vetores: $a \cdot (u + v) = a \cdot v + a \cdot u$, para qualquer $a \in \mathbb{R}$ e quaisquer $u, v \in V$;
  8. Distributiva da soma de escalares em relação a um vetor: $(a + b) \cdot v = a \cdot v + b \cdot v$, para quaisquer $a, b \in \mathbb{R}$ e qualquer $v \in V$.

Exemplo 35. Naturalmente, $\mathbb{R}$, $\mathbb{R}^2$, $\mathbb{R}^3$, ..., $\mathbb{R}^n$ são espaços vetoriais sobre $\mathbb{R}$, já que as propriedades acima foram todas inspiradas nas propriedades válidas para vetores de $\mathbb{R}^n$.

Exemplo 36. O conjunto $M_{m\times n}$ de todas as matrizes de ordem $m \times n$ é um espaço vetorial, com as operações naturais que já utilizamos anteriormente: Se

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \quad \text{e} \quad B = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{bmatrix},$$

definimos

$$A+B = \begin{bmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \cdots & a_{1n}+b_{1n} \\ a_{21}+b_{21} & a_{22}+b_{22} & \cdots & a_{2n}+b_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1}+b_{m1} & a_{m2}+b_{m2} & \cdots & a_{mn}+b_{mn} \end{bmatrix} \quad \text{e} \quad k\cdot A = \begin{bmatrix} k\cdot a_{11} & k\cdot a_{12} & \cdots & k\cdot a_{1n} \\ k\cdot a_{21} & k\cdot a_{22} & \cdots & k\cdot a_{2n} \\ \vdots & \vdots & & \vdots \\ k\cdot a_{m1} & k\cdot a_{m2} & \cdots & k\cdot a_{mn} \end{bmatrix}.$$

É imediato verificar que valem todas as 8 propriedades acima (faça!).

Exemplo 37. O conjunto de todas as funções de um intervalo I em $\mathbb{R}$

$$\mathcal{F}(I;\mathbb{R}) = \{f : I \to \mathbb{R} \text{ funções}\}.$$

forma um espaço vetorial sobre $\mathbb{R}$, com as operações:

Licença CC-BY-SA-3.0. Contato: livroscolaborativos@gmail.com

- Dadas duas funções f e g, a função $f+g: I \to \mathbb{R}$ é definida por
$$\big(f+g\big)(x) = f(x) + g(x).$$
- Dada uma função f e um número real k, a função $k \cdot f: I \to \mathbb{R}$ é definida por
$$\big(k \cdot f\big)(x) = k \cdot f(x).$$

Exercício 4. Justifique que o espaço gerado por dois elementos $\text{Span}\{\vec{u}, \vec{v}\}$ é um espaço vetorial.

5.2.1 Subespaços vetoriais

Consideramos V um espaço vetorial e $H \subseteq V$ um subconjunto. Nós dizemos que H é uma **subespaço vetorial de** V se, além de ser um subconjunto, H é por si só um espaço vetorial.

Na prática, para decidir se um subconjunto H é subespaço de V, não precisamos verificar todas as oito propriedades da definição de espaço vetorial. Temos o seguinte:

Teorema 38 (Critério para subespaços)**.** *Um subconjunto $H \subseteq V$ é subespaço de V se, e somente se, as duas propriedades abaixo são verificadas:*

- $0 \in H$*;*
- $a \cdot u + b \cdot v \in V$ *para quaisquer* $a, b \in \mathbb{R}$ *e quaisquer* $u, v \in V$*.*

Exemplo 39. A reta $H_1 = \{y = 2x\}$ em $\mathbb{R}^2$ é um subespaço de $\mathbb{R}^2$. Temos que

- $\vec{0} = (0, 0)$ pertence a reta, pois satisfaz a equação da reta: $0 = 2 \cdot 0$;
- se (x_1, y_1) e (x_2, y_2) pertencem à reta, vamos verificar que $(ax_1 + bx_2, ay_1 + by_2)$ também pertence:
$$ay_1 + by_2 = a \cdot (2x_1) + b \cdot (2x_2) = 2(ax_1 + bx_2).$$
Portanto, H é um subespaço de $\mathbb{R}^2$.

Exemplo 40. A reta $H_2 = \{y = 2x + 3\}$ não é um subespaço de $\mathbb{R}^2$. De fato, $\vec{0} = (0, 0)$ não pertence a reta. De um modo geral, as retas que passam pela origem são subespaços e as retas que não passam pela origem não são.

Exemplo 41. Mais geralmente, para $\vec{u}, \vec{v} \in \mathbb{R}^n$, o espaço gerado $\text{Span}\{\vec{u}, \vec{v}\}$ é um subespaço de $\mathbb{R}^n$. Caso os vetores sejam LI, este espaço pode ser pensado como o plano que contém os vetores $\vec{u}, \vec{v}$ dentro de $\mathbb{R}^n$ (por quê?).

Poderíamos argumentar que a reta do Exemplo 39 é um subespaço, pois todos os pontos da reta satisfazem:
$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ 2x \end{bmatrix} = x \cdot \begin{bmatrix} 1 \\ 2 \end{bmatrix},$$
isto é, $H = \text{Span}\left\{\begin{bmatrix} 1 \\ 2 \end{bmatrix}\right\}$.

Exemplo 42. Um plano que passa pela origem $2x - y + 3z = 0$ é um subespaço de $\mathbb{R}^3$, pois podemos escrever pontos do plano como
$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x \\ 2x + 3z \\ z \end{bmatrix} = x \cdot \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} + z \cdot \begin{bmatrix} 0 \\ 3 \\ 1 \end{bmatrix},$$
isto é,
$$H = \text{Span}\left\{\begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 3 \\ 1 \end{bmatrix}\right\}$$
Por outro lado, um plano que não passa pela origem, eg, $2x - y + 3z = 1$, não pode ser um subespaço, pois $\vec{0} \notin H$.

Licença CC-BY-SA-3.0. Contato: livroscolaborativos@gmail.com

Exemplo 43. O conjunto $\mathcal{P}(\mathbb{R})$ das funções polinomiais com coeficientes reais é um subespaço do espaço das funções contínuas do Exemplo 37. De fato, o polinômio nulo está neste espaço. Além disso, dados dois polinômios

$$p(x) = a_n x^n + \cdots a_2 x^2 + a_1 x + a_0 \quad \text{e} \quad q(x) = b_m x^m + \cdots b_2 x^2 + b_1 x + b_0,$$

sabemos que $a \cdot p + b \cdot q$ vai ser um polinômio cujo grau é no máximo o maior dos valores m ou n.

Exercício 5. Considere o conjunto

$$\mathcal{P}_3(\mathbb{R}) = \{p(x) = a_3 x^3 + \cdots a_2 x^2 + a_1 x + a_0 \text{ tais que } a_0, a_1, a_2, a_3 \in \mathbb{R}\}$$

o conjunto de todos os polinômios de grau 3. Mostre que:

1. $\mathcal{P}_3(\mathbb{R})$ é um subespaço de $\mathcal{P}(\mathbb{R})$;
2. $\mathcal{P}_3(\mathbb{R})$ é um subespaço de $\mathcal{F}(I; \mathbb{R})$.

É possível generalizar estas conclusões para o conjunto $\mathcal{P}_n(\mathbb{R})$ dos polinômios de grau n?

5.2.2 Espaço nulo

O **espaço nulo** de uma matriz A de ordem $m \times n$ é o conjunto definido por

$$\operatorname{Nul} A = \left\{\vec{x} \in \mathbb{R}^n \mid A\vec{x} = \vec{0}\right\}.$$

Para verificar que $\operatorname{Nul} A$ é um subespaço de $\mathbb{R}^n$, consideramos $\vec{u}, \vec{v} \in \operatorname{Nul} A$ e $a, b \in \mathbb{R}$. Queremos ver que $a\vec{u} + b\vec{v}$ também pertence a $\operatorname{Nul} A$: Pela linearidade do produto de matrizes, temos

$$A(a\vec{u} + b\vec{v}) = a \cdot A\vec{u} + b \cdot A\vec{v} = a \cdot \vec{0} + b \cdot \vec{0} = \vec{0}.$$

Portanto, como também $\vec{0} \in \operatorname{Nul} A$, concluimos que $\operatorname{Nul} A$ é subespaço de $\mathbb{R}^n$.

Exemplo 44 (Retirado do David C. Lay)**.** Encontrar um conjunto gerador para o espaço nulo da matriz

$$\begin{bmatrix} -3 & 6 & -1 & 1 & -7 \\ 1 & -2 & 2 & 3 & -1 \\ 2 & -4 & 5 & 8 & -4 \end{bmatrix}$$

Nós queremos determinar as soluções do sistema linear homogêneo

$$\begin{bmatrix} -3 & 6 & -1 & 1 & -7 \\ 1 & -2 & 2 & 3 & -1 \\ 2 & -4 & 5 & 8 & -4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Poderíamos escrever a matriz associada aumentada e escalonar, mas como este é um sistema homogêneo, não faz diferença escrevermos a última coluna com os zeros ou não (nenhuma operação elementar fará com que desapareçam os zeros). Logo, vamos obter a forma escalonada reduzida da matriz A (contas como exercício):

$$\begin{bmatrix} -3 & 6 & -1 & 1 & -7 \\ 1 & -2 & 2 & 3 & -1 \\ 2 & -4 & 5 & 8 & -4 \end{bmatrix} \sim \begin{bmatrix} 1 & -2 & 0 & -1 & 3 \\ 0 & 0 & 1 & 2 & -2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \sim \begin{cases} x_1 - 2x_2 - x_4 + 3x_5 = 0 \\ x_3 + 2x_4 - 2x_5 = 0 \\ x_2, x_4, x_5 \text{ livres.} \end{cases}$$

Assim, podemos escrever qualquer vetor de $\operatorname{Nul} A$ como

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 2x_2 + x_4 - 3x_5 \\ x_2 \\ -2x_4 + 2x_5 \\ x_4 \\ x_5 \end{bmatrix} = x_2 \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} 1 \\ 0 \\ -2 \\ 1 \\ 0 \end{bmatrix} + x_5 \begin{bmatrix} -3 \\ 0 \\ 2 \\ 0 \\ 1 \end{bmatrix}.$$

Licença CC-BY-SA-3.0. Contato: livroscolaborativos@gmail.com

Isto significa que

$$\operatorname{Nul} A = \operatorname{Span}\left\{ \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 0 \\ 2 \\ 0 \\ 1 \end{bmatrix} \right\}. \triangleleft$$

Uma das importâncias teóricas do espaço nulo é

Proposição 45. *Seja A uma matriz $m \times n$. Então a aplicação linear $\vec{x} \mapsto A\vec{x}$ é injetora se, e somente se,* $\operatorname{Nul} A = \{0\}$.

Em outras palavras, caso o sistema linear homogêneo $A\vec{x} = \vec{0}$ possua somente a solução trivial, podemos inferir que $A\vec{x} = \vec{b}$ possui no máximo uma solução, para qualquer $\vec{b}$!

5.2.3 Espaço coluna

O **espaço coluna** de uma matriz A de ordem $m \times n$ é o espaço gerado pelas colunas de A:

$$A = \begin{bmatrix} | & | & & | \\ \vec{a}_1 & \vec{a}_2 & \cdots & \vec{a}_n \\ | & | & & | \end{bmatrix} \rightsquigarrow \operatorname{Col} A = \operatorname{Span}\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n\}.$$

Exercício 6. Verificar que $\operatorname{Col} A$ é um subespaço de $\mathbb{R}^n$.

Lembrando ainda que

$$x_1\vec{a}_1 + x_2\vec{a}_2 + \cdots + x_n\vec{a}_n = \vec{b} \iff A\vec{x} = \vec{b},$$

podemos reescrever a definição como

$$\operatorname{Col} A = \{\vec{b} \in \mathbb{R}^m \mid \text{existe } \vec{x} \text{ tal que } \vec{b} = A\vec{x}\}.$$

O espaço coluna trata portanto de vetores na imagem da aplicação linear. Temos:

Proposição 46. *Seja A uma matriz $m \times n$. Então a aplicação linear $\vec{x} \mapsto A\vec{x}$ é sobrejetora se, e somente se,* $\operatorname{Col} A = \mathbb{R}^m$.

A definição do espaço coluna já deixa claro um conjunto gerador (as colunas de A!). Não precisamos fazer cálculos como fizemos para encontrar geradores para $\operatorname{Nul} A$.

Licença CC-BY-SA-3.0. Contato: livroscolaborativos@gmail.com

6

Semanas 6 e 7

6.1 Espaços vetoriais

Nós vimos que um espaço vetorial (sobre o conjunto $\mathbb{R}$ de escalares) é um conjunto V equipado com as operações de soma de vetores e de multiplicação por escalar e que satisfazem as propriedades usuais dos espaços $\mathbb{R}^n$.

6.2 Bases de espaços vetoriais

Uma **base** para um espaço vetorial V é um conjunto $\mathcal{B}$ que:

(a) gera V e;

(b) é linearmente independente.

Exemplo 47. $\mathbb{R}^n$ é um espaço vetorial sobre $\mathbb{R}$. Vimos anteriormente que os vetores

$$\vec{e}_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \vec{e}_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \cdots, \vec{e}_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

formam um conjunto linearmente independente. E também é verdade que eles geram $\mathbb{R}^n$, pois qualquer vetor pode ser escrito como combinação linear deles:

$$\vec{v} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix} + \cdots + x_n \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix} = x_1\vec{e}_1 + x_2\vec{e}_2 + \cdots + x_n\vec{e}_n.$$

Portanto, o conjunto de n elementos $\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\}$ é uma base de $\mathbb{R}^n$.

Várias propriedades interessantes aparecem naturalmente como consequência da definição de base. Suponhamos que um conjunto $\mathcal{B} = \{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_d\} \subset V$ seja uma base de um espaço vetorial V. Já que $\operatorname{Span} \mathcal{B} = V$, qualquer elemento de V pode ser representado como

$$\vec{v} = x_1\vec{v}_1 + x_2\vec{v}_2 + \cdots + x_d\vec{v}_d.$$

Além disso, o fato de ser $\mathcal{B}$ um conjunto linearmente independente nos diz que esta forma de representar o vetor $\vec{v}$ é única! De fato, vamos verificar que qualquer representação é, na realidade, a mesma representação que tínhamos acima: se tivermos

$$\vec{v} = k_1\vec{v}_1 + k_2\vec{v}_2 + \cdots + k_d\vec{v}_d,$$

deveremos também ter

$$(x_1 - k_1)\vec{v}_1 + (x_2 - k_2)\vec{v}_2 + \cdots + (x_d - k_d)\vec{v}_d = \vec{v} - \vec{v} = \vec{0}.$$

Sendo a base $\mathcal{B}$ formada por vetores LI, concluimos que os coeficientes da combinação linear acima devem ser nulos, isto é, $x_1 = k_1, x_2 = k_2, \ldots, x_d = k_d$. Portanto,

Todo vetor pode ser representado de maneira única como combinação linear dos elementos de uma base.

Uma segunda propriedade fundamental é a seguinte:

Toda base de um espaço vetorial fixado possui o mesmo número de elementos.

Justificativa. Sejam duas bases $\mathcal{B}_1 = \{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_d\} \subset V$ e $\mathcal{B}_2 = \{\vec{w}_1, \vec{w}_2, \ldots, \vec{w}_k\} \subset V$ do mesmo espaço vetorial V. Nós podemos escrever os elementos da base $\mathcal{B}_1$ como combinações lineares dos elementos da base $\mathcal{B}_2$:

$$\begin{cases} \vec{v}_1 &= a_{11}\vec{w}_1 + a_{21}\vec{w}_2 + \cdots + a_{k1}\vec{w}_k \\ \vec{v}_2 &= a_{12}\vec{w}_1 + a_{22}\vec{w}_2 + \cdots + a_{k2}\vec{w}_k \\ &\vdots \\ \vec{v}_d &= a_{1d}\vec{w}_1 + a_{2d}\vec{w}_2 + \cdots + a_{kd}\vec{w}_k \end{cases}$$

O problema é que, se tivéssemos quantidades diferentes de vetores nas bases, digamos $k < d$, então a equação

$$b_1\vec{v}_1 + b_2\vec{v}_2 + \cdots + b_d\vec{v}_d = \vec{0} \tag{6.1}$$

possuiria soluções não triviais, contradizendo o fato de ser $\mathcal{B}_1$ um conjunto linearmente independente. De fato, utilizando que $\mathcal{B}_1$ é linearmente independente, encontrar $b_1, b_2, \ldots, b_k$ na equação (6.1) fica equivalente a resolver

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1d} \\ a_{21} & a_{22} & \cdots & a_{2d} \\ \vdots & \vdots & & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kd} \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_d \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix},$$

que possui soluções não triviais quando $k < d$.

Argumentando de forma parecida, concluimos também que não podemos ter $d < k$. Portanto, $d = k$. □

O número de elementos de uma base do espaço vetorial V é chamado de **dimensão** de V. Graças à propriedade acima, tanto faz a base que considerarmos, pois todas tem o mesmo número de elementos. A dimensão de V, pelo que vimos até agora, indica o número de parâmetros necessários para representar qualquer vetor de V.

Quando um vetor $\vec{v}$ é escrito em termos dos elementos de uma base $\mathcal{B} = \{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_d\}$, é comum utilizarmos a notação

$$\vec{v} = x_1\vec{v}_1 + x_2\vec{v}_2 + \cdots + x_d\vec{v}_d = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{d-1} \\ x_d \end{bmatrix}_{\mathcal{B}} \text{ ou ainda } \left[\vec{v}\right]_{\mathcal{B}} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{d-1} \\ x_d \end{bmatrix}.$$

Exemplo 48. O conjunto

$$\mathcal{P}_3(\mathbb{R}) = \left\{p(x) = a_3x^3 + a_2x^2 + a_1x + a_0 \text{ tais que } a_0, a_1, a_2, a_3 \in \mathbb{R}\right\}$$

de todos os polinômios de grau 3 é um espaço vetorial. Uma base para este espaço pode ser encontrada "de olho": é fácil ver que o conjunto de polinômios

$$\left\{x^3, x^2, x, 1\right\}$$

gera o espaço $\mathcal{P}_3(\mathbb{R})$. Para ver que são LI, precisamos verificar que se tivermos

$$a_3x^3 + a_2x^2 + a_1x + a_0 = 0, \quad \text{para todo } x \in \mathbb{R},$$

Licença CC-BY-SA-3.0. Contato: livroscolaborativos@gmail.com

então todos os coeficientes devem ser nulos. Isto segue do Teorema Fundamental da Álgebra[1], que diz que todo polinômio não nulo de grau 3 possui exatamente 3 raízes complexas; logo, no máximo 3 raízes reais. Portanto, já que no nosso caso, todos os valores de x reais são raízes, nosso polinômio deve ser o constante igual a zero, de modo que todos os coeficientes são zero.

Concluimos desta maneira que $\mathcal{B} = \{x^3, x^2, x, 1\}$ é uma base de $\mathcal{P}_3(\mathbb{R})$, que consequentemente tem dimensão 4.

Desta maneira, qualquer elemento de $\mathcal{P}_3(\mathbb{R})$ pode ser representado por 4 parâmetros. Por exemplo,

$$p(x) = x^3 - 2x^2 + 3 \ \rightsquigarrow \ [p(x)]_{\mathcal{B}} = \begin{bmatrix} 1 \\ -2 \\ 0 \\ 3 \end{bmatrix},$$

$$q(x) = 2x^3 + x^2 + x - 1 \ \rightsquigarrow \ [q(x)]_{\mathcal{B}} = \begin{bmatrix} 2 \\ 1 \\ 1 \\ -1 \end{bmatrix}.$$

Nota que a ordem que escrevemos os elementos de uma base altera as representações $[p]_{\mathcal{B}}$ e $[q]_{\mathcal{B}}$.

Exemplo 49. Em um curso de equações diferenciais, se aprende que o conjunto de todas as soluções de uma equação linear de segunda ordem da forma

$$y''(t) + f(x)y'(x) + g(x)y(x) = 0$$

é uma espaço vetorial cuja dimensão é 2. Desta forma, por exemplo, sabendo que $y_1(x) = e^{2x}$ e $y_2(x) = e^{-x}$ satisfazem a equação

$$y''(t) - y'(x) - 2y(x) = 0, \tag{6.2}$$

podemos concluir que todas as soluções desta EDO podem ser escritas como combinação linear destas duas:

$$y(x) \text{ é solução de (6.2)} \iff y(x) = C_1 e^{2x} + C_2 e^{-x}, \ \text{ para algumas constantes } C_1, C_2 \in \mathbb{R}.$$

6.2.1 Espaço nulo

Vimos que o espaço nulo (núcleo) de uma matriz A de ordem $m \times n$ é o conjunto definido por

$$\operatorname{Nul} A = \{\vec{x} \in \mathbb{R}^n \mid A\vec{x} = \vec{0}\}.$$

Nas notas da semana passada, vimos também o seguinte exemplo:

Exemplo 50. Encontrar uma base para o espaço nulo da matriz

$$A = \begin{bmatrix} 1 & 0 & 3 & 1 & 2 \\ -1 & 2 & -2 & 0 & 2 \\ 6 & -4 & 1 & 1 & -4 \end{bmatrix}.$$

Por escalonamento:

$$A \sim \begin{bmatrix} 1 & 0 & 3 & 1 & 2 \\ 0 & 2 & 1 & 1 & 4 \\ 0 & 4 & 17 & 5 & 16 \end{bmatrix} \sim \cdots \sim \begin{bmatrix} 1 & 0 & 0 & 2/5 & 2/5 \\ 0 & 1 & 0 & 2/5 & 26/15 \\ 0 & 0 & 1 & 1/5 & 8/15 \end{bmatrix} \sim \begin{cases} x_1 = -\dfrac{2}{5}x_4 - \dfrac{2}{5}x_5 \\ x_2 = -\dfrac{2}{5}x_4 - \dfrac{26}{15}x_5 \\ x_3 = -\dfrac{1}{5}x_4 - \dfrac{8}{15}x_5 \end{cases}$$

[1]Mais geralmente, o Teorema Fundamental da Álgebra afirma que todo polinômio não nulo de grau n possui exatamente n raízes complexas; logo, no máximo n raízes reais.

Licença CC-BY-SA-3.0. Contato: livroscolaborativos@gmail.com

Logo,

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = x_4 \begin{bmatrix} -2/5 \\ -2/5 \\ -1/5 \\ 1 \\ 0 \end{bmatrix} + x_5 \begin{bmatrix} -2/5 \\ -26/15 \\ -8/15 \\ 0 \\ 1 \end{bmatrix}$$

de modo que $\dim \operatorname{Nul} A = 2$ e uma base é

$$\left\{ \begin{bmatrix} -2/5 \\ -2/5 \\ -1/5 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -2/5 \\ -26/15 \\ -8/15 \\ 0 \\ 1 \end{bmatrix} \right\}.$$

Já que estes vetores são LI, temos uma base para $\operatorname{Nul} A$. $\lhd$

É sempre possível obter uma base para $\operatorname{Nul} A$ desta forma. Enfatizamos esta propriedade abaixo:

> Uma base para $\operatorname{Nul} A$ pode ser encontrada por escalonamento ao resolver o sistema homogêneo associado à matriz A. Além disso, a dimensão do espaço nulo é igual ao número de variáveis livres do sistema.

6.2.2 Espaço coluna

Vimos também o espaço coluna de uma matriz A de ordem $m \times n$, que é o espaço gerado pelas colunas de A:

$$A = \begin{bmatrix} | & | & & | \\ \vec{a}_1 & \vec{a}_2 & \cdots & \vec{a}_n \\ | & | & & | \end{bmatrix} \rightsquigarrow \operatorname{Col} A = \operatorname{Span}\{\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n\}.$$

Desta forma, a definição de $\operatorname{Col} A$ já deixa claro um conjunto gerador: as colunas de A! Para determinar uma base, nós devemos "excluir" as colunas que podem ser geradas pelas demais, de modo a obter um conjunto linearmente independente.

> Uma base para $\operatorname{Col} A$ é formada pelas colunas pivô da matriz A. Para descobrir quais são as colunas pivô, procedemos por escalonamento. As *colunas da matriz original* formam uma base de $\operatorname{Col} A$. A dimensão do espaço coluna é igual ao número de colunas pivô da matriz.

Exemplo 51. Encontrar uma base para o espaço coluna da matriz

$$A = \begin{bmatrix} 1 & 0 & 3 & 1 & 2 \\ -1 & 2 & -2 & 0 & 2 \\ 6 & -4 & 1 & 1 & -4 \end{bmatrix}.$$

Vimos acima que a forma escalonada reduzida de A é

$$A \sim \begin{bmatrix} 1 & 0 & 0 & 2/5 & 2/5 \\ 0 & 1 & 0 & 2/5 & 26/15 \\ 0 & 0 & 1 & 1/5 & 8/15 \end{bmatrix},$$

de modo que todas as colunas pivô são as três primeiras. Porntanto, uma base de $\operatorname{Col} A$ é

$$\left\{ \begin{bmatrix} 1 \\ -1 \\ 6 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \\ -4 \end{bmatrix}, \begin{bmatrix} 3 \\ -2 \\ 1 \end{bmatrix} \right\}.$$

Licença CC-BY-SA-3.0. Contato: livroscolaborativos@gmail.com

6.3 Teorema do núcleo e da imagem

Nas seções anteriores, vimos que

$$\dim \operatorname{Col} A = \dim \text{ da imagem de } A = \text{ número de colunas pivô e}$$

$$\dim \operatorname{Nul} A = \text{ número de variáveis livres do sistema homogêneo associado.}$$

Ora, mas ou uma coluna é pivô ou sua variável associada pode ser considerada uma variável livre, de modo que temos o

Teorema 52 (Teorema do núcleo e da imagem)**.** *Para qualquer matriz A de ordem $m \times n$, temos*

$$\boxed{\dim \operatorname{Col} A + \dim \operatorname{Nul} A = n = \text{ número de colunas de } A.}$$

Observação 53. A dimensão do espaço coluna de A ou, o que é a mesma coisa, dimensão da imagem da transformação linear $\vec{x} \mapsto A\vec{x}$, é também conhecida na literatura como o **posto de** A. Assim, o Teorema do núcleo e da imagem pode ser reenunciado como: Se A é uma matriz $m \times n$, então

$$\operatorname{posto} A + \dim \operatorname{Nul} A = n.$$

Observação 54. O teorema afirma que há uma relação entre as dimensões dos espaços coluna e nulo. No entanto, não esqueça que estes são subespaços de universos diferentes! Temos

$$\operatorname{Nul} A \subseteq \mathbb{R}^n \text{ enquanto que } \operatorname{Col} A \subseteq \mathbb{R}^m.$$

Apesar da conclusão do teorema parecer simples, é muito útil na análise de sistemas lineares.

Exemplo 55. Todo sistema linear de ordem 5×9 deve possuir um espaço nulo de dimensão 4 ou mais. De fato, como a matriz associada tem apenas cinco linhas, pode ter no máximo 5 posições de pivô, de modo que

$$\dim \operatorname{Col} A \leq 5.$$

Portanto

$$\dim \operatorname{Nul} A = 9 - \dim \operatorname{Col} A \geq 9 - 5 = 4.$$

Desta forma, se resolvermos um sistema homogêneo desta ordem e encontrarmos um conjunto solução de dimensão 3, certamente há algum um erro nas contas!

6.4 Matriz de mudança de coordenadas

Se conhecemos as coordenadas do vetor $\vec{x}$ em uma base $\mathcal{B} = \{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$:

$$[\vec{x}]_{\mathcal{B}} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}_{\mathcal{B}} = b_1\vec{v}_1 + b_2\vec{v}_2 + \cdots + b_n\vec{v}_n,$$

podemos obter as coordenadas usuais (na base canônica) de $\vec{x}$ de forma sistemática. Em outras palavras, queremos encontrar $x_1, x_2, \ldots, x_n$ tais que

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1\vec{e}_1 + x_2\vec{e}_2 + \cdots + x_n\vec{e}_n.$$

Uma maneira é considerar a matriz cujas colunas são os vetores $\vec{v}_i$:

$$A = \begin{bmatrix} | & | & & | \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_n \\ | & | & & | \end{bmatrix} = \begin{bmatrix} v_{11} & v_{12} & \cdots & v_{1n} \\ v_{21} & v_{22} & \cdots & v_{2n} \\ \vdots & \vdots & & \vdots \\ v_{n1} & v_{n2} & \cdots & v_{nn} \end{bmatrix}.$$

Licença CC-BY-SA-3.0. Contato: livroscolaborativos@gmail.com

Funciona porque

$$b_1\vec{v}_1 + b_2\vec{v}_2 + \cdots + b_n\vec{v}_n = \sum_{j=1}^{n} b_j\vec{v}_j = \sum_{j=1}^{n} b_j \sum_{i=1}^{n} v_{ij}\vec{e}_i = \sum_{i=1}^{n}\left(\sum_{j=1}^{n} b_j v_{ij}\right)\vec{e}_i$$
$$= \left(\sum_{j=1}^{n} b_j v_{1j}\right)\vec{e}_1 + \left(\sum_{j=1}^{n} b_j v_{2j}\right)\vec{e}_2 + \cdots + \left(\sum_{j=1}^{n} b_j v_{nj}\right)\vec{e}_n.$$

Assim, devemos ter

$$\begin{cases} x_1 &=& v_{11}b_1 + v_{12}b_2 + \cdots + v_{1n}b_n \\ x_2 &=& v_{21}b_1 + v_{22}b_2 + \cdots + v_{2n}b_n \\ & \vdots & \\ x_n &=& v_{n1}b_1 + v_{n2}b_2 + \cdots + v_{nn}b_n \end{cases} \quad \text{i.e} \quad \vec{x} = A[\vec{x}]_{\mathcal{B}}.$$

A matriz A construida como acima é chamada de **matriz de mudança de coordenadas** da base $\mathcal{B}$ para a base usual (canônica) $\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\}$.

A matriz de mudança de coordenadas da base canônica para a base $\mathcal{B}$ pode ser obtida a partir da inversa de A:

$$[\vec{x}]_{\mathcal{B}} = A^{-1} \cdot \vec{x}.$$

(Notamos que a matriz A é necessariamente invertível, já que suas colunas são elementos de uma base; em particular, são linearmente independentes).

Exemplo 56. Consideramos a seguinte base para $\mathbb{R}^2$:

$$\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \end{bmatrix} \right\}.$$

Este conjunto é de fato uma base, pois são linearmente independentes e são dois (que é a dimensão de $\mathbb{R}^2$). Sabendo que as coordenadas de $\vec{u}$ e $\vec{v}$ nesta base $\mathcal{B}$ são

$$[u]_{\mathcal{B}} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}_{\mathcal{B}} \quad \text{e} \quad [v]_{\mathcal{B}} = \begin{bmatrix} 3 \\ -1 \end{bmatrix}_{\mathcal{B}},$$

encontrar as componentes de $\vec{u}$ e $\vec{v}$ na base canônica $\{\vec{e}_1, \vec{e}_2\}$ de $\mathbb{R}^2$.

Pelo que analisamos acima, a matriz de mudança de coordenadas que procuramos é a matriz

$$A = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$$

Logo,

$$\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix}_{\mathcal{B}} = \begin{bmatrix} 0 \\ 2 \end{bmatrix} \text{ e } \vec{v} = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 3 \\ -1 \end{bmatrix}_{\mathcal{B}} = \begin{bmatrix} 4 \\ 2 \end{bmatrix}.$$

Observação 57. A matriz de mudança de coordenadas é uma forma de trocar de uma base para outra de forma sistemática. Com o intuito de tornar mais intuitiva a construção, vamos repetir a argumentação que fizemos no início desta seção para o caso 2×2. Na base $\mathcal{B}$ do exemplo anterior, a representação

$$\begin{bmatrix} b_1 \\ b_2 \end{bmatrix}_{\mathcal{B}} \text{ significa que } \vec{v} = b_1 \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix} + b_2 \cdot \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

Os vetores da base $\mathcal{B}$ são representados na base canônica por

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \vec{e}_1 + \vec{e}_2 \quad \text{e} \quad \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} = -\vec{e}_1 + \vec{e}_2.$$

Logo,

$$\vec{v} = b_1(\vec{e}_1 + \vec{e}_2) + b_2(-\vec{e}_1 + \vec{e}_2) = (b_1 - b_2)\vec{e}_1 + (b_1 + b_2)\vec{e}_2,$$

de modo que as componentes de $\vec{v}$ na base canônica são

$$\begin{cases} x_1 = b_1 - b_2 \\ x_2 = b_1 + b_2 \end{cases}.$$

Em notação matricial:

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} b_1 \\ b_2 \end{bmatrix}.$$

Licença CC-BY-SA-3.0. Contato: livroscolaborativos@gmail.com

7

Semana 9

7.1 Determinantes

O determinante é um número que está associado com uma matriz quadrada.[1] Para os nossos propósitos neste curso, o determinante é principalmente utilizado para decidir se uma matriz é invertível. No entanto, o determinante tem outras interpretações. Veremos brevemente como é utilizá-lo no cálculo de volumes de paralelepípedos. Além disso, aparece em aplicações variadas, como a fórmula de mudança de variáveis em integrais múltiplas.

Vejamos inicialmente o caso 2×2. Consideramos a matriz

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

No caso em que ambas as entradas a e c são nulas, sabemos de antemão que A não pode ser uma matriz invertível, pois neste caso sua primeira coluna não possui posição de pivô. Suponhamos que $a \neq 0$ (caso contrário, poderíamos fazer uma troca de linhas). Por eliminação Gaussiana, chegamos a

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \xrightarrow{-\frac{c}{a}\ell_1 + \ell_2 \text{ em } \ell_2} \begin{bmatrix} a & b \\ 0 & -\dfrac{bc}{a} + d \end{bmatrix} = \begin{bmatrix} a & b \\ 0 & \dfrac{ad - bc}{a} \end{bmatrix}.$$

Assim, A é invertível (ou, equivalentemente, as colunas de A são linearmente independentes) se, e somente se, o valor numérico $ad - bc$ é diferente de 0. Esta é a nossa definição de **determinante** para uma matriz A de ordem 2×2:

$$\det A \stackrel{\text{def}}{=} ad - bc.$$

Outras notações bastante utilizadas de determinante são

$$\det \begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad \text{ou} \quad \begin{vmatrix} a & b \\ c & d \end{vmatrix}.$$

Cuidado para não confundir! Com estas notações, $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ representa uma matriz enquanto que $\begin{vmatrix} a & b \\ c & d \end{vmatrix}$ representa um número, o determinante de A.

A nossa discussão acima de imediato implica a seguinte propriedade:

$$\boxed{A \text{ é invertível} \iff \det A \neq 0.}$$

Exemplo 58. A matriz $A = \begin{bmatrix} 1 & 2 \\ 3 & -1 \end{bmatrix}$ é invertível pois $\det A = 1 \cdot (-1) - 3 \cdot 2 = -7 \neq 0$. Em particular, podemos utilizar o Teorema da Matriz Invertível (Teorema ???) para concluir que a

[1]Como afirma ???, surgiu primeiro para sistemas lineares (e não matrizes).

transformação linear associada é invertível (injetiva e sobrejetiva), as colunas são linearmente independentes, espaço nulo tem dimensão zero e o posto é 2, etc.

Por outro lado, $A = \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix}$ não é invertível, já que $\det A = 1 \cdot 6 - 3 \cdot 2 = 0$. $\lhd$

Agora, vamos fazer as contas também no caso 3×3. Considere a matriz

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}.$$

Suponhamos que $a_{11} \neq 0$. Por escalonamento:

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \xrightarrow[-\frac{a_{31}}{a_{11}}\ell_1+\ell_3 \text{ em } \ell_3]{-\frac{a_{21}}{a_{11}}\ell_1+\ell_2 \text{ em } \ell_2} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & \dfrac{a_{11}a_{22} - a_{21}a_{12}}{a_{11}} & \dfrac{a_{11}a_{23} - a_{21}a_{13}}{a_{11}} \\ 0 & \dfrac{a_{11}a_{32} - a_{31}a_{12}}{a_{11}} & \dfrac{a_{11}a_{33} - a_{31}a_{13}}{a_{11}} \end{bmatrix}.$$

Podemos ainda simplificar os denominadores

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \sim \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{11}a_{22} - a_{21}a_{12} & a_{11}a_{23} - a_{21}a_{13} \\ 0 & a_{11}a_{32} - a_{31}a_{12} & a_{11}a_{33} - a_{31}a_{13} \end{bmatrix} \overset{\text{notação}}{===} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & A_{33} & A_{32} \\ 0 & A_{23} & A_{22} \end{bmatrix}.$$

Em breve (esperamos que) ficará clara a escolha da notação acima. O passo seguinte no escalonamento será, supondo que $A_{33} \neq 0$, eliminar o elemento A_{23} que está na posição 32. Temos assim

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \sim \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & A_{33} & A_{32} \\ 0 & A_{23} & A_{22} \end{bmatrix} \sim \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & A_{33} & A_{32} \\ 0 & 0 & A_{22}A_{33} - A_{32}A_{23} \end{bmatrix}. \tag{7.1}$$

Nosso raciocínio é que nossa matriz A é uma matriz invertível se esta última coluna possuir uma posição de pivô, isto é, se $A_{22}A_{33} - A_{32}A_{23} \neq 0$. Este último pode ser escrito mais explicitamente como

$$\begin{aligned} A_{22}A_{33} - A_{32}A_{23} &= (a_{11}a_{22} - a_{21}a_{12})(a_{11}a_{33} - a_{31}a_{13}) - (a_{11}a_{32} - a_{31}a_{12})(a_{11}a_{23} - a_{21}a_{13}) \\ &= a_{11}^2 a_{22}a_{33} - a_{11}a_{22}a_{31}a_{13} - a_{21}a_{12}a_{11}a_{33} + \cancel{a_{21}a_{12}a_{31}a_{13}} + \\ &\quad - a_{11}^2 a_{32}a_{23} + a_{11}a_{32}a_{21}a_{13} + a_{31}a_{12}a_{11}a_{23} - \cancel{a_{31}a_{12}a_{21}a_{13}} \\ &= a_{11}\big(a_{11}a_{22}a_{33} - a_{22}a_{31}a_{13} - a_{21}a_{12}a_{33} - a_{11}a_{32}a_{23} + a_{32}a_{21}a_{13} + a_{31}a_{12}a_{23}\big) \end{aligned}$$

Destas considerações, segue que A é invertível sempre que

$$a_{11}a_{22}a_{33} - a_{22}a_{31}a_{13} - a_{21}a_{12}a_{33} - a_{11}a_{32}a_{23} + a_{32}a_{21}a_{13} + a_{31}a_{12}a_{23} \neq 0.$$

Definimos o **determinante** de uma matriz A de ordem 3×3 por (note que apenas mudamos a ordem dos termos):

$$\det A \overset{\text{def}}{=} a_{11}a_{22}a_{33} - a_{11}a_{32}a_{23} - a_{12}a_{21}a_{33} + a_{12}a_{31}a_{23} + a_{13}a_{21}a_{32} - a_{13}a_{31}a_{22}.$$

Existe uma forma de memorização deste determinante que usualmente é ensinado no ensino médio. No entanto, vamos utilizar um outro método que poderá ser aplicado para matrizes de qualquer ordem!

Observamos que a expressão acima está cheia de simetrias, por exemplo, cada um dos elementos da matriz A aparece exatamente duas vezes. Além disso, aparece uma vez com sinal positivo e outra com sinal negativo. Podemos escrever:

$$\begin{aligned} \det A &= a_{11}\big(a_{22}a_{33} - a_{32}a_{23}\big) - a_{12}\big(a_{21}a_{33} + a_{31}a_{23}\big) + a_{13}\big(a_{21}a_{32} - a_{31}a_{22}\big). \\ &= a_{11} \cdot \det \begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix} - a_{12} \cdot \det \begin{bmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{bmatrix} + a_{13} \cdot \det \begin{bmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} \end{aligned}$$

Licença CC-BY-SA-3.0. Contato: livroscolaborativos@gmail.com

Esta última fórmula (que é apenas uma outra forma de escrever a nossa definição de determinante de uma matriz de ordem 3×3), apesar de aparentemente complicada, nos permite entender como que os coeficientes de uma matriz aparecem na definição de $\det A$. Vamos escrever novamente:

$$\det \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = a_{11} \cdot \det \begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix} - a_{12} \cdot \det \begin{bmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{bmatrix} + a_{13} \cdot \det \begin{bmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix}.$$

Podemos pensar como segue:

- Nós vamos percorrer a primeira linha da esquerda para a direita, alternando o sinal e multiplicando por determinantes menores.
- O primeiro elemento é o elemento da primeira linha é a_{11}. *Não alteramos* o sinal e multiplicamos por um **determinante menor**, obtido ao desconsiderar a primeira linha e a primeira coluna (ou, em outras palavras, a linha e a coluna do elemento a_{11}):

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \rightsquigarrow \begin{bmatrix} \square & \square & \square \\ \square & a_{22} & a_{23} \\ \square & a_{32} & a_{33} \end{bmatrix} \rightsquigarrow A_{11} \stackrel{\text{def}}{=} \begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix}.$$

  Denotamos por A_{11} a matriz obtida ao remover a linha e a coluna no elemento a_{11}.
- Em seguida, vamos para o segundo elemento da primeira linha, que é a_{12}. *Alteramos* o sinal e multiplicamos pelo determinante menor da matriz A_{12}, obtida de A ao eliminar a linha e a coluna de a_{12}:

$$\begin{bmatrix} \square & \square & \square \\ a_{21} & \square & a_{23} \\ a_{31} & \square & a_{33} \end{bmatrix} \rightsquigarrow A_{12} \stackrel{\text{def}}{=} \begin{bmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{bmatrix}.$$

- Finalmente consideramos a_{31}. *Não alteramos* o sinal e multiplicamos pelo determinante menor da matriz A_{13}, obtida de A ao eliminar a linha e a coluna de a_{13}:

$$\begin{bmatrix} \square & \square & \square \\ a_{21} & a_{22} & \square \\ a_{31} & a_{32} & \square \end{bmatrix} \rightsquigarrow A_{13} \stackrel{\text{def}}{=} \begin{bmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix}.$$

- Podemos então escrever

$$\boxed{\det A = a_{11} \det A_{11} - a_{12} \det A_{12} + a_{13} \det A_{13}.} \tag{7.2}$$

Exemplo 59. Calcular o determinante de

$$A = \begin{bmatrix} 2 & 4 & 3 \\ 1 & 2 & -1 \\ 0 & 2 & 1 \end{bmatrix}.$$

A notação de "barrinhas" para o determinante é particularmente adequada para escrever o determinante como aparece na fórmula (7.2), pois assim podemos ir mentalmente desconsiderando (ou tapando com um lápis) as linhas e colunas que não devemos escrever (indentifique que tudo o que fizemos foi escrever a fórmula (7.2)):

$$\det A = \begin{vmatrix} 2 & 4 & 3 \\ 1 & 2 & -1 \\ 0 & 2 & 1 \end{vmatrix} = 2 \cdot \begin{vmatrix} 2 & -1 \\ 2 & 1 \end{vmatrix} - 4 \cdot \begin{vmatrix} 1 & -1 \\ 0 & 1 \end{vmatrix} + 3 \cdot \begin{vmatrix} 1 & 2 \\ 0 & 2 \end{vmatrix}.$$

Agora, já sabemos como calcular determinantes de matrizes 2×2, que é o que nos resta fazer:

$$\begin{aligned} \begin{vmatrix} 2 & 4 & 3 \\ 1 & 2 & -1 \\ 0 & 2 & 1 \end{vmatrix} &= 2\big(2 \cdot 1 - 2 \cdot (-1)\big) - 4\big(1 \cdot 1 - 0 \cdot (-1)\big) + 3\big(1 \cdot 2 - 0 \cdot 2\big) \\ &= 2 \cdot 4 - 4 \cdot 1 + 3 \cdot 2 = 10. \ \triangleleft \end{aligned}$$

Exemplo 60. Calcular o determinante de

$$B = \begin{bmatrix} 1 & -3 & -4 \\ 1 & 0 & -1 \\ 0 & 2 & 1 \end{bmatrix}.$$

A notação de "barrinhas" para o determinante é particularmente adequada para escrever o determinante como aparece na fórmula (7.2), pois assim podemos ir mentalmente desconsiderando (ou tapando com um lápis) as linhas e colunas que não devemos escrever (indentifique que tudo o que fizemos foi escrever a fórmula (7.2)):

$$\begin{aligned} \det B = \begin{vmatrix} 1 & -3 & -4 \\ 1 & 0 & -1 \\ 0 & 2 & 1 \end{vmatrix} &= \begin{vmatrix} 0 & -1 \\ 2 & 1 \end{vmatrix} - (-3) \cdot \begin{vmatrix} 1 & -1 \\ 0 & 1 \end{vmatrix} + (-4) \cdot \begin{vmatrix} 1 & 0 \\ 0 & 2 \end{vmatrix} \\ &= 2 + 3 \cdot 1 - 4 \cdot 2 = -3. \end{aligned}$$

Analise com atenção como os sinais alternam, independentemente dos sinais dos coeficientes da matriz! ◁

Como já mencionamos, na fórmula para o determinante de uma matriz A, cada um dos elementos de A aparece exatamente duas vezes. Isto significa que poderíamos ter rearranjado os termos da matriz não a partir da primeira linha, mas a partir de *qualquer linha ou qualquer coluna*. Mas devemos ter cuidado para que os sinais sejam levados em consideração de forma coerente.

Dada uma matriz quadrada A de ordem $n \times n$, obtemos uma **matriz menor** A_{ij} ao remover a linha i e coluna j. Agora é possível entender a notação escolhida no inicio deste capítulo, na fórmula (7.1). Definimos o **cofator** (i, j) de A por

$$C_{ij} \stackrel{\text{def}}{=} (-1)^{i+j} \det A_{ij}.$$

Este sinal ± 1 na definição do cofator é o que faz com que o sinal seja levado em consideração corretamente.

Teorema 61. *Podemos calcular o determinante de A a partir de qualquer linha ou de qualquer coluna. Mais precisamente:*

- *Se consideramos a linha i, então*

$$\det A = a_{i1}C_{i1} + a_{i2}C_{i2} + a_{i3}C_{i3}.$$

- *Se consideramos a coluna j, então*

$$\det A = a_{1j}C_{1j} + a_{2j}C_{2j} + a_{3j}C_{3j}.$$

É prático de calcular o determinante pensando como vínhamos fazendo antes, alternando os sinais. Isto é possível de fazer utilizando qualquer linha ou qualquer coluna. Basta descobrirmos com qual sinal devemos começar. Construimos uma "matriz" com os sinais que cada posição da matriz nos dá. Vamos colocar o sinal de $(-1)^{i+j}$ na posição ij da matriz:

$$\begin{bmatrix} (-1)^{1+1} & (-1)^{1+2} & (-1)^{1+3} \\ (-1)^{2+1} & (-1)^{2+2} & (-1)^{2+3} \\ (-1)^{3+1} & (-1)^{3+2} & (-1)^{3+3} \end{bmatrix} \leftrightsquigarrow \begin{bmatrix} + & - & + \\ - & + & - \\ + & - & + \end{bmatrix}.$$

Claro que poderíamos fazer esta matriz de sinais para matrizes de qualquer ordem.

Exemplo 62. Vamos calcular de várias maneiras o determinante da matriz

$$A = \begin{bmatrix} -1 & 1 & 4 \\ 3 & 0 & -1 \\ 1 & 0 & 3 \end{bmatrix}.$$

Licença CC-BY-SA-3.0. Contato: livroscolaborativos@gmail.com

Pela nossa definição

$$\begin{vmatrix} -1 & 1 & 4 \\ 3 & 0 & -1 \\ 1 & 0 & 3 \end{vmatrix} = -1 \begin{vmatrix} 0 & -1 \\ 0 & 3 \end{vmatrix} - 1 \begin{vmatrix} 3 & -1 \\ 1 & 3 \end{vmatrix} + 4 \begin{vmatrix} 3 & 0 \\ 1 & 0 \end{vmatrix} = (-1) \cdot 0 - 1(9+1) + 4 \cdot 0 = -10.$$

Uma boa escolha seria uma linha ou coluna que tenha o maior número de zeros! Pois assim, economizamos tanto nos cálculos quanto na escrita. Por exemplo, escolhemos a segunda coluna. Para saber o sinal adequado, podemos proceder da seguinte maneira: começando na posição 11 com o sinal "$+$", vamos alternando o sinal até completar a segunda coluna:

$$\begin{bmatrix} + & & \\ & & \\ & & \end{bmatrix} \rightsquigarrow \begin{bmatrix} + & - & \\ & & \\ & & \end{bmatrix} \rightsquigarrow \begin{bmatrix} + & - & \\ & + & \\ & & \end{bmatrix} \rightsquigarrow \begin{bmatrix} + & - & \\ & + & \\ & - & \end{bmatrix} \rightsquigarrow \text{sinais segunda coluna são } \begin{bmatrix} - \\ + \\ - \end{bmatrix}.$$

Assim, podemos calcular

$$\begin{vmatrix} -1 & 1 & 4 \\ 3 & 0 & -1 \\ 1 & 0 & 3 \end{vmatrix} = -1 \cdot \begin{vmatrix} 3 & -1 \\ 1 & 3 \end{vmatrix} + 0 - 0 = -1(9+1) = -10.$$

Observe que nem escrevemos as determinantes menores que estão multiplicados por zero. De fato, nem precisaríamos ter escrito os zeros, apenas o fizemos para exemplificar os sinais alternando de forma correta.

A segunda coluna, neste caso, era a melhor escolha para o cálculo do determinante, pois apenas um elemento é não nulo. De qualquer maneira, para praticar, vamos calcular ainda mais uma vez $\det A$, agora utilizando a terceira linha. Sinais que aparecem na frente dos coeficientes, de acordo com a terceira linha:

$$\begin{bmatrix} + & & \\ - & & \\ + & - & + \end{bmatrix} \rightsquigarrow \text{sinais terceira linha são } \begin{bmatrix} + & - & + \end{bmatrix}.$$

Logo,

$$\begin{vmatrix} -1 & 1 & 4 \\ 3 & 0 & -1 \\ 1 & 0 & 3 \end{vmatrix} = 1 \cdot \begin{vmatrix} 1 & 4 \\ 0 & -1 \end{vmatrix} - 0 + 3 \cdot \begin{vmatrix} -1 & 1 \\ 3 & 0 \end{vmatrix} = 1 \cdot (-1) + 3 \cdot (-3) = -10.$$

Como exercício, calcule o determinante utilizando outras linhas ou colunas. $\lhd$

7.2 Determinantes de matrizes de ordem maior

Seja A uma matriz quadrada, de ordem $n \times n$. O **determinante** de A é definido recursivamente:

$$\boxed{\det A = a_{11} \det A_{11} - a_{12} \det A_{12} + a_{13} \det A_{13} - a_{14} \det A_{14} + \cdots + (-1)^{1+n} a_{1n} \det A_{1n}.} \tag{7.3}$$

ou, na notação dos cofatores:

$$\boxed{\det A = a_{11}C_{11} + a_{12}C_{12} + a_{13}C_{13} + a_{14}C_{14} + \cdots + a_{1n}C_{1n}.}$$

Nossa definição é de fato recorrente, pois para calcular $\det A$, de acordo com a definição, nós precisaremos calcular vários determinantes de ordem $n-1$. Estes por sua vez, consistem de vários determinante de ordem $n-2$, e assim por diante. Isto implica, em particular, que o cálculo de determinantes é, em geral, uma tarefa bastante trabalhosa.

Assim como na seção anterior, podemos utilizar qualquer linha ou coluna desde que com os sinais corretos:

Teorema 63. *Podemos calcular o determinante de A a partir de qualquer linha ou de qualquer coluna. Mais precisamente:*

Licença CC-BY-SA-3.0. Contato: livroscolaborativos@gmail.com

- *Se consideramos a linha i, então*
$$\det A = a_{i1}C_{i1} + a_{i2}C_{i2} + a_{i3}C_{i3} + \cdots + a_{in}C_{in}.$$
- *Se consideramos a coluna j, então*
$$\det A = a_{1j}C_{1j} + a_{2j}C_{2j} + a_{3j}C_{3j} + \cdots + a_{nj}C_{nj}.$$

Exemplo 64. Calcular o determinante da matriz de ordem 4×4:

$$A = \begin{bmatrix} -2 & 3 & 0 & 4 \\ 1 & 0 & -1 & 0 \\ 3 & 2 & 1 & 1 \\ -2 & 2 & 0 & 1 \end{bmatrix}.$$

Na tentativa de evitar muitas contas, vamos escolher para começar, uma linha ou coluna que possua o menor número de zeros possível. Neste caso, poderia ser a segunda linha ou a terceira coluna. Vamos escolher a segunda linha (calcule, como exercício, o determinante utilizando a terceira coluna). Os sinais são:

$$\begin{bmatrix} + & & & \\ - & + & - & + \\ & & & \\ & & & \end{bmatrix} \rightsquigarrow \text{ sinais segunda linha são } \begin{bmatrix} - & + & - & + \end{bmatrix}.$$

Logo,

$$\begin{vmatrix} -2 & 3 & 0 & 4 \\ 1 & 0 & -1 & 0 \\ 3 & 2 & 1 & 1 \\ -2 & 2 & 0 & 1 \end{vmatrix} = -1 \cdot \begin{vmatrix} 3 & 0 & 4 \\ 2 & 1 & 1 \\ 2 & 0 & 1 \end{vmatrix} + 0 - (-1) \cdot \begin{vmatrix} -2 & 3 & 4 \\ 3 & 2 & 1 \\ -2 & 2 & 1 \end{vmatrix} + 0.$$

Perceba que os dois zeros evitaram que calculássemos dois determinantes de ordem 3. Em seguida, calculamos cada um dos determinantes de ordem 3 (no primeiro deles, escolhemos a segunda coluna, por possuir dois zeros; no segundo, qualquer escolha seria parecida, já que a matriz não tem entradas nulas – escolhemos a terceira coluna):

$$\begin{aligned} \begin{vmatrix} -2 & 3 & 0 & 4 \\ 1 & 0 & -1 & 0 \\ 3 & 2 & 1 & 1 \\ -2 & 2 & 0 & 1 \end{vmatrix} &= - \begin{vmatrix} 3 & 0 & 4 \\ 2 & 1 & 1 \\ 2 & 0 & 1 \end{vmatrix} + \begin{vmatrix} -2 & 3 & 4 \\ 3 & 2 & 1 \\ -2 & 2 & 1 \end{vmatrix} \\ &= -1 \cdot \begin{vmatrix} 3 & 4 \\ 2 & 1 \end{vmatrix} + \left(4 \cdot \begin{vmatrix} 3 & 2 \\ -2 & 2 \end{vmatrix} - 1 \cdot \begin{vmatrix} -2 & 3 \\ -2 & 2 \end{vmatrix} + 1 \cdot \begin{vmatrix} -2 & 3 \\ 3 & 2 \end{vmatrix} \right) \\ &= (-1) \cdot (-5) + \Big(4 \cdot 10 - 1 \cdot 2 + 1 \cdot (-13) \Big) = 30. \ \lhd \end{aligned}$$

Exemplo 65. Calcular o determinante da matriz de ordem 5×5:

$$A = \begin{bmatrix} 2 & 0 & 0 & 8 & 0 \\ 1 & -7 & -5 & 0 & 0 \\ 3 & 8 & 6 & 0 & 0 \\ 0 & 7 & 5 & 4 & 0 \\ 2 & 3 & 1 & 1 & 1 \end{bmatrix}.$$

Começamos pela última coluna, pois esta possui apenas uma entrada não nula. Assim, nosso determinante já é reduzido a calcular apenas um determinante de ordem 4×4 (em contraste com calcular cinco determinantes 4×4).

Análise dos sinais da quinta coluna:

$$\begin{bmatrix} + & - & + & - & + \\ & & & & - \\ & & & & + \\ & & & & - \\ & & & & + \end{bmatrix} \rightsquigarrow \text{ sinais quinta coluna são } \begin{bmatrix} + \\ - \\ + \\ - \\ + \end{bmatrix}.$$

Licença CC-BY-SA-3.0. Contato: livroscolaborativos@gmail.com

Assim:

$$\begin{vmatrix} 2 & 0 & 0 & 8 & 0 \\ 1 & -7 & -5 & 0 & 0 \\ 3 & 8 & 6 & 0 & 0 \\ 0 & 7 & 5 & 4 & 0 \\ 2 & 3 & 1 & 1 & 1 \end{vmatrix} = 0 - 0 + 0 - 0 + 1 \cdot \begin{vmatrix} 2 & 0 & 0 & 8 \\ 1 & -7 & -5 & 0 \\ 3 & 8 & 6 & 0 \\ 0 & 7 & 5 & 4 \end{vmatrix} = \begin{vmatrix} 2 & 0 & 0 & 8 \\ 1 & -7 & -5 & 0 \\ 3 & 8 & 6 & 0 \\ 0 & 7 & 5 & 4 \end{vmatrix}.$$

Em seguida, escolhemos (por exemplo) a primeira linha da nova matriz 4×4:

$$\det A = \begin{vmatrix} 2 & 0 & 0 & 8 \\ 1 & -7 & -5 & 0 \\ 3 & 8 & 6 & 0 \\ 0 & 7 & 5 & 4 \end{vmatrix} = 2 \cdot \begin{vmatrix} -7 & -5 & 0 \\ 8 & 6 & 0 \\ 7 & 5 & 4 \end{vmatrix} - 8 \cdot \begin{vmatrix} 1 & -7 & -5 \\ 3 & 8 & 6 \\ 0 & 7 & 5 \end{vmatrix}.$$

Finalmente, temos dois determinantes de matrizes de ordem 3×3 para calcular:

$$\begin{aligned} \det A &= 2 \cdot 4 \cdot \begin{vmatrix} -7 & -5 \\ 8 & 6 \end{vmatrix} - 8 \left(\begin{vmatrix} 8 & 6 \\ 7 & 5 \end{vmatrix} - 3 \cdot \begin{vmatrix} -7 & -5 \\ 7 & 5 \end{vmatrix} \right) \\ &= 8 \cdot (-2) - 8(-2 + -3 \cdot 0) = -16 + 16 = 0. \lhd \end{aligned}$$

Estes exemplos já devem deixar claro que o cálculo de determinantes é demasiado trabalhoso, exceto em alguns casos que a matriz tem muitas entradas nulas. Veremos nas próximas seções algumas propriedades e aplicações.

7.3 Propriedades do determinante

Nesta seção, vamos apontar as principais propriedades do determinante. A prova rigorosa destas propriedades será adiada para o apêndice desta seção.

Teorema 66. *Valem as seguintes propriedades:*

(i) *O determinante depende linearmente de cada uma das linhas, isto é, se fizermos uma combinação linear de uma linha apenas, poderíamos ter feito uma cobinação linear dos determinantes:*

$$\det \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ \alpha a_{i1} + \beta b_{i1} & \cdots & \alpha a_{in} + \beta b_{in} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} = \alpha \det \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{i1} & \cdots & a_{in} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} + \beta \det \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ b_{i1} & \cdots & b_{in} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}.$$

(ii) *Se uma linha de A for composta só por zeros, então* $\det A = 0$*.*

(iii) *O determinante de uma matriz triangular é igual ao produto dos elementos da diagonal principal.*

(iv) *A operação elementar "trocar duas linhas de lugar" altera o sinal do determinante.*

(v) *A operação elementar de somar o múltiplo de uma linha à outra não altera o determinante. Em outras palavras, se um múltiplo de uma linha de A for somado à outra linha formando a matriz B, então* $\det A = \det B$*.*

(vi) *Uma matriz A é invertível se, e somente se,* $\det A \neq 0$.

(vii) *Para quaisquer duas matrizes A e B de mesma ordem,* $\det(AB) = \det A \det B$.

$(viii)$ *O determinante da matriz transposta de A é igual ao determinante de A, isto é,* $\det(A^T) = \det A$.

(ix) *Todos os itens acima que envolvem operações com linhas poderiam ser enunciados com "colunas" no lugar de "linhas".*

Licença CC-BY-SA-3.0. Contato: livroscolaborativos@gmail.com

Várias destas propriedades já devem ser familiares para os leitores deste livro, com a possível exceção do comportamento do determinante com as operações elementares de escalonamento. E estas vão ser muito úteis no cálculo do determinante. Vamos enfatizar estas propriedades abaixo. O método para calcular o determinante por escalonamento, segue o seguinte raciocínio:

- Caso seja necessário uma troca de linhas para que a posição de pivô fique com um elemento não nulo, somos permitidos de fazer a troca, *desde que alterando o sinal do determinante*, como nos diz a propriedade (iv) acima;
- De acordo com a propriedade (v), eliminar os elementos abaixo da posição de pivô não altera o determinante. **Um cuidado**: multiplicar linhas por escalares altera o determinante! Desta maneira, esta operação elementar significa estritamente fazer uma operação do tipo
$$k\ell_i + \ell_j \text{ em } \ell_j.$$
Observe que "adicionamos um múltiplo da linha i na linha j". Atentem para o fato de que não pode haver coeficiente diferente de 1 em ℓ_j.
- Caso queiramos, para simplificar as contas, multiplicar ou dividir uma linha por um fator qualquer, podemos fazer uma aplicação cuidadosa da linearidade enunciada na propriedade (i) acima. Considerando $\beta = 0$, esta propriedade se transforma em "colocar um fator α de uma linha em evidência":
$$\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ \alpha a_{i1} & \cdots & \alpha a_{in} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} = \alpha \cdot \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{i1} & \cdots & a_{in} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}.$$

Exemplo 67. Vamos calcular o determinante da matriz A do Exemplo 64 utilizando as propriedades acima (em particular o escalonamento). Este método é particularmente útil quando as matrizes não possuem muitas entradas nulas. Já que a segunda linha possui um "1" na primeira entrada, vamos fazer uma troca de linhas para facilitar as contas (cuidado com o sinal!). Em seguida, eliminamos os elementos da primeira coluna.

$$\begin{vmatrix} -2 & 3 & 0 & 4 \\ 1 & 0 & -1 & 0 \\ 3 & 2 & 1 & 1 \\ -2 & 2 & 0 & 1 \end{vmatrix} = - \begin{vmatrix} 1 & 0 & -1 & 0 \\ -2 & 3 & 0 & 4 \\ 3 & 2 & 1 & 1 \\ -2 & 2 & 0 & 1 \end{vmatrix} = - \begin{vmatrix} 1 & 0 & -1 & 0 \\ 0 & 3 & -2 & 4 \\ 0 & 2 & 4 & 1 \\ 0 & 2 & -2 & 1 \end{vmatrix} = - \begin{vmatrix} 1 & 0 & -1 & 0 \\ 0 & 3 & -2 & 4 \\ 0 & 0 & 16/3 & -5/3 \\ 0 & 0 & -2/3 & -5/3 \end{vmatrix}.$$

As divisões nos denominadores nos atrapalham um pouco na hora de fazer a conta. Podemos retirá-los dali, desde que cuidadosamente (o mesmo para o sinal de "-1"), colocando-os em evidência (note que devemos fazê-lo para cada linha!):

$$\begin{vmatrix} -2 & 3 & 0 & 4 \\ 1 & 0 & -1 & 0 \\ 3 & 2 & 1 & 1 \\ -2 & 2 & 0 & 1 \end{vmatrix} = -\frac{1}{3} \begin{vmatrix} 1 & 0 & -1 & 0 \\ 0 & 3 & -2 & 4 \\ 0 & 0 & 16 & -5 \\ 0 & 0 & -2/3 & -5/3 \end{vmatrix} = \frac{1}{9} \begin{vmatrix} 1 & 0 & -1 & 0 \\ 0 & 3 & -2 & 4 \\ 0 & 0 & 16 & -5 \\ 0 & 0 & 2 & 5 \end{vmatrix}$$

Finalmente, podemos fazer uma troca de linhas e eliminar o elemento "16":

$$\begin{vmatrix} -2 & 3 & 0 & 4 \\ 1 & 0 & -1 & 0 \\ 3 & 2 & 1 & 1 \\ -2 & 2 & 0 & 1 \end{vmatrix} = -\frac{1}{9} \begin{vmatrix} 1 & 0 & -1 & 0 \\ 0 & 3 & -2 & 4 \\ 0 & 0 & 2 & 5 \\ 0 & 0 & 0 & -45 \end{vmatrix} = -\frac{1}{9} \cdot 1 \cdot 3 \cdot 2 \cdot (-45) = 30. \lhd$$

Observação 68. No exemplo anterior, vimos como calcular o determinante utilizando escalonamento de maneira "straightforward". Como pode-se perceber, o método não parece muito melhor do que calcular o determinante utilizando expansão por cofatores. Vamos ver que, de

Licença CC-BY-SA-3.0. Contato: livroscolaborativos@gmail.com

fato, o melhor é *misturar* o método de cofatores com as propriedades acima! Vamos novamente calcular

$$\begin{vmatrix} -2 & 3 & 0 & 4 \\ 1 & 0 & -1 & 0 \\ 3 & 2 & 1 & 1 \\ -2 & 2 & 0 & 1 \end{vmatrix}.$$

Observe que a terceira coluna tem duas entradas nulas. Podemos ainda utilizar uma operação elementar para eliminar uma das entradas: por exemplo, substituir $\ell_3 + \ell_2$ em ℓ_2. Assim:

$$\begin{vmatrix} -2 & 3 & 0 & 4 \\ 1 & 0 & -1 & 0 \\ 3 & 2 & 1 & 1 \\ -2 & 2 & 0 & 1 \end{vmatrix} = \begin{vmatrix} -2 & 3 & 0 & 4 \\ 4 & 2 & 0 & 1 \\ 3 & 2 & 1 & 1 \\ -2 & 2 & 0 & 1 \end{vmatrix}. \qquad \left(\text{lembrando sinais} \begin{bmatrix} + & - & + & \\ & & - & \\ & & + & \\ & & - & \end{bmatrix} \right)$$

Note que a terceira coluna agora ficou com apenas uma entrada não nula; logo, utilizando a terceira coluna para o cálculo de $\det A$ (como na seção anterior), obtemos

$$\begin{vmatrix} -2 & 3 & 0 & 4 \\ 1 & 0 & -1 & 0 \\ 3 & 2 & 1 & 1 \\ -2 & 2 & 0 & 1 \end{vmatrix} = \begin{vmatrix} -2 & 3 & 4 \\ 4 & 2 & 1 \\ -2 & 2 & 1 \end{vmatrix}$$

Podemos também utilizar a propriedade (ix) do teorema para colocar em evidência um "-2" da primeira coluna e, em seguida, continuar com o cálculo:

$$\begin{vmatrix} -2 & 3 & 0 & 4 \\ 1 & 0 & -1 & 0 \\ 3 & 2 & 1 & 1 \\ -2 & 2 & 0 & 1 \end{vmatrix} = -2 \cdot \begin{vmatrix} 1 & 3 & 4 \\ -2 & 2 & 1 \\ 1 & 2 & 1 \end{vmatrix} = -2 \cdot \begin{vmatrix} 1 & 3 & 4 \\ 0 & 8 & 9 \\ 0 & -1 & -3 \end{vmatrix} = -2 \cdot \begin{vmatrix} 1 & 3 & 4 \\ 0 & 1 & 3 \\ 0 & 0 & -15 \end{vmatrix} = -2 \cdot (-15) = 30. \lhd$$

Exemplo 69. Por fim, recalculamos também o determinante da matriz do Exemplo 65, utilizando as propriedades desta seção. Vamos fazer as contas de forma um pouco mais rápidas. Tente acompanhar o que está sendo feito de um passo para outro! Iniciamos aproveitando o fato de a última coluna ter muitos zeros.

$$\begin{vmatrix} 2 & 0 & 0 & 8 & 0 \\ 1 & -7 & -5 & 0 & 0 \\ 3 & 8 & 6 & 0 & 0 \\ 0 & 7 & 5 & 4 & 0 \\ 2 & 3 & 1 & 1 & 1 \end{vmatrix} = \begin{vmatrix} 2 & 0 & 0 & 8 \\ 1 & -7 & -5 & 0 \\ 3 & 8 & 6 & 0 \\ 0 & 7 & 5 & 4 \end{vmatrix} \overset{-2\ell_4+\ell_1 \text{ em } \ell_1}{=} \begin{vmatrix} 2 & -14 & -10 & 0 \\ 1 & -7 & -5 & 0 \\ 3 & 8 & 6 & 0 \\ 0 & 7 & 5 & 4 \end{vmatrix} = 4 \cdot \begin{vmatrix} 2 & -14 & -10 \\ 1 & -7 & -5 \\ 3 & 8 & 6 \end{vmatrix}.$$

Em seguida, eliminamos o 2 e o 3 da primeira coluna sem trocar linhas de lugar (quanto mais trocarmos linhas, mais riscos corremos de errar o sinal):

$$\begin{vmatrix} 2 & 0 & 0 & 8 & 0 \\ 1 & -7 & -5 & 0 & 0 \\ 3 & 8 & 6 & 0 & 0 \\ 0 & 7 & 5 & 4 & 0 \\ 2 & 3 & 1 & 1 & 1 \end{vmatrix} = 4 \cdot \begin{vmatrix} 0 & 0 & 0 \\ 1 & -7 & -5 \\ 0 & 29 & -9 \end{vmatrix} 4 \cdot 0 = 0.$$

Nota: antes de eliminarmos o 3, já reparamos que o determinante vai ser nulo, graças à propriedade (ii). $\lhd$

7.3.1 Demonstração das propriedades do Teorema 66

Em construção.

Licença CC-BY-SA-3.0. Contato: livroscolaborativos@gmail.com

7.4 Uma aplicação em cálculo de várias variáveis

Como já é conhecido de primeiros cursos de cálculo[2], é possível fazer mudanças de coordenadas em integrais unidimensionais (também conhecido como integrar por substituição): para $I \subseteq \mathbb{R}$ um intervalo e $\phi : [a, b] \to I$ é diferenciável com derivada contínua e $f : I \to \mathbb{R}$ é uma função contínua, então

$$\int_{\phi(a)}^{\phi(b)} f(x)\ \mathrm{d}x = \int_a^b f\big(\phi(t)\big)\phi'(t)\ \mathrm{d}t.$$

Intuitivamente, pensamos na substituição $x = \phi(t) \implies \mathrm{d}x = \phi'(t)\,\mathrm{d}t$.

Um tópico que nem sempre é abordado em cursos de Cálculo em várias variáveis é a fórmula de mudança de variáveis para integrais múltiplas, onde o determinante aparece de forma fundamental.

Teorema 70. *Seja $U \subset \mathbb{R}^2$ um conjunto aberto e $\phi : U \to \mathbb{R}^n$ uma função vetorial cujas componentes são todas diferenciáveis e com derivadas parciais contínuas. Então, para qualquer $f : U \to \mathbb{R}$ contínua, vale a fórmula de mudança de variáveis:*

$$\iiint_{\phi(U)} f(\vec{v})\ \mathrm{d}\vec{v} = \iiint_U f\big(\varphi(\vec{u})\big) J\varphi(\vec{u})\ \mathrm{d}\vec{u},$$

onde $J\varphi(\vec{u})$ é o ***Jacobiano****: valor absoluto do determinante da matriz formada com as derivadas parciais de $\phi = (\phi_1, \phi_2, \ldots, \phi_n)$:*

$$J\varphi(\vec{u}) = \left|\det\left(\frac{\partial \phi_i}{\partial x_j}\right)\right| = \left|\det\begin{bmatrix} \dfrac{\partial \phi_1}{\partial x_1} & \dfrac{\partial \phi_1}{\partial x_2} & \cdots & \dfrac{\partial \phi_1}{\partial x_n} \\ \dfrac{\partial \phi_2}{\partial x_1} & \dfrac{\partial \phi_2}{\partial x_2} & \cdots & \dfrac{\partial \phi_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial \phi_n}{\partial x_1} & \dfrac{\partial \phi_n}{\partial x_2} & \cdots & \dfrac{\partial \phi_n}{\partial x_n} \end{bmatrix}\right|$$

7.4.1 Coordenadas polares

Para uma região R do plano $\mathbb{R}^2$, vamos verificar como fazer para escrever uma integral dupla em coordenadas polares. Sabemos que

$$\iint_R f\ \mathrm{d}A = \iint_R f(x, y)\ \mathrm{d}x\,\mathrm{d}y.$$

A função ϕ que muda de coordenadas polares para Cartesianas pode ser escrita como

$$\phi(r, \theta) = (r\cos\theta, r\,\mathrm{sen}\,\theta),\ \text{onde } r > 0 \text{ e } \theta \in (0, 2\pi).$$

Isto é, as componentes são

$$x = \phi_1(r, \theta) = r\cos\theta, \quad y = \phi_2(r, \theta) = r\,\mathrm{sen}\,\theta.$$

Assim, a matriz das derivadas parciais é

$$\begin{bmatrix} \dfrac{\partial \phi_1}{\partial r} & \dfrac{\partial \phi_1}{\partial \theta} \\ \dfrac{\partial \phi_2}{\partial r} & \dfrac{\partial \phi_2}{\partial \theta} \end{bmatrix} = \begin{bmatrix} \cos\theta & -r\,\mathrm{sen}\,\theta \\ \mathrm{sen}\,\theta & r\cos\theta \end{bmatrix} \text{ cujo Jacobiano é } J\phi = r\cos^2\theta + r\,\mathrm{sen}^2\,\theta = r.$$

Portanto, a fórmula de mudança de variáveis implica que, em coordenadas polares:

$$\iint_R f\ \mathrm{d}A = \iint_R f(x, y)\ \mathrm{d}x\,\mathrm{d}y = \iint_{\phi^{-1}(R)} f(r\cos\theta, r\,\mathrm{sen}\,\theta)\, r\ \mathrm{d}r\,\mathrm{d}\theta.$$

[2]Esta seção é opcional e necessita de conhecimentos básicos de Cálculo.

Licença CC-BY-SA-3.0. Contato: livroscolaborativos@gmail.com

7.4.2 Coordenadas esféricas

Para uma região R do espaço $\mathbb{R}^3$, vamos verificar como fazer para escrever uma integral tripla em coordenadas esféricas. O raciocínio segue as mesmas linhas da subseção anterior para coordenadas polares. Sabemos que

$$\iiint_R f \, \mathrm{d}A = \iiint_R f(x,y,z) \, \mathrm{d}x \, \mathrm{d}y \, \mathrm{d}z.$$

A função ϕ que muda de coordenadas esféricas para Cartesianas pode ser escrita como

$$\phi(\rho,\theta,\phi) = (\rho \operatorname{sen}\phi\cos\theta, \rho\operatorname{sen}\phi\operatorname{sen}\theta, \rho\cos\phi), \text{ onde } \rho > 0, \theta \in (0,2\pi) \text{ e } \phi \in (0,\pi).$$

Isto é, as componentes são

$$x = \phi_1(\rho,\theta,\phi) = \rho\operatorname{sen}\phi\cos\theta, \quad y = \phi_2(\rho,\theta,\phi) = \rho\operatorname{sen}\phi\operatorname{sen}\theta \quad \text{e} \quad z = \phi_3(\rho,\theta,\phi) = \rho\cos\phi.$$

Assim, a matriz das derivadas parciais é

$$\begin{bmatrix} \dfrac{\partial\phi_1}{\partial\rho} & \dfrac{\partial\phi_1}{\partial\theta} & \dfrac{\partial\phi_1}{\partial\phi} \\ \dfrac{\partial\phi_2}{\partial\rho} & \dfrac{\partial\phi_2}{\partial\theta} & \dfrac{\partial\phi_2}{\partial\phi} \\ \dfrac{\partial\phi_3}{\partial\rho} & \dfrac{\partial\phi_3}{\partial\theta} & \dfrac{\partial\phi_3}{\partial\phi} \end{bmatrix} = \begin{bmatrix} \operatorname{sen}\phi\cos\theta & -\rho\operatorname{sen}\phi\operatorname{sen}\theta & \rho\cos\phi\cos\theta \\ \operatorname{sen}\phi\operatorname{sen}\theta & \rho\operatorname{sen}\phi\cos\theta & \rho\cos\phi\operatorname{sen}\theta \\ \cos\phi & 0 & -\rho\operatorname{sen}\phi \end{bmatrix}.$$

cujo Jacobiano é (usando, por exemplo a segunda coluna para expandir em cofatores)

$$\begin{aligned} J\phi &= \left| \rho\operatorname{sen}\phi\operatorname{sen}\theta \begin{vmatrix} \operatorname{sen}\phi\operatorname{sen}\theta & \rho\cos\phi\operatorname{sen}\theta \\ \cos\phi & -\rho\operatorname{sen}\phi \end{vmatrix} + \rho\operatorname{sen}\phi\cos\theta \begin{vmatrix} \operatorname{sen}\phi\cos\theta & \rho\cos\phi\cos\theta \\ \cos\phi & -\rho\operatorname{sen}\phi \end{vmatrix} \right| \\ &= \left| \rho\operatorname{sen}\phi\operatorname{sen}\theta\Big(-\rho\operatorname{sen}^2\phi\operatorname{sen}\theta - \rho\cos^2\phi\operatorname{sen}\theta\Big) + \rho\operatorname{sen}\phi\cos\theta\Big(-\rho\operatorname{sen}^2\phi\cos\theta - \rho\cos^2\phi\cos\theta\Big) \right| \\ &= \left| \rho\operatorname{sen}\phi\operatorname{sen}\theta\Big(-\rho\operatorname{sen}\theta\Big) + \rho\operatorname{sen}\phi\cos\theta\Big(-\rho\cos\theta\Big) \right| = \left| -\rho^2\operatorname{sen}\phi\Big(\operatorname{sen}^2\theta + \cos^2\theta\Big) \right| \\ &= \left| -\rho^2\operatorname{sen}\phi \right| = \rho^2\operatorname{sen}\phi. \end{aligned}$$

Na última igualdade, utilizamos que $\phi \in (0,\pi)$ implica $\operatorname{sen}\phi > 0$, de modo que o valor absoluto está considerado corretamente. Portanto, a fórmula de mudança de variáveis implica que, em coordenadas polares:

$$\iiint_R f \, \mathrm{d}A = \iint_R f(x,y,z) \, \mathrm{d}x \, \mathrm{d}y \, \mathrm{d}z = \iiint_{\phi^{-1}(R)} f(\rho\operatorname{sen}\phi\cos\theta, \rho\operatorname{sen}\phi\operatorname{sen}\theta, \rho\cos\phi)\, \rho^2\operatorname{sen}\phi \, \mathrm{d}\rho \, \mathrm{d}\theta \, \mathrm{d}\phi,$$

como é de costume em cursos de cálculo.

Licença CC-BY-SA-3.0. Contato: livroscolaborativos@gmail.com

8

Semana 10

8.1 Autovalores, autovetores e autoespaços associados

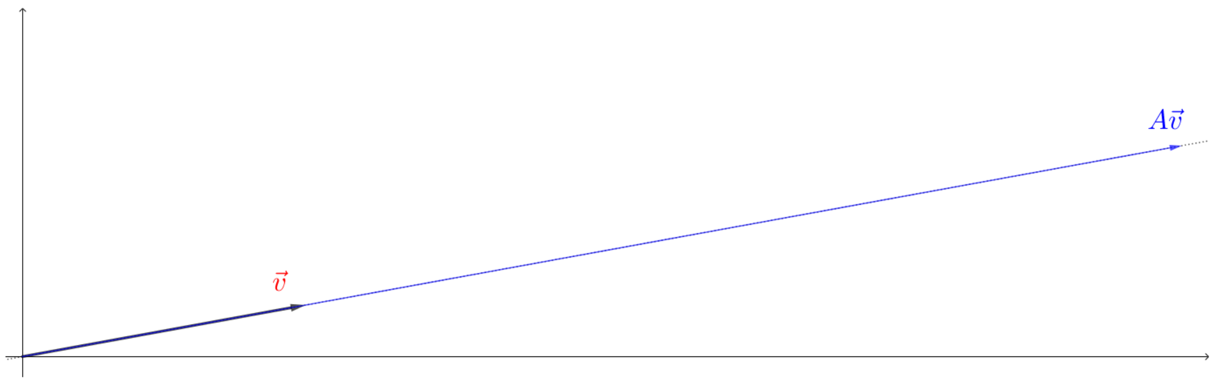
Associados com uma transformação linear T estão os seus autovetores, que, como veremos, são direções especiais para esta transformação T. Por esta razão, são também conhecidos como vetores próprios ou vetores característicos de T. Aparecem em muitas aplicações, pois nos ajudam a entender mais profundamente a transformação linear T.

Dada uma matriz *quadrada* A de ordem $n \times n$, com entradas reais, nós dizemos que um número $\lambda \in \mathbb{R}$ é um **autovalor** de A quando existe um vetor *não nulo* $\vec{v}$ tal que

$$A\vec{v} = \lambda\vec{v}.$$

Neste caso, $\vec{v}$ é dito um **autovetor** de A associado a λ.

Geometricamente, $\vec{v}$ é um vetor que não muda de direção quando aplicamos a matriz A. No entanto, quando permitimos que λ seja um número complexo, esta interpretação geométrica é perdida.



Ver também animação em wikipedia.org

Exemplo 71. Considere a matriz

$$A = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 4 \\ 0 & 4 & 9 \end{bmatrix}.$$

Temos que $\lambda_1 = 2$ é um autovalor desta matriz, porque o vetor

$$\vec{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \quad \text{satisfaz} \quad \begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 4 \\ 0 & 4 & 9 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = 2 \cdot \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}.$$

Assim, podemos dizer que $\vec{v}_1$ é um autovetor de A associado ao autovalor 2. Colocando de outra maneira, podemos dizer também que

$$\vec{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} \quad \text{é um autovetor de } A\text{, pois} \quad \begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 4 \\ 0 & 4 & 9 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 11 \\ 22 \end{bmatrix} = 11 \cdot \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}.$$

Neste caso, concluimos também que 11 é um autovalor de A. $\triangleleft$

Observamos que, se $\vec{v}$ for um autovetor de uma matriz A associado a um autovalor λ, então qualquer múltiplo escalar $\alpha\vec{v}$ também é um autovetor de A associado a λ:

$$A(\alpha\vec{v}) = \alpha A\vec{v} = \alpha\lambda\vec{v} = \lambda(\alpha\vec{v}).$$

Da mesma maneira, se dois vetores $\vec{v}$ e $\vec{w}$ forem autovetores associados a um mesmo autovalor λ, então a soma $\vec{v} + \vec{w}$ também é um autovetor associado a λ:

$$A(\vec{v} + \vec{w}) = A\vec{v} + A\vec{w} = \lambda\vec{v} + \lambda\vec{w} = \lambda(\vec{v} + \vec{v}).$$

Logo, concluimos que o conjunto de todos os autovetores associados a um autovalor λ, união com o vetor nulo, forma um subespaço vetorial de $\mathbb{R}^n$, denominado **autoespaço** associado ao autovalor λ:

$$\text{autoespaço associado a } \lambda = \left\{\vec{v} \in \mathbb{R}^n \mid \vec{v} \text{ é autovetor associado a } \lambda\right\} \cup \{\,\vec{0}\,\}.$$

Vamos analisar uma forma de encontrar os autovalores de uma matriz A de forma sistemática. Temos as seguintes equivalências:

$$\begin{aligned}
&\lambda \text{ é um autovalor de } A\\
\iff &\text{existe } \vec{v} \neq \vec{0} \text{ tal que } A\vec{v} = \lambda\vec{v}\\
\iff &\text{existe } \vec{v} \neq \vec{0} \text{ tal que } (A - \lambda I)\vec{v} = \vec{0}\\
\iff &\text{o sistema homogêneo } (A - \lambda I)\vec{v} = \vec{0} \text{ admite solução não trivial}\\
\iff &A - \lambda I \text{ não é invertível}\\
\iff &\det(A - \lambda I) = 0.
\end{aligned}$$

Esta última equação pode ser utilizada para encontrar os autovalores: são as raízes do **polinômio característico**

$$p(\lambda) = \det(A - \lambda I).$$

A equação $\det(A - \lambda I) = 0$ é também chamada de **equação característica**. Observe, pelas propriedades dos determinantes, que p é de fato um polinômio e tem grau n, que é igual a ordem da matriz A. É assim consequência do Teorema Fundamental da Álgebra que existem no máximo n autovalores reais (já que um polinômio de grau $n \geq 1$ possui exatamente n raízes complexas).

Exemplo 72. Vamos encontrar os autovalores de

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

Precisamos calcular

$$\det \begin{bmatrix} 4-\lambda & 3 \\ 1 & 2-\lambda \end{bmatrix} = (4-\lambda)(2-\lambda) - 3 \cdot 1 = \lambda^2 - 6\lambda + 5,$$

que tem raízes

$$\lambda = \frac{6 \pm \sqrt{36 - 20}}{2} = 3 \pm 2.$$

Portanto, A tem dois autovalores reais: $\lambda_1 = 1$ e $\lambda_2 = 5$.

Exemplo 73. Encontrar os autovalores de

$$A = \begin{bmatrix} 5 & 8 & 16 \\ 4 & 1 & 8 \\ -4 & -4 & -11 \end{bmatrix}.$$

Licença CC-BY-SA-3.0. Contato: livroscolaborativos@gmail.com

Precisamos calcular

$$
\begin{aligned}
p(\lambda) = \det(A - \lambda I) = \det \begin{bmatrix} 5-\lambda & 8 & 16 \\ 4 & 1-\lambda & 8 \\ -4 & -4 & -11-\lambda \end{bmatrix} \\
&= (5-\lambda) \cdot \begin{vmatrix} 1-\lambda & 8 \\ -4 & -11-\lambda \end{vmatrix} - 8 \cdot \begin{vmatrix} 4 & 8 \\ -4 & -11-\lambda \end{vmatrix} + 16 \cdot \begin{vmatrix} 4 & 1-\lambda \\ -4 & -4 \end{vmatrix} \\
&= (5-\lambda)\big[(1-\lambda)(-11-\lambda)+32\big] - 8\big[-44-4\lambda+32\big] + 16(-16+4-4\lambda) \\
&= (5-\lambda)\big[\lambda^2+10\lambda+21\big] + 32\lambda + 12\cdot 8 - 64\lambda - 12\cdot 16 \\
&= -\lambda^3 - 5\lambda^2 + 29\lambda + 105 - 32\lambda - 96 \\
&= -\lambda^3 - 5\lambda^2 - 3\lambda + 9.
\end{aligned}
$$

As possíveis raízes racionais desse polinômio só podem ser os divisores do termo independente acima: $\pm 1, \pm 3, \pm 9$. Verificamos que 1 é raiz. Logo, dividindo $p(\lambda)$ por $\lambda - 1$:

$$-\lambda^3 - 5\lambda^2 - 3\lambda + 9 = -(\lambda-1)(\lambda^2+6\lambda+9) = -(\lambda-1)(\lambda+3)^2.$$

Portanto, os autovalores de A são $\lambda_1 = 1$ e $\lambda_2 = 3$. Observamos que 3 é uma raiz de multiplicidade 2 do polinômio característico. $\lhd$

Um grande problema é que, em geral, encontrar raízes de polinômios é difícil. De fato, mesmo quando o grau do polinômio é baixo, encontrar raízes de polinômios pode ser bastante complicado. Já há muito tempo é conhecido[1] que polinômios de grau 5 ou mais podem não possuir fórmulas para cálculo de raízes a partir de radicais.

Por outro lado, uma vez que os autovalores são conhecidos, encontrar os autovetores é um cálculo direto: basta resolver o sistema linear homogêneo

$$(A - \lambda I)\vec{v} = \vec{0}.$$

Isto é o mesmo que dizer que o autoespaço associado a λ é o espaço nulo $\mathrm{Nul}(A - \lambda I)$. Vejamos como encontrar os autoespaços das matrizes dos exemplos anteriores.

Exemplo 74 (de volta ao Exemplo 72)**.** Para encontrar os autovetores de

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

associados com o autovalor $\lambda_1 = 1$, vamos resolver o sistema homogêneo:

$$\begin{bmatrix} 4-\lambda & 3 \\ 1 & 2-\lambda \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \rightsquigarrow \begin{bmatrix} 3 & 3 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Já que o sistema é homogêneo, não é necessário escrever a última coluna de zeros na matriz aumentada associada (no entanto, é necessário lembrar que há uma coluna de zeros). Por escalonamento

$$\begin{bmatrix} 3 & 3 \\ 1 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \rightsquigarrow \begin{cases} v_1 + v_2 = 0 \\ 1 \text{ variável livre.} \end{cases}$$

Em forma vetorial paramétrica:

$$\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} -v_2 \\ v_2 \end{bmatrix} = v_2 \begin{bmatrix} -1 \\ 1 \end{bmatrix} \implies \mathrm{Nul}(A - I) = \mathrm{Span}\left\{ \begin{bmatrix} -1 \\ 1 \end{bmatrix} \right\}.$$

Para encontrar os autovetores associados a $\lambda_2 = 5$, resolvemos

$$\begin{bmatrix} 4-\lambda & 3 \\ 1 & 2-\lambda \end{bmatrix} = \begin{bmatrix} -1 & 3 \\ 1 & -3 \end{bmatrix} \sim \begin{bmatrix} 1 & -3 \\ 0 & 0 \end{bmatrix} \rightsquigarrow \begin{cases} v_1 - 3v_2 = 0 \\ 1 \text{ variável livre.} \end{cases}$$

Em forma vetorial paramétrica:

$$\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 3v_2 \\ v_2 \end{bmatrix} = v_2 \begin{bmatrix} 3 \\ 1 \end{bmatrix} \implies \mathrm{Nul}(A - 5I) = \mathrm{Span}\left\{ \begin{bmatrix} 3 \\ 1 \end{bmatrix} \right\}.$$

[1] Este resultado é uma consequência da famosa Teoria de Galois.

Exemplo 75 (de volta ao Exemplo 73)**.** Vamos encontrar os autovetores de

$$A = \begin{bmatrix} 5 & 8 & 16 \\ 4 & 1 & 8 \\ -4 & -4 & -11 \end{bmatrix}.$$

- associados com o autovalor $\lambda_1 = 1$, vamos resolver o sistema homogêneo:

$$\begin{bmatrix} 5-\lambda_1 & 8 & 16 \\ 4 & 1-\lambda_1 & 8 \\ -4 & -4 & -11-\lambda_1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Por escalonamento,

$$\begin{bmatrix} 4 & 8 & 16 \\ 4 & 0 & 8 \\ -4 & -4 & -12 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 4 \\ 0 & -8 & -8 \\ 0 & 4 & 4 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 4 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix} \leftrightsquigarrow \begin{cases} v_1 + 2v_3 = 0 \\ v_2 + v_3 = 0 \\ v_3 \text{ livre} \end{cases}$$

Em forma paramétrica, os autovetores são

$$\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} -2v_3 \\ -v_3 \\ v_3 \end{bmatrix} = v_3 \begin{bmatrix} -2 \\ -1 \\ 1 \end{bmatrix} \implies \operatorname{Nul}(A - I) = \operatorname{Span}\left\{ \begin{bmatrix} -2 \\ -1 \\ 1 \end{bmatrix} \right\}.$$

- associados com o autovalor $\lambda_2 = -3$, vamos resolver o sistema homogêneo:

$$\begin{bmatrix} 5-\lambda_2 & 8 & 16 \\ 4 & 1-\lambda_2 & 8 \\ -4 & -4 & -11-\lambda_2 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Por escalonamento,

$$\begin{bmatrix} 8 & 8 & 16 \\ 4 & 4 & 8 \\ -4 & -4 & -8 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \leftrightsquigarrow \begin{cases} v_1 + v_2 + 2v_3 = 0 \\ \text{2 variáveis livres} \end{cases}$$

Em forma paramétrica, os autovetores são

$$\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} -v_2 - 2v_3 \\ v_2 \\ v_3 \end{bmatrix} = v_2 \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + v_3 \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix} \implies \operatorname{Nul}(A-I) = \operatorname{Span}\left\{ \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix} \right\}. \lhd$$

Observamos que, nos exemplos anteriores, a dimensão do autoespaço associado ficou igual à multiplicidade do autovalor. **Isto nem sempre é verdade**. Como exercício, verifique que a dimensão do autoespaço associado ao autovalor $\lambda = 1$ da matriz

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

é igual a 1, embora a multiplicidade do autovalor seja 2.

De forma geral, chamamos a multiplicidade do autovalor de **multiplicidade algébrica**, enquanto que a dimensão do autoespaço associado é chamada de **multiplicidade geométrica**.

8.2 Diagonalização

Matrizes diagonais são matrizes da forma

$$\begin{bmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ 0 & 0 & \lambda_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_n \end{bmatrix}.$$

Licença CC-BY-SA-3.0. Contato: livroscolaborativos@gmail.com

Já que são triangulares, seus autovalores são os elementos da diagonal principal. Observamos também que os autovetores associados são os elementos da base canônica de $\mathbb{R}^n$.

Na realidade, o fato de se conseguir formar uma base com autovetores de uma matriz está intimamente relacionado com se obter uma matriz diagonal semelhante a A. De fato, vamos justificar que, sempre que for possível formar uma base do espaço $\mathbb{R}^n$ apenas com autovetores de uma matriz quadrada A, então é possível fazer uma mudança de base de modo que a obter uma matriz diagonal, com os autovalores na diagonal principal. Este procedimento é chamado de **diagonalização** da matriz A, por motivos óbvios. O método, que é surpreendentemente simples (mas, na prática, trabalhoso), consiste em montar uma matriz P com os autovetores de A e efetuar multiplicações de matrizes:

$$P = \begin{bmatrix} | & | & & | \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_n \\ | & | & & | \end{bmatrix} \implies P^{-1}AP = D = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}.$$

Justificativa do método de diagonalização. Vamos verificar que $P^{-1}AP = D$ mostrando que as colunas da matriz $P^{-1}AP$ são iguais as colunas de D. Pela forma que construimos a matriz P, sabemos que a primeira coluna de P é $\vec{v}_1$:

$$P\vec{e}_1 = \vec{v}_1 \qquad \text{ou, ainda,} \qquad \vec{e}_1 = P^{-1}\vec{v}_1.$$

Assim, já que $A\vec{v}_1 = \lambda_1\vec{v}_1$, temos

$$\begin{aligned} \text{primeira coluna de } P^{-1}AP &= P^{-1}AP\vec{e}_1 \\ &= P^{-1}A\vec{v}_1 \\ &= P^{-1}\lambda_1\vec{v}_1 \\ &= \lambda_1 P^{-1}\vec{v}_1 \\ &= \lambda_1\vec{e}_1 \\ &= \text{primeira coluna de } D. \end{aligned}$$

Este mesmo raciocício, repetido n vezes, para cada uma das colunas, prova que todas as colunas são iguais; portanto, $P^{-1}AP = D$. □

Exemplo 76 (de volta ao Exemplo 73)**.** Vimos que, para a matriz

$$A = \begin{bmatrix} 5 & 8 & 16 \\ 4 & 1 & 8 \\ -4 & -4 & -11 \end{bmatrix},$$

temos autovalores $\lambda_1 = 1$ e $\lambda_2 = -3$ e autoespaços associados:

$$\text{Nul}(A - I) = \text{Span}\left\{ \begin{bmatrix} -2 \\ -1 \\ 1 \end{bmatrix} \right\}; \qquad \text{Nul}(A - I) = \text{Span}\left\{ \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix} \right\}.$$

Montamos a matriz

$$P = \begin{bmatrix} -2 & -1 & -2 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

Por escalonamento, podemos calcular a matriz inversa

$$P^{-1} = \begin{bmatrix} -1 & -1 & -2 \\ -1 & 0 & -2 \\ 1 & 1 & 3 \end{bmatrix}$$

e temos

$$\begin{aligned} P^{-1}AP &= \begin{bmatrix} -1 & -1 & -2 \\ -1 & 0 & -2 \\ 1 & 1 & 3 \end{bmatrix} \begin{bmatrix} 5 & 8 & 16 \\ 4 & 1 & 8 \\ -4 & -4 & -11 \end{bmatrix} \begin{bmatrix} -2 & -1 & -2 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} -1 & -1 & -2 \\ 3 & 0 & 6 \\ -3 & -3 & -9 \end{bmatrix} \begin{bmatrix} -2 & -1 & -2 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & -3 \end{bmatrix}. \lhd \end{aligned}$$

Licença CC-BY-SA-3.0. Contato: livroscolaborativos@gmail.com

Embora o método de diagonalização seja bastante transparente, pois basta seguir uma sequência de passos, ele é muito trabalhoso! Não devemos nos iludir com a aparente simplicidade das contas no Exemplo 76, pois as contas já haviam sido feitas anteriormente (além de que "escondemos" o cálculo da inversa P^{-1}). Na totalidade, os passos envolvem:

- Calcular o polinômio característico $p(\lambda) = \det(A - \lambda I)$. Aqui já entra uma quantidade significativa de cálculos, cuja quantidade aumenta de acordo com a dimensão do problema;
- Encontrar as raízes do polinômio característico. Este passo pode ser muito complicado, já que muitas vezes não é possível encontrar raízes explicitamente;
- Depois de conhecidos os autovalores, calcular cada um dos autoespaços associados, o que demanda a resolução de tantos sistemas lineares quantos forem os autovalores distintos da matriz;
- Caso todos os geradores de todos os autoespaços formem uma base para o espaço $\mathbb{R}^n$, a matriz A é diagonalizável e formamos a matriz P como indicado acima;
- Observe que, se estivermos apenas interessados em encontrar uma base de autovetores e uma matriz diagonal equivalente a A, então não há necessidade de se calcular P^{-1}, pois já sabemos que o método funciona e que $P^{-1}AP = D$. No entanto, uma descrição mais completa pode ser necessária nas principais aplicações, de modo que ainda seria preciso uma quantidade grande de cálculos para determinar a matriz P^{-1}.

Nem todas as matrizes são diagonalizáveis, como acima. Já vimos que isto é possível quando uma base de autovetores existe. No exemplo

$$A = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}$$

não há dois autovetores linearmente independentes e, portanto, não é possível diagonalizar a matriz A.

Em geral, valem as seguintes propriedades: seja A uma matriz de orden $n \times n$. Então:

- Se A possui n autovalores reais distintos, então A possui uma base de autovetores e é diagonalizável, pois possui um autovetor associado a cada um dos seus autovalores distintos.
- No caso de A possuir autovalores reais com multiplicidade maior do que 1, A apenas será diagonalizável se cada autovalor de multiplicidade k tiver k autovetores linearmente independentes. Em outras palavras, só será possível formar uma base de autovetores de A quando
$$\dim \mathrm{Nul}(A - \lambda I) = \text{ multiplicidade do autovalor } \lambda.$$
Na notação introduzida não seção anterior, isto é o mesmo que dizer que, se A for diagonalizável, então as multiplicidades algébrica e geométrica são iguais.
- Caso $\dim \mathrm{Nul}(A - \lambda I)$ seja menor do que a multiplicidade do autovalor λ ou A possua autovalores complexos, então A não é diagonalizável.

Exemplo 77. A matriz

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

tem polinômio característico

$$p(\lambda) = \det \begin{bmatrix} 2-\lambda & 1 \\ -1 & 2-\lambda \end{bmatrix} = (2-\lambda)^2 + 1 = \lambda^2 - 4\lambda + 5$$

cujas raízes são $\lambda_1 = \dfrac{4+3i}{2}$ e $\lambda_2 = \dfrac{4-3i}{2}$. Logo, A não é diagonalizável.

Licença CC-BY-SA-3.0. Contato: livroscolaborativos@gmail.com

9

Semana 11

Nesta semana, nós começamos uma abordagem um pouco mais geométrica de espaços vetoriais: ângulo entre vetores, comprimento de vetores, ortogonalidade. Estes conceitos já são de nossa familiaridade da geometria euclideana (em $\mathbb{R}^2$ e $\mathbb{R}^3$). No entanto, é também nosso objetivo estender estes conceitos para dimensões maiores do que 3.

9.1 Comprimento, ângulos e o produto escalar

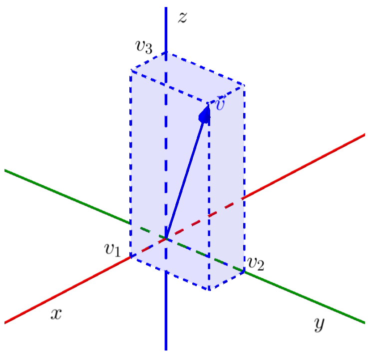
Se, na base canônica (isto é, no sistema de coordenadas usual) do espaço $\mathbb{R}^3$, representamos um vetor por

$$\vec{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix},$$

então o seu **comprimento** (ou **magnitude** ou **norma**) é dado por

$$\|\vec{v}\| = \sqrt{v_1^2 + v_2^2 + v_3^2}.$$

Esta é uma instância do Teorema de Pitágoras[1].



A definição acima pode ser generalizada para dimensão n qualquer:

$$\vec{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \quad \rightsquigarrow \quad \|\vec{v}\| = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}.$$

O **produto escalar** ou **produto interno** entre dois vetores $\vec{v}$ e $\vec{w}$ é um número (também dito escalar) associado com estes dois vetores. Em coordenadas, se $\vec{v} = (v_1, v_2, v_3)$ e $\vec{w} = (w_1, w_2, w_3)$

[1] Duas aplicações do Teorema de Pitágoras usual do plano. Uma para obter o comprimento $\sqrt{v_1^2 + v_2^2}$ da diagonal que está no plano xy e a outra para obter $\|\vec{v}\|$.

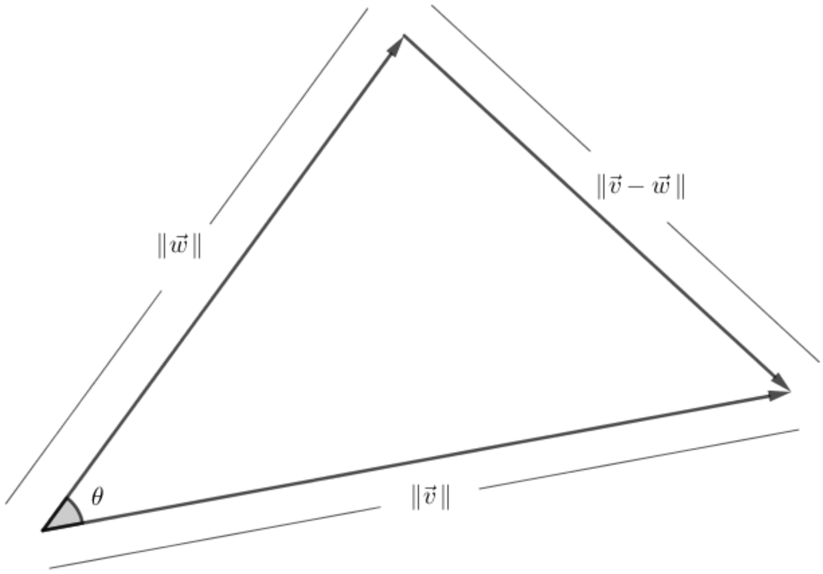
são vetores do espaço tridimensional $\mathbb{R}^3$, temos

$$\vec{v} \cdot \vec{w} = v_1 w_1 + v_2 w_2 + v_3 w_3.$$

Vejamos que realmente o produto escalar tem a ver com o ângulo entre dois vetores. Isto é uma consequência da Lei dos Cossenos, que nos diz que

$$\|\vec{v} - \vec{w}\|^2 = \|\vec{v}\|^2 + \|\vec{w}\|^2 - 2\|\vec{v}\|\|\vec{w}\| \cos\theta,$$

onde θ é o ângulo entre $\vec{v}$ e $\vec{w}$.



Escrevendo estes comprimentos em termos das coordenadas e fazendo os cancelamentos necessários, chegamos na seguinte identidade:

$$v_1 w_1 + v_2 w_2 + v_3 w_3 = \|\vec{v}\|\|\vec{w}\| \cos\theta, \quad \text{isto é,} \quad \vec{v} \cdot \vec{w} = \|\vec{v}\|\|\vec{w}\| \cos\theta.$$

Motivados pelos conceitos acima, consideramos

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad \text{e} \quad \vec{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

vetores de $\mathbb{R}^n$ representados na base canônica, e definimos o **produto escalar** ou **produto interno** entre $\vec{x}$ e $\vec{y}$:

$$\vec{x} \cdot \vec{y} = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n.$$

O ângulo entre $\vec{x}$ e $\vec{y}$, para n qualquer, pode então ser *definido* pela equação

$$\cos\theta = \frac{\vec{x} \cdot \vec{y}}{\|\vec{x}\|\|\vec{y}\|}.$$

Embora abstratos, estes conceitos são bem fundamentados geometricamente. Dois vetores (linearmente independentes) de um espaço vetorial n-dimensional geram um subespaço vetorial de dimensão 2. Este pode, por sua vez, pode ser pensado como um plano que passa pela origem de $\mathbb{R}^n$. Assim, podemos imaginar o ângulo entre estes vetores naturalmente.

Proposição 78 (Principais propriedades do produto escalar)**.** *Para $\vec{x}, \vec{y}, \vec{z} \in \mathbb{R}^n$ e $c \in \mathbb{R}$ um escalar, temos as seguintes propriedades*

1. $\vec{x} \cdot \vec{y} = \vec{y} \cdot \vec{x}$

2. $(\vec{x}+\vec{y})\cdot\vec{z} = \vec{x}\cdot\vec{z}+\vec{y}\cdot\vec{z}$

3. $(c\vec{x})\cdot\vec{y} = c\vec{x}\cdot\vec{y} = \vec{x}\cdot(c\vec{y})$

4. $\vec{x}\cdot\vec{x} \geq 0$ *e* $\vec{x}\cdot\vec{x} = 0$ *se, e somente se,* $\vec{x} = \vec{0}$*;*

5. $\|\vec{x}\| = \sqrt{\vec{x}\cdot\vec{x}}$*;*

6. $\|c\vec{x}\| = |c|\|\vec{x}\|$*.*

Estas propriedades são todas de fácil verificação (e são deixadas para o leitor verificar, como exercício).

Exemplo 79. Considere os vetores de $\mathbb{R}^5$ escritos na base canônica:

$$\vec{x} = \begin{bmatrix} -1 \\ 2 \\ 0 \\ 3 \\ -1 \end{bmatrix} \quad \text{e} \quad \vec{y} = \begin{bmatrix} 4 \\ 3 \\ -1 \\ -11 \\ 1 \end{bmatrix}$$

Assim, temos

$$\vec{x}\cdot\vec{y} = (-1)\cdot 4 + 2\cdot 3 + 0\cdot(-1) + 3\cdot(-11) + (-1)\cdot 1 = -4+6+0-33-1 = -31.$$

Da mesma forma, podemos calcular os comprimentos

$$\|\vec{x}\| = \sqrt{(-1)^2+2^2+0^2+3^2+(-1)^2} = \sqrt{15} \quad \text{e} \quad \|\vec{y}\| = \sqrt{4^2+3^2+(-1)^2+(-11)^2+1^2} = \sqrt{148}.$$

Logo, o ângulo θ entre estes vetores satisfaz (utilizamos uma calculadora para calcular as raízes quadradas e a função inversa do cosseno):

$$\cos\theta = \frac{\vec{x}\cdot\vec{y}}{\|\vec{x}\|\|\vec{y}\|} = \frac{-31}{\sqrt{15}\sqrt{148}} \simeq -0,65794 \implies \theta \simeq 2,28887634 \text{ radianos} \simeq 131,12^{\text{o}}. \lhd$$

Exemplo 80. Considere um vetor tridimensional

$$\vec{u} = \begin{bmatrix} 2 \\ 3 \\ -1 \end{bmatrix}.$$

Já sabemos a multiplicação de um escalar pelo vetor $\vec{u}$ muda o comprimeto (e talvez também o sentido) do vetor. Por exemplo, se multiplicarmos por 2, o comprimento será multiplicado por 2; se dividirmos por 3, o comprimento será dividido por 3. Se nós dividirmos pelo próprio comprimento de $\vec{u}$ nós obtemos um vetor unitário. Para verificar analíticamente, basta utilizar a Propriedade *6* acima:

$$c = \frac{1}{\|\vec{u}\|} \implies \left\|\frac{\vec{u}}{\|\vec{u}\|}\right\| = \|c\vec{u}\| = c\|\vec{u}\| = \frac{1}{\|\vec{u}\|}\|\vec{u}\| = 1.$$

Então se quisermos encontrar um vetor unitário na direção de $\vec{u}$ acima, calculamos

$$\|\vec{u}\| = \sqrt{4+9+1} = \sqrt{14} \quad \rightsquigarrow \quad \vec{v} = \frac{\vec{u}}{\|\vec{u}\|} = \frac{\vec{u}}{\sqrt{14}} = \begin{bmatrix} 2/\sqrt{14} \\ 3/\sqrt{14} \\ -1/\sqrt{14} \end{bmatrix}$$

é o vetor procurado. Este processo de obter um vetor unitário na mesma direção de um vetor previamente fixado é chamado de **normalização**. Observe que existem apenas dois vetores unitários na mesma direção de $\vec{u}$, a saber, $\vec{v}$ e $-\vec{v}$. $\lhd$

Licença CC-BY-SA-3.0. Contato: livroscolaborativos@gmail.com

9.2 Ortogonalidade

Como vimos na seção anterior, o produto escalar está relacionado com o ângulo entre dois vetores pela fórmula

$$\cos\theta = \frac{\vec{x}\cdot\vec{y}}{\|\vec{x}\|\|\vec{y}\|}.$$

Quando este ângulo θ é igual a $\pi/2$ (que é equivalente a 90^o), vale que $\cos\theta = 0$ e logo $\vec{x}\cdot\vec{y} = 0$.

Nós vamos dizer que os dois vetores $\vec{x}$ e $\vec{y}$ são **ortogonais** quando satisfazem

$$\vec{x}\cdot\vec{y} = 0.$$

Poderíamos também usar a palavra perpendiculares, mas por algum motivo esta terminologia é bastante menos comum quando lidamos com vetores em espaços vetoriais de dimensão maior do que 3.

Dizemos também que um conjunto de vetores é um **conjunto ortogonal** se todo par de vetores do conjunto for ortogonal. Em outras palavras, um conjunto $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k\}$ é um conjunto ortonogonal se, para qualquer escolha de índices $i \neq j$, tivermos $\vec{v}_i \cdot \vec{v}_j = 0$.

Se, além de o conjunto ser ortogonal, todos os seus elementos serem unitários (estarem normalizados), então nós dizemos que o conjunto é um **conjunto ortonormal**. De outra maneira, um conjunto $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k\}$ é um conjunto ortonormal se, para qualquer índice i, $\|\vec{v}_i\| = 1$ e, para qualquer escolha de índices $i \neq j$, tivermos $\vec{v}_i \cdot \vec{v}_j = 0$.[2]

Exemplo 81. Os vetores

$$\vec{x} = \begin{bmatrix} -1 \\ 2 \end{bmatrix} \quad \text{e} \quad \vec{y} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

são ortogonais, pois $\vec{x}\cdot\vec{y} = (-1)\cdot 2 + 2\cdot 1 = 0$, enquanto que os vetores

$$\vec{w} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \quad \text{e} \quad \vec{z} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

não são ortogonais, já que $\vec{w}\cdot\vec{z} = 1\cdot 2 + 2\cdot 1 = 4$. Nestes casos, podemso dizer que o conjunto $\{\vec{x}, \vec{y}\}$ é um conjunto ortogonal enquanto $\{\vec{w}, \vec{z}\}$ não é.

Exemplo 82. A base canônica do espaço $\mathbb{R}^n$:

$$\vec{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \ \vec{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \ \vec{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \ \cdots, \vec{e}_n = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

é um conjunto ortogonal, porque quando escolhermos $i \neq j$, temos

$$\vec{e}_i \cdot \vec{e}_j = 1\cdot 0 + 0\cdot 1 = 0.$$

É também ortonormal, já que, para todo i,

$$\|\vec{e}_i\| = \sqrt{0^2 + 0^2 + \cdots + 1^2 + \cdots + 0^2 + 0^2} = 1. \ \lhd$$

Uma primeira propriedade de ortogonalidade é uma versão para espaços vetoriais quaisquer do Teorema de Pitágoras, que aparece naturamente ao considerarmos vetores ortogonais.

Teorema 83 (Teorema de Pitágoras)**.** *Se dois vetores $\vec{x}$ e $\vec{y}$ são ortogonais, então*

$$\|\vec{x} + \vec{y}\|^2 = \|\vec{x}\|^2 + \|\vec{y}\|^2.$$

[2]Uma notação bastante comum, principalmente em matemática e física, é a do delta de Kronecker:

$$\delta_{ij} = \begin{cases} 1, \text{ se } i = j \\ 0, \text{ se } i \neq j \end{cases}.$$

Em notação matricial, $I_k = (\delta_{ij})$ é a matriz identidade. Nesta notação, podemos dizer, mais sucintamente, que um conjunto $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k\}$ é ortonormal se $\vec{v}_i \cdot \vec{v}_j = \delta_{ij}$ para quaisquer $i, j = 1, 2, 3, \ldots, k$.

Licença CC-BY-SA-3.0. Contato: livroscolaborativos@gmail.com

Demonstração. Esperamos que esta demonstração auxilie no compreendimento analítico e geométrico dos conceitos que introduzimos até agora. O vetor $\vec{x}+\vec{y}$ pode ser considerado a hipotenusa do triângulo retângulo de catetos $\vec{x}$ e $\vec{y}$ (fazer uma figura!).

Pelas propriedades do comprimento e do produto escalar, temos que, para *quaisquer* dois vetores $\vec{x}$ e $\vec{y}$ (não necessariamente ortogonais), vale que

$$\|\vec{x}+\vec{y}\|^2 = (\vec{x}+\vec{y})\cdot(\vec{x}+\vec{y}) = \vec{x}\cdot\vec{x}+\vec{x}\cdot\vec{y}+\vec{y}\cdot\vec{x}+\vec{y}\cdot\vec{y} = \|\vec{x}\|^2 + 2(\vec{x}\cdot\vec{y}) + \|\vec{y}\|^2.$$

De certa maneira, esta é uma forma de reescrever a Lei dos Cossenos. Quando os vetores $\vec{x}$ e $\vec{y}$ forem ortogonais, temos $\vec{x}\cdot\vec{y}=0$, concluimos que

$$\|\vec{x}+\vec{y}\|^2 = \|\vec{x}\|^2 + \|\vec{y}\|^2.$$ □

Notamos que, de acordo com a nossa definição, o vetor nulo é ortogonal a qualquer outro. Vejamos, no entanto, que vetores não nulos satisfazem propriedades adicionais.

Teorema 84. *Todo conjunto ortogonal $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k\}$ formado por vetores não nulos é um conjunto linearmente independente (LI).*

Demonstração. Considerando uma combinação linear

$$c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_k\vec{v}_k = \vec{0}$$

e fazendo o produto escalar com qualquer dos vetores $\vec{v}_j$ na equação acima, concluimos que

$$(c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_k\vec{v}_k)\cdot\vec{v}_j = \vec{0}\cdot\vec{v}_j.$$

Agora, a condição de ortogonalidade nos diz que $\vec{v}_i\cdot\vec{v}_j = 0$ sempre que $i\neq j$. Portanto,

$$c_j\|\vec{v}_j\| = c_j(\vec{v}_j\cdot\vec{v}_j) = 0 \overset{\vec{v}_j\neq\vec{0}}{\Longrightarrow} c_j = 0.$$

Isto é a definição de um conjunto ser LI: qualquer combinação linear que resulte no vetor nulo deve ter todos os coeficientes nulos. □

Exemplo 85. O conjunto

$$\left\{ \vec{u}_1 = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix},\ \vec{u}_2 = \begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix},\ \vec{u}_3 = \begin{bmatrix} 2 \\ -2 \\ 1 \end{bmatrix} \right\}$$

pois

$$\begin{aligned} \vec{u}_1\cdot\vec{u}_2 &= 2+2-4 = 0, \\ \vec{u}_1\cdot\vec{u}_3 &= 2-4+2 = 0, \\ \vec{u}_2\cdot\vec{u}_3 &= 4-2-2 = 0. \end{aligned}$$

Logo, pelo teorema acima, o conjunto $\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$ é linearmente independente. Neste exemplo, são todos elementos de $\mathbb{R}^3$; portanto, formam uma base para $\mathbb{R}^3$ que é também ortogonal (voltaremos a falar sobre bases ortogonais em breve). ◁

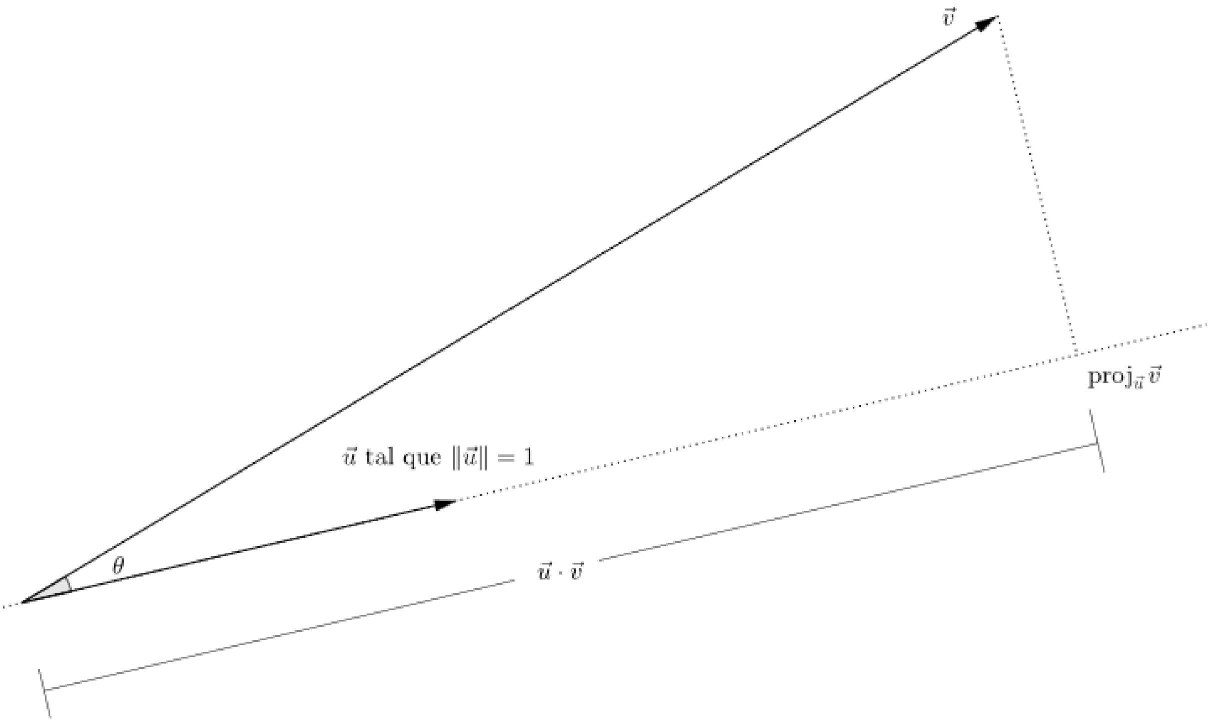
Uma das aplicações mais importantes do produto escalar reside no fato de nós conseguirmos fazer projeções ortogonais em direções fixadas. Por exemplo, suponhamos que em um sistema físico, um vetor $\vec{v}$ representa a força aplicada em um ponto, como na figura abaixo. Gostaríamos de encontrar a componente do vetor na direção da reta tracejada. Para isto, podemos considerar um vetor *unitário* $\vec{u}$ sobre a reta. Assim, a componente de $\vec{v}$ na direção de $\vec{u}$ é igual a $k\vec{u}$, onde k tem a ver com o cosseno do ângulo entre os vetores:

$$\cos\theta = \frac{\text{cateto adjacente}}{\text{hipotenusa}} = \frac{k}{\|\vec{v}\|}.$$

Utilizando novamente que $\vec{u}$ é um vetor unitário (isto é, $\|\vec{u}\| = 1$), concluimos que

$$k = \|\vec{v}\|\cos\theta = \|\vec{u}\|\|\vec{v}\|\cos\theta = \vec{u}\cdot\vec{v}.$$

Licença CC-BY-SA-3.0. Contato: livroscolaborativos@gmail.com



Assim sendo, definimos a **projeção ortogonal de $\vec{v}$ em um vetor unitário $\vec{u}$** como sendo

$$\boxed{\mathrm{proj}_{\vec{u}}\,\vec{v} = (\vec{u}\cdot\vec{v})\,\vec{u}.}$$

No **caso onde $\vec{u}$ não está normalizado**, devemos primeiramente normalizar o vetor, de modo que a projeção ortogonal é dada por

$$\boxed{\mathrm{proj}_{\vec{u}}\,\vec{v} = \left(\frac{\vec{u}}{\|\vec{u}\|}\cdot\vec{v}\right)\frac{\vec{u}}{\|\vec{u}\|} = \frac{\vec{u}\cdot\vec{v}}{\|\vec{u}\|^2}\,\vec{u} = \frac{\vec{u}\cdot\vec{v}}{\vec{u}\cdot\vec{u}}\,\vec{u}.}$$

Exemplo 86. Considere vetores no espaço de quatro dimensões $\mathbb{R}^4$. Vamos determinar a projeção do vetor $\vec{v} = \begin{bmatrix} 1 \\ -1 \\ 3 \\ 4 \end{bmatrix}$ na direção da reta que tem a mesma direção do vetor $\vec{u} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ -1 \end{bmatrix}$.
Observamos inicialmente que $\|\vec{u}\| = \sqrt{4} = 2$ não é unitário. Temos que

$$\mathrm{proj}_{\vec{u}}\,\vec{v} = \frac{\vec{u}\cdot\vec{v}}{\|\vec{u}\|^2}\,\vec{u} = \frac{1-1+3-4}{4}\begin{bmatrix} 1 \\ 1 \\ 1 \\ -1 \end{bmatrix} = -\frac{1}{4}\begin{bmatrix} 1 \\ 1 \\ 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 1/4 \\ 1/4 \\ 1/4 \\ -1/4 \end{bmatrix}.$$

A projeção de $\vec{v}$ sobre o vetor $\vec{e}_3$ da base canônica de $\mathbb{R}^4$ é (observe que $\|\vec{e}_3\| = 1$)

$$\mathrm{proj}_{\vec{e}_3}\,\vec{v} = (\vec{e}_3\cdot\vec{v})\,\vec{e}_3 = (0+0+3+0)\begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 3 \\ 0 \end{bmatrix}.$$

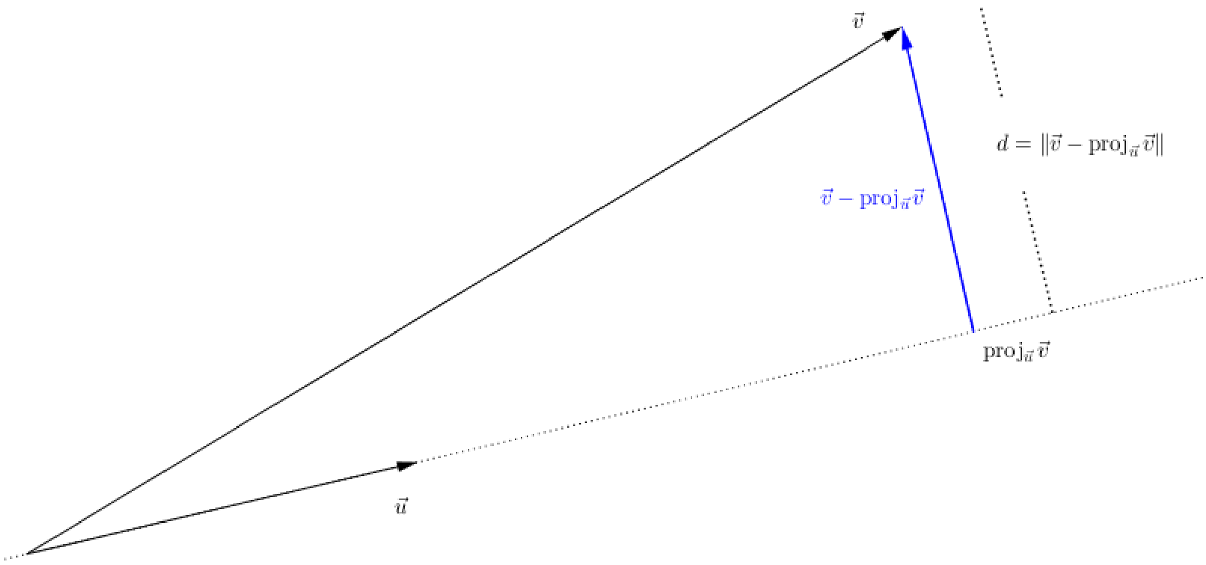
Observe que a projeção sobre o vetor $\vec{e}_3$ da base canônica apenas "enfatiza" a terceira componente do vetor $\vec{v}$. Isto é bastante natural, pois $\vec{v}\cdot\vec{e}_3$ deve fornecer o comprimento da projeção de $\vec{v}$ na direção do vetor unitário $\vec{e}_3$, que deve ser igual a terceira componente do vetor. É justamente assim que usamos a base canônica para representar vetores geometricamente. De forma geral, podemos escrever

$$\vec{v} = (\vec{v}\cdot\vec{e}_1)\vec{e}_1 + (\vec{v}\cdot\vec{e}_2)\vec{e}_2 + (\vec{v}\cdot\vec{e}_3)\vec{e}_3 + (\vec{v}\cdot\vec{e}_4)\vec{e}_4.$$

Licença CC-BY-SA-3.0. Contato: livroscolaborativos@gmail.com

Este tipo de represetação será explorado com mais detalhes na próxima seção. $\triangleleft$

As projeções ortogonais também podem ser utilizadas para o cálculo de distâncias de ponto à retas que passam pela origem: Pela interpretação geométrica da subtração de vetores, obtemos um vetor perpendicular à reta com a direção de $\vec{u}$ e que termina onde termina o vetor $\vec{v}$ (ver figura).



Desta maneira, o comprimento deste vetor, a saber $\|\vec{u} - \mathrm{proj}_{\vec{u}}\,\vec{v}\|$, é a distância procurada.[3]

Exemplo 87. Calcular a distância do ponto $P = (0, 2, 3)$ e a reta que passa pela origem e tem a direção do vetor $\vec{u} = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$.

Para isto, consideramos o vetor $\vec{v}$ cujas componentes são as componentes do ponto P. Assim, estamos na situação da discussão acima e vale que

$$\mathrm{proj}_{\vec{u}}\,\vec{v} = \frac{7}{6}\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \implies \vec{u} - \mathrm{proj}_{\vec{u}}\,\vec{v} = \begin{bmatrix} 0 \\ 2 \\ 3 \end{bmatrix} - \frac{7}{6}\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} -7/6 \\ -1/3 \\ 11/6 \end{bmatrix}.$$

Portanto,

$$d = \|\vec{u} - \mathrm{proj}_{\vec{u}}\,\vec{v}\| = \frac{\sqrt{49 + 1 + 121}}{6} = \frac{\sqrt{171}}{6} \simeq 2,18.$$

Exercício 7. Como poderíamos fazer se a reta não passasse pela origem? O mesmo raciocínio funciona?

9.3 Bases ortogonais e ortonormais

Vimos que uma das importâncias do produto escalar é estender o conceito de ortogonalidade para espaços vetoriais mais gerais. A ortogonalidade traz consigo a noção de projeção ortogonal. Agora vamos ver algumas vantagens em se considerar bases de espaços vetoriais que são também conjuntos ortogonais; estas são conhecidas como **bases ortogonais**.

Exemplo 88. A base canônica de $\mathbb{R}^n$ forma um conjunto ortogonal e é, portanto, uma base ortogonal. Assim como no Exemplo 86, é possível representar qualquer vetor $\vec{v}$ como

$$\vec{v} = (\vec{v} \cdot \vec{e}_1)\vec{e}_1 + (\vec{v} \cdot \vec{e}_2)\vec{e}_2 + \cdots + (\vec{v} \cdot \vec{e}_n)\vec{e}_n.$$

[3]Lembre: a distância de um ponto a uma reta é a menor distância entre este ponto e os pontos da reta; por esta razão que procuramos um vetor perpendicular à reta.

Licença CC-BY-SA-3.0. Contato: livroscolaborativos@gmail.com

Vejamos que esta propriedade da base canônica é, na realidade, uma propriedade de todas as bases ortogonais:

Teorema 89. *Seja $\mathcal{B} = \{\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_n\}$ uma base ortogonal de $\mathbb{R}^n$. Então qualquer vetor $\vec{v} \in \mathbb{R}^n$ pode ser representado como*

$$\vec{v} = \frac{\vec{v} \cdot \vec{v}_1}{\vec{v}_1 \cdot \vec{v}_1}\, \vec{v}_1 + \frac{\vec{v} \cdot \vec{v}_2}{\vec{v}_2 \cdot \vec{v}_2}\, \vec{v}_2 + \cdots + \frac{\vec{v} \cdot \vec{v}_n}{\vec{v}_n \cdot \vec{v}_n}\, \vec{v}_n.$$

Podemos escrever, de outra maneira, que

$$\vec{v} = \operatorname{proj}_{\vec{v}_1} \vec{v} + \operatorname{proj}_{\vec{v}_2} \vec{v} + \cdots + \operatorname{proj}_{\vec{v}_n} \vec{v}.$$

Este resultado mostra que bases ortogonais são especiais e boas de trabalhar. Lembra que anteriormente, para encontrar os coeficientes de um vetor em uma base, tínhamos que encontrar $c_1, c_2, \cdots, c_n$ resolvendo o sistema linear cuja forma vetorial é

$$c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_n\vec{v}_n = \vec{v}.$$

O teorema acima afirma, em outras palavras, quando a base é ortogonal, os coeficientes do vetor $\vec{v}$ na base $\mathcal{B}$ são os números

$$c_j = \frac{\vec{v} \cdot \vec{v}_j}{\vec{v}_j \cdot \vec{v}_j}.$$

Justificativa do Teorema 89. Se a representação do vetor $\vec{v}$ na base $\mathcal{B}$ for

$$\vec{v} = c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_n\vec{v}_n,$$

então, fazendo o produto escalar desta expressão com $\vec{v}_j$ e utilizando as propriedades de ortogonalidade, concluimos que

$$\vec{v} \cdot \vec{v}_j = \big(c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_n\vec{v}_n\big) \cdot \vec{v}_j = c_1(\vec{v}_1 \cdot \vec{v}_j) + c_2(\vec{v}_2 \cdot \vec{v}_j) + \cdots + c_n(\vec{v}_n \cdot \vec{v}_j) = c_j\big(\vec{v}_j \cdot \vec{v}_j\big),$$

pois $\vec{v}_i \cdot \vec{v}_j = 0$ sempre que $i \neq j$. Daí, concluimos, como queríamos, que os coeficientes devem satisfazer

$$c_j = \frac{\vec{v} \cdot \vec{v}_j}{\vec{v}_j \cdot \vec{v}_j}.$$

□

Exemplo 90. Vamos justificar que os vetores

$$\vec{u} = \begin{bmatrix} 4 \\ 2 \end{bmatrix} \quad \text{e} \quad \vec{v} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$$

formam uma base ortogonal para $\mathbb{R}^2$.

Em primeiro lugar, formam uma base, pois são vetores linearmente independentes e são dois (que é a dimensão de $\mathbb{R}^2$). Agora, calculamos

$$\vec{u} \cdot \vec{v} = 4 \cdot (-1) + 2 \cdot 2 = 0.$$

Segue que $\vec{u}$ e $\vec{v}$ são também ortogonais e, logo, formam uma base ortogonal de $\mathbb{R}^2$.

Para escrever um vetor $\vec{x} = \begin{bmatrix} 2 \\ -7 \end{bmatrix}$ nesta base, podemos calcular as projeções nos elementos da base (isto apenas vale se os elementos da base são ortogonais!):

$$\vec{x} = \frac{\vec{x} \cdot \vec{u}}{\vec{u} \cdot \vec{u}}\, \vec{u} + \frac{\vec{x} \cdot \vec{v}}{\vec{v} \cdot \vec{v}}\, \vec{v} = \frac{-6}{20}\, \vec{u} + \frac{-16}{5}\, \vec{v} = -0.3\, \vec{u} - 3.2\, \vec{v}. \lhd$$

Observe no exemplo acima que é muito mais fácil obter as componentes por ortogonalidade do que resolver sistemas lineares. Isto mesmo em dimensão baixa! Imagine se fosse um sistema 5×5 ou ainda maior.

Note que a base canônica, como no Exemplo 88, tem a propriedade adicional que $\|\vec{e}_j\| = 1$ para todo índice j. Isto fez com que a representação de vetores do Teorema 89 assumisse um formato ainda mais simples. Estas bases ortogonais que tem por elementos vetores unitários

Licença CC-BY-SA-3.0. Contato: livroscolaborativos@gmail.com

são conhecidas como **bases ortonormais**. Mais explicitamente, uma base ortonormal é uma base $\{\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_n\}$ de um espaço vetorial que adicionalmente satisfaz

$$\vec{v}_i \cdot \vec{v}_j = 0 \text{ para } i \neq j \text{ e também } \|\vec{v}_j\| = 1 \text{ para todo } j \ \left(\iff \vec{v}_j \cdot \vec{v}_j = 1 \right)$$

Uma consequência imediata é que, sendo $\{\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_n\}$ uma base ortonormal de $\mathbb{R}^n$, podemos escrever

$$\vec{v} = (\vec{v} \cdot \vec{v}_1)\vec{v}_1 + (\vec{v} \cdot \vec{v}_2)\vec{v}_2 + \cdots + (\vec{v} \cdot \vec{v}_n)\vec{v}_n.$$

Exemplo 91. Os vetores

$$\vec{u} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \ \vec{v} = \begin{bmatrix} 0 \\ \sqrt{2}/2 \\ \sqrt{2}/2 \end{bmatrix}, \text{ e } \vec{w} = \begin{bmatrix} 0 \\ \sqrt{2}/2 \\ -\sqrt{2}/2 \end{bmatrix}$$

formam uma base ortonormal de $\mathbb{R}^3$: verifique! Qualquer vetor $\vec{x} \in \mathbb{R}^3$ pode ser escrito como

$$\vec{x} = (\vec{x} \cdot \vec{u})\vec{u} + (\vec{x} \cdot \vec{v})\vec{v} + (\vec{x} \cdot \vec{w})\vec{w}$$

Por exemplo,

$$\vec{x} = \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix} = 2\vec{u} + 0\vec{v} - \sqrt{2}\vec{w} = 2 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + \sqrt{2} \begin{bmatrix} 0 \\ \sqrt{2}/2 \\ -\sqrt{2}/2 \end{bmatrix}.$$

Observamos finalmente que bases ortogonais podem ser facilmente transformadas em bases ortonormais pelo processo de normalização.

Exemplo 92. Os vetores

$$\vec{u} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \ \vec{v} = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \text{ e } \vec{w} = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$$

formam uma base ortogonal de $\mathbb{R}^3$. Normalizando cada um dos vetores, podemos obter uma base ortonormal de $\mathbb{R}^3$:

$$\begin{cases} \|\vec{u}\| = \sqrt{1+1+1} = \sqrt{3} \\ \|\vec{v}\| = \sqrt{1+1} = \sqrt{2} \\ \|\vec{w}\| = \sqrt{1+4+1} = \sqrt{6} \end{cases}$$

Isto implica que os vetores

$$\frac{\vec{u}}{\|\vec{u}\|} = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}, \ \frac{\vec{v}}{\|\vec{v}\|} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{2} \\ 0 \\ -1/\sqrt{2} \end{bmatrix} \text{ e } \frac{\vec{w}}{\|\vec{w}\|} = \frac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{6} \\ -2/\sqrt{6} \\ 1/\sqrt{6} \end{bmatrix}$$

formam um base ortonormal de $\mathbb{R}^3$. $\triangleleft$

Licença CC-BY-SA-3.0. Contato: livroscolaborativos@gmail.com

10

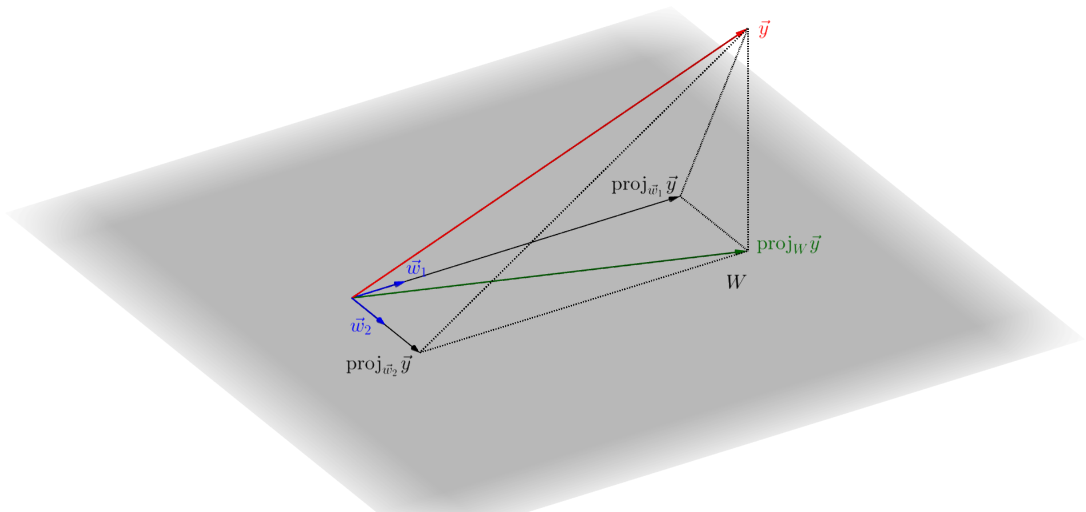
Semana 12

10.1 Projeção ortogonais sobre subespaços

No último capítulo, estudamos projeções ortogonais e vimos que a projeção ortogonal de $\vec{y}$ na direção de um vetor $\vec{w}$ pode ser calculada por

$$\operatorname{proj}_{\vec{w}} \vec{y} = \frac{\vec{y} \cdot \vec{w}}{\vec{w} \cdot \vec{w}}\, \vec{w}$$

Gostaríamos de endereçar agora a seguinte questão: como obter a projeção de um vetor $\vec{y} \in \mathbb{R}^3$ sobre um plano W?



Vamos trabalhar geometricamente: suponhamos que o plano

$$W = \operatorname{Span}\{\vec{w}_1, \vec{w}_2\}$$

e que os vetores $\vec{w}_1$ e $\vec{w}_2$ são ortogonais. Em outras palavras, $\{\vec{w}_1, \vec{w}_2\}$ é uma base ortogonal de W. Como veremos abaixo, o Processo de Gram-Schmidt fornece um algoritmo para obter uma base ortogonal de um espaço vetorial a partir de uma outra base qualquer de modo que sempre é possível considerar uma base ortogonal e, então, conseguimos calcular projeções.

A **projeção ortogonal sobre o subespaço** W, denotada por $\operatorname{proj}_W \vec{y}$, é definida como

$$\operatorname{proj}_W \vec{y} = \operatorname{proj}_{\vec{w}_1} \vec{y} + \operatorname{proj}_{\vec{w}_2} \vec{y} = \frac{\vec{y} \cdot \vec{w}_1}{\vec{w}_1 \cdot \vec{w}_1}\, \vec{w}_1 + \frac{\vec{y} \cdot \vec{w}_2}{\vec{w}_2 \cdot \vec{w}_2}\, \vec{w}_2.$$

Notamos que este vetor pertence a W, já que é uma combinação linear dos elementos de uma base de W. Mas mais fundamental é o fato de a projeção ser feita em uma direção ortogonal ao plano: passamos a verificar que

$$\vec{y} - \operatorname{proj}_W \vec{y} \;\text{ é ortogonal a } W.$$

Qualquer vetor de W pode ser escrito como

$$\vec{w} = \frac{\vec{w}\cdot\vec{w}_1}{\vec{w}_1\cdot\vec{w}_1}\,\vec{w}_1 + \frac{\vec{w}\cdot\vec{w}_2}{\vec{w}_2\cdot\vec{w}_2}\,\vec{w}_2.$$

Exercício 8. Utilize as propriedades do produto escalar e a ortogonalidade da base $\{\vec{w}_1,\vec{w}_2\}$ para verificar que

$$\vec{w}\cdot\left(\vec{y} - \mathrm{proj}_W\,\vec{y}\right) = 0.$$

Exemplo 93. Vamos calcular a projeção do vetor

$$\vec{y} = \begin{bmatrix}1\\1\\1\end{bmatrix}$$

sobre o plano gerado pelos vetores

$$\vec{w}_1 = \begin{bmatrix}4\\2\\-7\end{bmatrix} \quad \text{e} \quad \vec{w}_2 = \begin{bmatrix}1\\5\\2\end{bmatrix}.$$

Observamos inicialemente que *somos permitidos* de usar a fórmula acima, já que os vetores $\vec{w}_1$ e $\vec{w}_2$ formam uma base ortogonal de W:

$$\vec{w}_1\cdot\vec{w}_2 = 4 + 10 - 14 = 0.$$

Calculamos

$$\begin{aligned}\mathrm{proj}_W\,\vec{y} &= \frac{\vec{y}\cdot\vec{w}_1}{\vec{w}_1\cdot\vec{w}_1}\,\vec{w}_1 + \frac{\vec{y}\cdot\vec{w}_2}{\vec{w}_2\cdot\vec{w}_2}\,\vec{w}_2 = \frac{4+2-7}{16+4+49}\,\vec{w}_1 + \frac{1+5+2}{1+25+4}\,\vec{w}_2 \\ &= -\frac{1}{69}\begin{bmatrix}4\\2\\-7\end{bmatrix} + \frac{4}{15}\begin{bmatrix}1\\5\\2\end{bmatrix} = \begin{bmatrix}24/115\\30/23\\73/115\end{bmatrix}. \lhd\end{aligned}$$

Desenvolvemos acima o conceito de projeção ortogonal sobre um subespaço do espaço $\mathbb{R}^3$ de dimensão três apenas pela conveniência da visualização geométrica. Na verdade, o conceito pode ser definido em qualquer dimensão.

Seja $W \subset \mathbb{R}^n$ um subespaço vetorial de dimensão k e $\{\vec{w}_1,\vec{w}_2,\ldots,\vec{w}_k\}$ uma base ortogonal de W. Dado um vetor $\vec{y}\in\mathbb{R}^n$ qualquer, definimos a **projeção ortogonal de $\vec{y}$ sobre** W como

$$\mathrm{proj}_W\,\vec{y} = \frac{\vec{y}\cdot\vec{w}_1}{\vec{w}_1\cdot\vec{w}_1}\,\vec{w}_1 + \frac{\vec{y}\cdot\vec{w}_2}{\vec{w}_2\cdot\vec{w}_2}\,\vec{w}_2 + \cdots + \frac{\vec{y}\cdot\vec{w}_k}{\vec{w}_k\cdot\vec{w}_k}\,\vec{w}_k$$

Pode-se verificar, como no Exercício 8 acima, que

$$\vec{w}\cdot\left(\vec{y} - \mathrm{proj}_W\,\vec{y}\right) = 0, \text{ para qualquer } \vec{w}\in W,$$

isto é, a projeção $\mathrm{proj}_W\,\vec{y}$ pertence a W e $\vec{y} - \mathrm{proj}_W\,\vec{y}$ é ortogonal a todo elemento de W, assim como nossa intuição esperaria.

Exemplo 94. Calcular a projeção do vetor

$$\vec{y} = \begin{bmatrix}1\\0\\-1\\-3\\2\end{bmatrix}$$

sobre o espaço tridimensional W gerado pelos vetores

$$\vec{w}_1 = \begin{bmatrix}0\\4\\0\\6\\0\end{bmatrix}, \quad \vec{w}_2 = \begin{bmatrix}1\\0\\2\\0\\-1\end{bmatrix} \quad \text{e} \quad \vec{w}_3 = \begin{bmatrix}2\\-3\\0\\-2\\2\end{bmatrix}.$$

Licença CC-BY-SA-3.0. Contato: livroscolaborativos@gmail.com

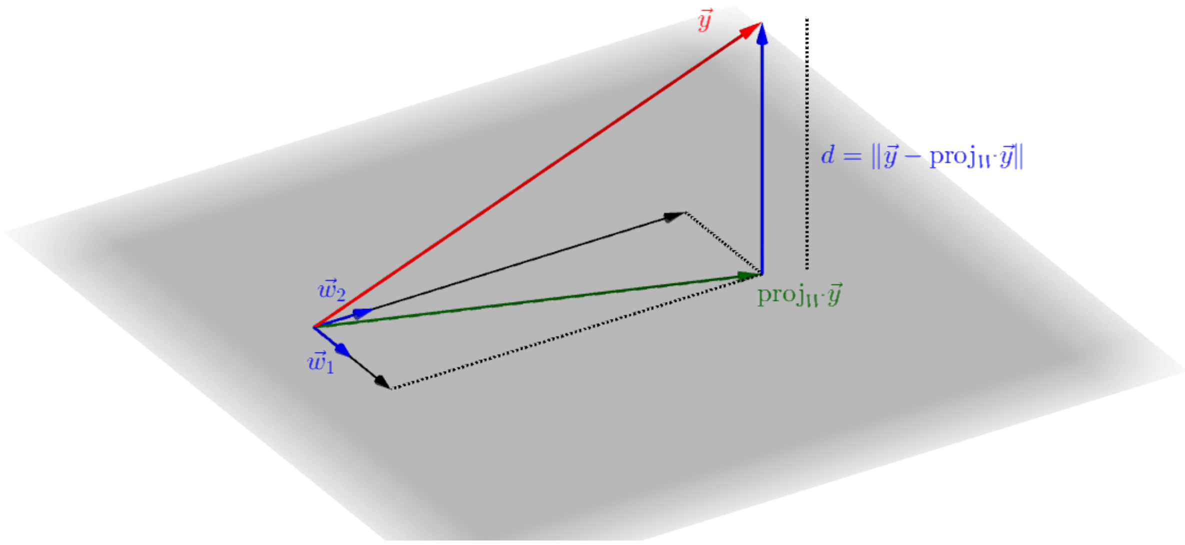
Observar que a base $\{\vec{w}_1, \vec{w}_2, \vec{w}_3\}$ de W é ortogonal! Devemos então calcular

$$\begin{aligned}\text{proj}_W \vec{y} &= \frac{\vec{y}\cdot\vec{w}_1}{\vec{w}_1\cdot\vec{w}_1}\,\vec{w}_1 + \frac{\vec{y}\cdot\vec{w}_2}{\vec{w}_2\cdot\vec{w}_2}\,\vec{w}_2 + \frac{\vec{y}\cdot\vec{w}_3}{\vec{w}_3\cdot\vec{w}_3}\,\vec{w}_3 \\ &= \frac{-18}{14+36}\,\vec{w}_1 + \frac{1-2-2}{1+4+1}\,\vec{w}_2 + \frac{2+6+4}{4+9+4+4}\,\vec{w}_3 \\ &= -\frac{9}{25}\begin{bmatrix}0\\4\\0\\6\\0\end{bmatrix} - \frac{1}{2}\begin{bmatrix}1\\0\\2\\0\\-1\end{bmatrix} + \frac{4}{7}\begin{bmatrix}2\\-3\\0\\-2\\2\end{bmatrix} = \begin{bmatrix}9/14\\-552/175\\-1\\-578/175\\23/14\end{bmatrix}.\end{aligned}$$

Confira estas contas com uma calculadora (os coeficientes dos vetores nem sempre são bonitinhos!)

10.2 Distância a subespaços e a melhor aproximação

A projeção ortogonal pode ser utilizada para calcular distâncias entre o pontos e subespaços.



Dado um ponto $P = (x_1, x_2, \ldots, x_n)$, formamos o vetor da origem ao ponto P

$$\vec{y} = \vec{OP} = \begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix}$$

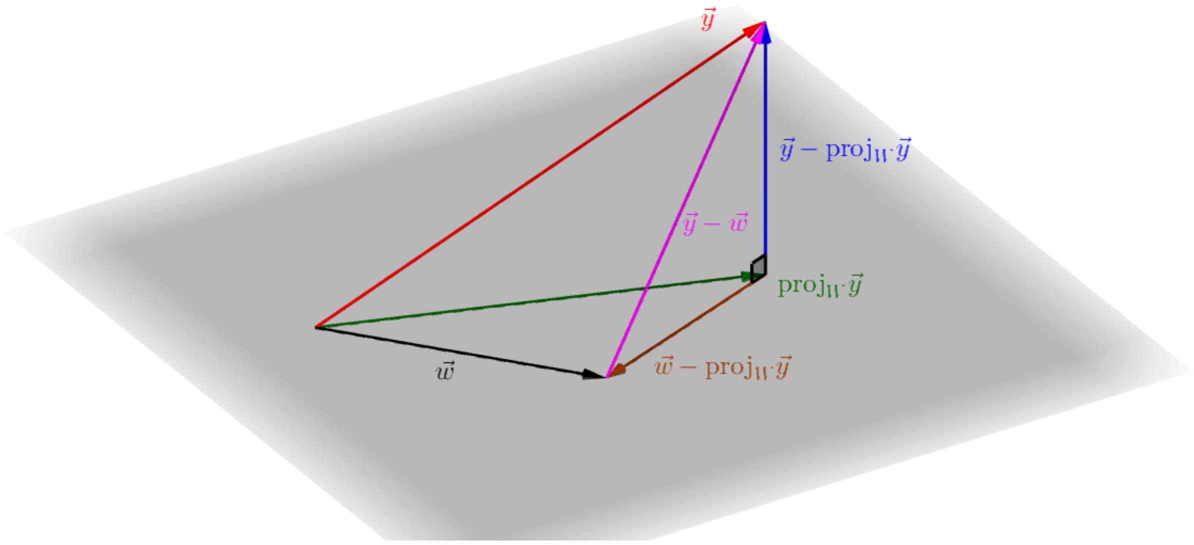
cujas componentes são as coordenadas de P. Como vimos na seção anterior, o vetor $\vec{y} - \text{proj}_W \vec{y}$ é ortogonal ao subespaço W. Sendo a distância de P até W denotada por $\text{dist}(P, W)$ e definida como a menor distância entre P e elementos de W, temos que

$$\text{dist}(P, W) = \|\vec{y} - \text{proj}_W \vec{y}\|.$$

Justificativa desta afirmação. Dado qualquer outro vetor $\vec{w} \in W$, o Teorema de Pitágoras implica que (ver figura)

$$\|\vec{y} - \vec{w}\|^2 = \|\vec{w} - \text{proj}_W \vec{y}\|^2 + \|\vec{y} - \text{proj}_W \vec{y}\|^2 > \|\vec{y} - \text{proj}_W \vec{y}\|^2.$$

Em outras palavras, $\|\vec{y} - \text{proj}_W \vec{y}\|$ é de fato a menor distância entre P e elementos de W. $\square$



Exemplo 95. Calcular a distância do ponto $P = (1, 2, -1, 0, 7)$ até o plano $W = \mathrm{Span}\{\vec{u}, \vec{v}\}$ (contido em $\mathbb{R}^5$) gerado pelos vetores (que são ortogonais!)

$$\vec{u} = \begin{bmatrix} 1 \\ 0 \\ -8 \\ 0 \\ 5 \end{bmatrix} \text{ e } \vec{v} = \begin{bmatrix} 1 \\ -1 \\ 2 \\ 0 \\ 3 \end{bmatrix}.$$

Para isto, consideramos o vetor

$$\vec{y} = \vec{OP} = \begin{bmatrix} 1 \\ 2 \\ -1 \\ 0 \\ 7 \end{bmatrix}$$

e vamos calcular $\|\vec{y} - \mathrm{proj}_W \vec{y}\|$:

$$\vec{y} - \mathrm{proj}_W \vec{y} = \vec{y} - \frac{\vec{y}\cdot\vec{u}}{\vec{u}\cdot\vec{u}}\,\vec{u} - \frac{\vec{y}\cdot\vec{v}}{\vec{v}\cdot\vec{v}}\,\vec{v} = \begin{bmatrix} 1 \\ 2 \\ -1 \\ 0 \\ 7 \end{bmatrix} - \frac{44}{90}\begin{bmatrix} 1 \\ 0 \\ -8 \\ 0 \\ 5 \end{bmatrix} - \frac{18}{15}\begin{bmatrix} 1 \\ -1 \\ 2 \\ 0 \\ 3 \end{bmatrix}$$

$$= \begin{bmatrix} 1 \\ 2 \\ -1 \\ 0 \\ 7 \end{bmatrix} - \frac{22}{45}\begin{bmatrix} 1 \\ 0 \\ 2 \\ 0 \\ 5 \end{bmatrix} - \frac{18}{15}\begin{bmatrix} 1 \\ -1 \\ 2 \\ 0 \\ 3 \end{bmatrix} = \begin{bmatrix} -31/45 \\ 16/5 \\ -197/45 \\ 0 \\ 61/45 \end{bmatrix}$$

Desta forma, a distância de P a W é

$$\mathrm{dist}(P, W) = \|\vec{y} - \mathrm{proj}_W \vec{y}\| = \sqrt{\frac{961}{2025} + \frac{256}{25} + \frac{38809}{2025} + \frac{3721}{2025}} = \sqrt{\frac{45795}{2025}}$$

$$= \frac{\sqrt{45795}}{45} \simeq 4,75. \triangleleft$$

10.3 O Processo de ortogonalização de Gram–Schmidt

O **Processo de Gram–Schmidt** é um algoritmo para obter uma base ortogonal (ou ortonormal) a partir de uma base qualquer. De maneira mais geral, o método permite transformar um

Licença CC-BY-SA-3.0. Contato: livroscolaborativos@gmail.com

conjunto de vetores linearmente independentes em um conjunto ortogonal que gera o mesmo espaço vetorial.

Vamos começar com um exemplo.

Exemplo 96. Consideramos os vetores

$$\vec{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \quad \text{e} \quad \vec{v}_3 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}.$$

O espaço gerado por $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ é todo o espaço $\mathbb{R}^3$, já que temos três vetores linearmente independentes em um espaço de dimensão três.

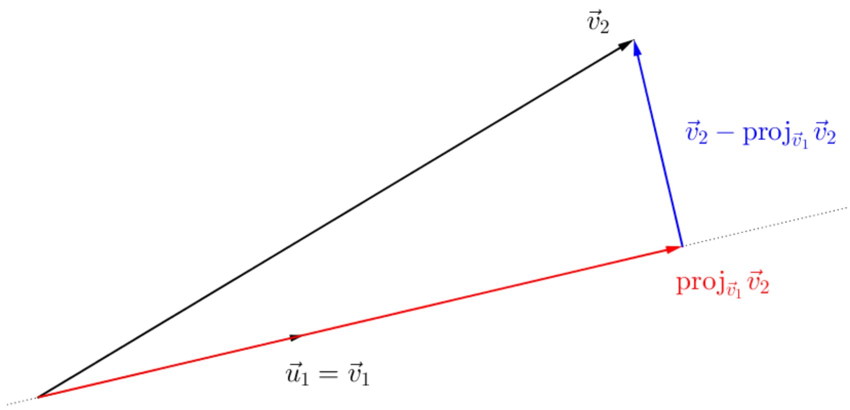
O processo consiste em fixar qualquer um dos vetores como o primeiro dos vetores do conjunto que virá a ser ortogonal. Por exemplo, fixamos

$$\vec{u}_1 = \vec{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}.$$

Em seguida, já vimos que, ao fazer a projeção de $\vec{v}_2$ sobre $\vec{v}_1$, o vetor

$$\vec{v}_2 - \operatorname{proj}_{\vec{v}_1} \vec{v}_2$$

é ortogonal a $\vec{v}_1$, como ilustra a figura.



Assim sendo, definimos o segundo vetor do nosso conjunto que será ortogonal como o vetor

$$\vec{u}_2 = \vec{v}_2 - \operatorname{proj}_{\vec{u}_1} \vec{v}_2 = \vec{v}_2 - \frac{\vec{v}_2 \cdot \vec{u}_1}{\vec{u}_1 \cdot \vec{u}_1} \vec{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1/2 \\ 1 \\ -1/2 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}.$$

Temos que $\vec{u}_1$ e $\vec{u}_2$ são ortogonais e que estão no mesmo plano, de modo que também temos

$$\operatorname{Span}\{\vec{u}_1, \vec{u}_2\} = \operatorname{Span}\{\vec{v}_1, \vec{v}_2\}.$$

Vamos escrever momentariamente $W = \operatorname{Span}\{\vec{u}_1, \vec{u}_2\}$.

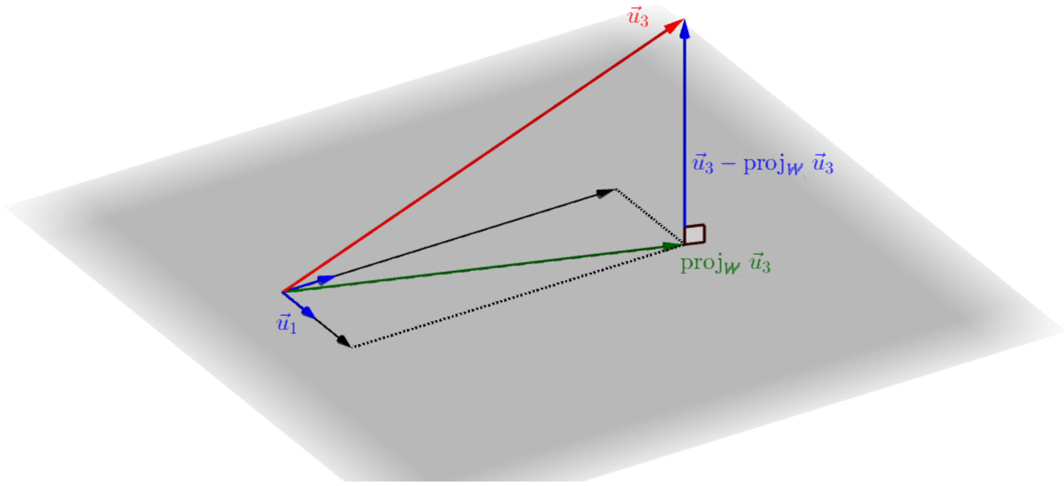
No próximo passo do processo, o terceiro vetor pode ser obtido como

$$\vec{u}_3 = \vec{v}_3 - \operatorname{proj}_W \vec{v}_3,$$

pois, desta forma, $\vec{u}_3$ é ortogonal a todos os vetores de W; em particular, é ortogonal a ambos $\vec{u}_1$ e $\vec{u}_2$. Além disso, como $\vec{u}_1$ e $\vec{u}_2$ já são vetores ortogonais, podemos calcular a projeção sobre W como de costume:

$$\vec{u}_3 = \vec{v}_3 - \operatorname{proj}_W \vec{v}_3 = \vec{v}_3 - \frac{\vec{v}_3 \cdot \vec{u}_1}{\vec{u}_1 \cdot \vec{u}_1} \vec{u}_1 - \frac{\vec{v}_3 \cdot \vec{u}_2}{\vec{u}_2 \cdot \vec{u}_2} \vec{u}_2.$$

Licença CC-BY-SA-3.0. Contato: livroscolaborativos@gmail.com



Calculando, temos (observe como dois fatores $1/2$ cancelam no último termo)

$$\vec{u}_3 = \begin{bmatrix}0\\1\\1\end{bmatrix} - \frac{1}{2}\begin{bmatrix}1\\0\\1\end{bmatrix} - \frac{2-1}{1+4+1}\begin{bmatrix}1\\2\\-1\end{bmatrix} = \begin{bmatrix}0\\1\\1\end{bmatrix} - \frac{1}{2}\begin{bmatrix}1\\0\\1\end{bmatrix} - \frac{1}{6}\begin{bmatrix}1\\2\\-1\end{bmatrix} = \begin{bmatrix}-2/3\\2/3\\2/3\end{bmatrix} = \frac{2}{3}\begin{bmatrix}-1\\1\\1\end{bmatrix}$$

Concluimos assim que o conjunto

$$\{\vec{u}_1, \vec{u}_2, \vec{u}_3\} = \left\{ \begin{bmatrix}1\\0\\1\end{bmatrix}, \begin{bmatrix}1/2\\1\\-1/2\end{bmatrix}, \begin{bmatrix}-2/3\\2/3\\2/3\end{bmatrix} \right\}$$

é ortogonal e gera o mesmo espaço que $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$.

Na realidade, se tivéssemos considerado múltiplos dos vetores acima, não perderíamos a propriedade de ortogonalidade, e temos

$$\text{Span}\{\vec{u}_1, \vec{u}_2, \vec{u}_3\} = \text{Span}\left\{ \begin{bmatrix}1\\0\\1\end{bmatrix}, \begin{bmatrix}1\\2\\-1\end{bmatrix}, \begin{bmatrix}-1\\1\\1\end{bmatrix} \right\}.$$

“Colocar em evidência” os fatores comuns desta forma e “desconsiderá-los” pode facilitar as contas em alguns casos.

Observamos também que, fazendo a normalização dos vetores, podemos obter uma base ortonormal:

$$\begin{cases} \|\vec{u}\| = \sqrt{1+0+1} = \sqrt{2} \\ \|\vec{v}\| = \sqrt{1+4+1} = \sqrt{6} \\ \|\vec{w}\| = \sqrt{1+1+1} = \sqrt{3} \end{cases}$$

e assim,

$$\frac{\vec{u}_1}{\|\vec{u}_1\|} = \frac{1}{\sqrt{3}}\begin{bmatrix}1\\0\\1\end{bmatrix} = \begin{bmatrix}1/\sqrt{2}\\0\\1/\sqrt{2}\end{bmatrix}, \quad \frac{\vec{u}_2}{\|\vec{u}_2\|} = \frac{1}{\sqrt{6}}\begin{bmatrix}1\\2\\-1\end{bmatrix} = \begin{bmatrix}1/\sqrt{6}\\2/\sqrt{6}\\-1/\sqrt{6}\end{bmatrix} \text{ e } \frac{\vec{u}_3}{\|\vec{u}_3\|} = \frac{1}{\sqrt{3}}\begin{bmatrix}-1\\1\\1\end{bmatrix} = \begin{bmatrix}-1/\sqrt{3}\\1/\sqrt{3}\\1/\sqrt{3}\end{bmatrix}$$

forma um conjunto ortonormal que gera o mesmo espaço vetorial que $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$. $\triangleleft$

Em seguida, mostramos como funciona o processo de maneira geral (um bom entendimento do exemplo acima deve deixar claro o porquê deste processo funcionar): sejam k vetores linearmente independentes $\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_k$ pertencentes ao espaço n-dimensional $\mathbb{R}^n$. Vamos denotar

$$W = \text{Span}\{\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_k\}$$

e construir uma base ortogonal (e em seguida ortonormal) para W. Começamos com

$$\boxed{\vec{u}_1 = \vec{v}_1}$$

e definimos

$$\boxed{\vec{u}_2 = \vec{v}_2 - \text{proj}_{\vec{u}_1} \vec{v}_2 = \vec{v}_2 - \frac{\vec{v}_2 \cdot \vec{u}_1}{\vec{u}_1 \cdot \vec{u}_1} \vec{u}_1.} \tag{10.1}$$

Temos que $\vec{u}_1$ e $\vec{u}_2$ são ortogonais e que estão no mesmo plano, de modo que também temos

$$\text{Span}\{\vec{u}_1, \vec{u}_2\} = \text{Span}\{\vec{v}_1, \vec{v}_2\} = W_2 \ \rightsquigarrow \ \text{notação.}$$

Em seguida, definimos

$$\boxed{\vec{u}_3 = \vec{v}_3 - \text{proj}_{W_2} \vec{v}_3 = \vec{v}_3 - \frac{\vec{v}_3 \cdot \vec{u}_1}{\vec{u}_1 \cdot \vec{u}_1} \vec{u}_1 - \frac{\vec{v}_3 \cdot \vec{u}_2}{\vec{u}_2 \cdot \vec{u}_2} \vec{u}_2.} \tag{10.2}$$

Temos que $\vec{u}_1, \vec{u}_2, \vec{u}_3$ ortogonais e

$$\text{Span}\{\vec{u}_1, \vec{u}_2, \vec{u}_3\} = \text{Span}\{\vec{v}_1, \vec{v}_2, \vec{v}_3\} = W_3 \ \rightsquigarrow \ \text{notação.}$$

Em seguida

$$\boxed{\vec{u}_4 = \vec{v}_4 - \text{proj}_{W_3} \vec{v}_4 = \vec{v}_4 - \frac{\vec{v}_4 \cdot \vec{u}_1}{\vec{u}_1 \cdot \vec{u}_1} \vec{u}_1 - \frac{\vec{v}_4 \cdot \vec{u}_2}{\vec{u}_2 \cdot \vec{u}_2} \vec{u}_2 - \frac{\vec{v}_4 \cdot \vec{u}_3}{\vec{u}_3 \cdot \vec{u}_3} \vec{u}_3.}$$

Temos que $\vec{u}_1, \vec{u}_2, \vec{u}_3, \vec{u}_4$ ortogonais e

$$\text{Span}\{\vec{u}_1, \vec{u}_2, \vec{u}_3, \vec{u}_4\} = \text{Span}\{\vec{v}_1, \vec{v}_2, \vec{v}_3, \vec{u}_4\} = W_4 \ \rightsquigarrow \ \text{notação.}$$

E assim seguimos fazendo as contas até chegar ao último vetor: sendo

$$\text{Span}\{\vec{u}_1, \vec{u}_2, \cdots, \vec{u}_{k-1}\} = \text{Span}\{\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_{k-1}\} = W_{k-1} \ \rightsquigarrow \ \text{notação,}$$

escrevemos

$$\boxed{\vec{u}_k = \vec{v}_k - \text{proj}_{W_{k-1}} \vec{v}_k = \vec{v}_k - \frac{\vec{v}_k \cdot \vec{u}_1}{\vec{u}_1 \cdot \vec{u}_1} \vec{u}_1 - \frac{\vec{v}_k \cdot \vec{u}_2}{\vec{u}_2 \cdot \vec{u}_2} \vec{u}_2 - \cdots - \frac{\vec{v}_k \cdot \vec{u}_{k-1}}{\vec{u}_{k-1} \cdot \vec{u}_{k-1}} \vec{u}_{k-1}.}$$

Temos que $\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_k\}$ é um conjunto ortogonal e

$$\text{Span}\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_k\} = \text{Span}\{\vec{v}_1, \vec{v}_2, \ldots, \vec{u}_k\} = W \ \rightsquigarrow \ \text{que é o espaço inicial,}$$

como queríamos.

Esta parte (obter o conjunto ortogonal) é a mais "pesada" do método. Por exemplo, se agora quiséssemos uma base ortonormal para W, basta considerar o conjunto

$$\left\{ \frac{\vec{u}_1}{\|\vec{u}_1\|}, \frac{\vec{u}_2}{\|\vec{u}_2\|}, \ldots, \frac{\vec{u}_k}{\|\vec{u}_k\|} \right\},$$

que continua sendo ortogonal, mas agora formado por vetores unitários.

Exemplo 97. [1] Vamos obter uma base ortonormal para o espaço coluna $\text{Col}\, A$ da matriz

$$A = \begin{bmatrix} 12 & -51 & 4 \\ 6 & 167 & -68 \\ -4 & 24 & -41 \end{bmatrix}.$$

Lembrando que $\text{Col}\, A$ é o espaço gerado pelas colunas de A (e observando que, neste caso, $\dim \text{Col}\, A = 3$), vamos aplicar o processo de Gram–Schmidt para os vetores

$$\vec{v}_1 = \begin{bmatrix} 12 \\ 6 \\ -4 \end{bmatrix}, \vec{v}_2 = \begin{bmatrix} -51 \\ 167 \\ 24 \end{bmatrix} \quad \text{e} \quad \vec{v}_3 = \begin{bmatrix} 4 \\ -68 \\ -41 \end{bmatrix}.$$

[1] A matriz A deste exemplo foi retirada de Wikipedia.org neste link permanente, em 21 de Novembro de 2017.

Licença CC-BY-SA-3.0. Contato: livroscolaborativos@gmail.com

O Processo começa como:

$$\vec{u}_1 = \vec{v}_1 = \begin{bmatrix} 12 \\ 6 \\ -4 \end{bmatrix},$$

$$\vec{u}_2 = \vec{v}_2 - \frac{\vec{v}_2 \cdot \vec{u}_1}{\vec{u}_1 \cdot \vec{u}_1}\vec{u}_1 = \begin{bmatrix} -51 \\ 167 \\ 24 \end{bmatrix} - \frac{-612+1002-96}{144+36+16}\begin{bmatrix} 12 \\ 6 \\ -4 \end{bmatrix} = \begin{bmatrix} -51 \\ 167 \\ 24 \end{bmatrix} - \frac{296}{196}\begin{bmatrix} 12 \\ 6 \\ -4 \end{bmatrix} = \begin{bmatrix} -69 \\ 158 \\ 30 \end{bmatrix}$$

Em seguida (conferir estas contas com o auxílio de uma calculadora)

$$\vec{u}_3 = \vec{v}_3 - \frac{\vec{v}_3 \cdot \vec{u}_1}{\vec{u}_1 \cdot \vec{u}_1}\vec{u}_1 - \frac{\vec{v}_3 \cdot \vec{u}_2}{\vec{u}_2 \cdot \vec{u}_2}\vec{u}_2 = \begin{bmatrix} 4 \\ -68 \\ -41 \end{bmatrix} - \frac{48-408+164}{196}\begin{bmatrix} 12 \\ 6 \\ -4 \end{bmatrix} - \frac{-276-10744-1230}{4761+24694+900}\begin{bmatrix} -69 \\ 158 \\ 30 \end{bmatrix}$$
$$= \begin{bmatrix} 4 \\ -68 \\ -41 \end{bmatrix} + \frac{\cancel{196}}{\cancel{196}}\begin{bmatrix} 12 \\ 6 \\ -4 \end{bmatrix} + \frac{12250}{30355}\begin{bmatrix} -69 \\ 158 \\ 30 \end{bmatrix} = \begin{bmatrix} -58/5 \\ 6/5 \\ -33 \end{bmatrix}.$$

Para obter uma base ortonormal para o espaço coluna, vamos ainda considerar

$$\frac{\vec{u}_1}{\|\vec{u}_1\|}, \frac{\vec{u}_2}{\|\vec{u}_2\|}, \frac{\vec{u}_3}{\|\vec{u}_3\|}.$$

Calculamos

$$\begin{cases} \|\vec{u}_1\| = \sqrt{196} = 14 \\ \|\vec{u}_2\| = \sqrt{30625} = 175 \\ \|\vec{u}_3\| = \dfrac{\sqrt{30625}}{5} = 35 \end{cases}$$

Portanto, uma base ortonormal para $\operatorname{Col} A$ é

$$\vec{n}_1 = \frac{\vec{u}_1}{\|\vec{u}_1\|} = \begin{bmatrix} 6/7 \\ 3/7 \\ -2/7 \end{bmatrix}, \quad \vec{n}_2 = \frac{\vec{u}_2}{\|\vec{u}_2\|} = \begin{bmatrix} -69/175 \\ 158/175 \\ 6/35 \end{bmatrix} \quad \text{e} \quad \vec{n}_3 = \frac{\vec{u}_3}{\|\vec{u}_3\|} = \begin{bmatrix} -58/175 \\ 6/175 \\ -33/35 \end{bmatrix}. \triangleleft$$

10.4 Fatoração QR e matrizes ortogonais

Como passamos a discutir abaixo, o método de Gram-Schmidt permite fatorar uma matriz A cujas colunas são linearmente independentes em um produto

$$A = QR,$$

onde Q é uma matriz cujas colunas formam um conjunto ortonormal e R é uma matriz triangular superior. Tal fatoração é muito importante do ponto de vista computacional, como veremos em breve, quando usarmos a fatoração QR como uma alternativa para resolução de sistemas lineareas por mínimos quadrados. Também é a base de um eficiente método numérico para encontrar autovalores de uma matriz quadrada.

Matrizes cujas colunas formam um conjunto ortonormal são muito úteis e por isso recebem um nome especial: são chamadas de **matrizes ortogonais**. Aqui cabe uma observação relevante: este é um exemplo de uma situação onde a terminologia clássica está longe de ser satisfatória: matrizes cujas colunas formam um conjunto *ortonormal* são chamadas de matrizes *ortogonais*.

Faremos agora uma exposição da fatoração QR usando uma matriz inicial A que tem 3 colunas, observando que a construção abaixo é facilmente generalizada para uma matriz de qualquer dimensão (desde que tenha colunas LI).

Dada uma matriz $A = \begin{bmatrix} \vec{v}_1 & \vec{v}_2 & \vec{v}_3 \end{bmatrix}$ cujas colunas são LI e são denotadas pelos vetores $\vec{v}_1$, $\vec{v}_2$, e $\vec{v}_3$, o processo de Gram-Schmidt começa com a construção de uma base ortogonal formada pelos vetores $\vec{u}_1$, $\vec{u}_2$, e $\vec{u}_3$, para o espaço-coluna da matriz A. Para isso, definimos

Licença CC-BY-SA-3.0. Contato: livroscolaborativos@gmail.com

$$\begin{cases} \vec{u}_1 = \vec{v}_1 \\ \vec{u}_2 = \vec{v}_2 - \operatorname{proj}_{\vec{u}_1} \vec{v}_2 \\ \vec{u}_3 = \vec{v}_3 - \operatorname{proj}_{\operatorname{Span}\{\vec{u}_1, \vec{u}_2\}} \vec{v}_3 \end{cases}$$

Desta forma, existem coeficientes b_1, c_1 e $c_2 \in \mathbb{R}$ tais que

$$\begin{cases} \vec{u}_1 = \vec{v}_1 \\ \vec{u}_2 = \vec{v}_2 - b_1\vec{u}_1 \\ \vec{u}_3 = \vec{v}_3 - c_1\vec{u}_1 - c_2\vec{u}_2 \end{cases}.$$

Estes coeficientes são justamente aqueles que aparecem nas fórmulas (10.1) e (10.2). Apenas denotamos desta forma para não poluir as equações e facilitar o entendimento.

Em um segundo momento, cada um dos três vetores $\vec{u}_1$, $\vec{u}_2$, e $\vec{u}_3$ pode ser normalizado, ou seja, podemos dividir cada um deles pelo seu comprimento para obter uma base ortonormal. Assim, $\vec{u}_1$, $\vec{u}_2$, e $\vec{u}_3$ formam uma base ortonormal para o espaço coluna de A, e existem coeficientes α_1, β_1, β_2, γ_1, γ_2 e $\gamma_3 \in \mathbb{R}$ tais que[2]

$$\begin{cases} \vec{u}_1 = \alpha_1\vec{v}_1 \\ \vec{u}_2 = \beta_2\vec{v}_2 - \beta_1\vec{u}_1 \\ \vec{u}_3 = \gamma_3\vec{v}_3 - \gamma_1\vec{u}_1 - \gamma_2\vec{u}_2 \end{cases}$$

Podemos agora inverter estas equações (que são do tipo triangular – faça as contas!) para escrever as colunas da matriz A em função dos vetores da base ortonormal $\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$ que acaba de ser obtida:

$$\begin{cases} \vec{v}_1 = \alpha_1^{-1}\vec{u}_1 \\ \vec{v}_2 = \beta_2^{-1}\beta_1\vec{u}_1 + \beta_2^{-1}\vec{u}_2 \\ \vec{v}_3 = \gamma_3^{-1}\gamma_1\vec{u}_1 + \gamma_3^{-1}\gamma_2\vec{u}_2 + \gamma_3^{-1}\vec{u}_3 \end{cases}$$

Matricialemente, estas equações podem ser escritas como

$$\begin{bmatrix} | & | & | \\ \vec{v}_1 & \vec{v}_2 & \vec{v}_3 \\ | & | & | \end{bmatrix} = \begin{bmatrix} | & | & | \\ \vec{u}_1 & \vec{u}_2 & \vec{u}_3 \\ | & | & | \end{bmatrix} \begin{bmatrix} \alpha_1^{-1} & \beta_2^{-1}\beta_1 & \gamma_3^{-1}\gamma_1 \\ 0 & \beta_2^{-1} & \gamma_3^{-1}\gamma_2 \\ 0 & 0 & \gamma_3^{-1} \end{bmatrix},$$

visto que cada coluna do produto da direita é definido como a combinação linear das colunas da primeira matriz com pesos dados pela respectiva coluna da segunda matriz.[3] Assim, definindo

$$Q = \begin{bmatrix} | & | & | \\ \vec{u}_1 & \vec{u}_2 & \vec{u}_3 \\ | & | & | \end{bmatrix} \quad \text{e} \quad R = \begin{bmatrix} \alpha_1^{-1} & \beta_2^{-1}\beta_1 & \gamma_3^{-1}\gamma_1 \\ 0 & \beta_2^{-1} & \gamma_3^{-1}\gamma_2 \\ 0 & 0 & \gamma_3^{-1} \end{bmatrix},$$

obtemos

$$A = QR,$$

que é o que queríamos obter.

Note também que R é invertível, já que sua diagonal principal contém apenas termos não-nulos.

A construção acima é facilmente generalizada para matrizes quaisquer e nos permite enunciar o seguinte fato (lembre que uma matriz ortogonal é uma matriz cujas colunas formam um conjunto ortonormal):

[2] Para simplificar a notação, continuamos escrevendo $\vec{u}_i$ quando, na realidade, deveríamos escrever $\vec{u}_i/\|\vec{u}_i\|$. Observe também que cada um dos coeficientes é obtido simplesmente pela normalização dos vetores. Mais explicitamente, temos

$$\alpha_1 = \frac{1}{\|\vec{u}_1\|}, \quad \beta_2 = \frac{1}{\|\vec{u}_2\|}, \quad \beta_1 = \frac{b_1}{\|\vec{u}_2\|}, \quad \gamma_3 = \frac{1}{\|\vec{u}_3\|}, \quad \gamma_1 = \frac{c_1}{\|\vec{u}_3\|} \quad \text{e} \quad \gamma_2 = \frac{c_2}{\|\vec{u}_3\|}.$$

[3] Comparar, por exemplo, com as Subseções 4.1.2 e 4.3.2, onde este tipo de interpretação para o produto de matrizes foi discutido.

Licença CC-BY-SA-3.0. Contato: livroscolaborativos@gmail.com

Teorema 98. *Qualquer matriz A cujas colunas são linearmente independentes pode ser fatorada na forma $A = QR$, onde Q é uma matriz ortogonal e R é uma matriz triangular superior invertível. Ainda, o espaço coluna das matrizes A e Q é o mesmo, ou seja, as colunas de Q formam uma base ortonormal para o subespaço gerado pelas colunas de A.*

Note que a fatoração QR está para o processo de Gram-Schmidt assim como a fatoração LU está para o método de eliminação gaussiana à forma escada

Bem, vimos que a matriz Q tem como colunas a base ortonormal gerada pelo processo de Gram-Schmidt. E como podemos calcular a matriz R no caso geral ?

Usando o seguinte fato surpreendente:

Teorema 99. *Uma matriz ortogonal é sempre invertível e sua inversa é dada pela sua transposta, ou seja $Q^{-1} = Q^T$.*

Demonstração. A prova deste teorema é uma simples verificação:

$$Q^TQ = \begin{bmatrix} — & \vec{u}_1 & — \\ — & \vec{u}_2 & — \\ — & \vec{u}_3 & — \end{bmatrix} \begin{bmatrix} | & | & | \\ \vec{u}_1 & \vec{u}_2 & \vec{u}_3 \\ | & | & | \end{bmatrix} = \begin{bmatrix} \vec{u}_1 \cdot \vec{u}_1 & \vec{u}_1 \cdot \vec{u}_2 & \vec{u}_1 \cdot \vec{u}_2 \\ \vec{u}_2 \cdot \vec{u}_1 & \vec{u}_2 \cdot \vec{u}_2 & \vec{u}_2 \cdot \vec{u}_3 \\ \vec{u}_3 \cdot \vec{u}_1 & \vec{u}_3 \cdot \vec{u}_2 & \vec{u}_3 \cdot \vec{u}_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = I. \qquad \square$$

Assim, se $A = QR$, basta multiplicar ambos os lados desta equação por Q^T pela esquerda para obtermos:

$$Q^TA = Q^TQR \implies R = Q^TA.$$

No último exemplo da seção anterior, obtemos através do método de Gram-Schmidt uma base ortonormal para $\operatorname{Col} A$ onde

$$A = \begin{bmatrix} 12 & -51 & 4 \\ 6 & 167 & -68 \\ -4 & 24 & -41 \end{bmatrix}.$$

Vamos definir a matriz Q como a matriz formada pelos vetores desta base ortonormal:

$$Q = \begin{bmatrix} 6/7 & -69/175 & -58/175 \\ 3/7 & 158/175 & 6/175 \\ -2/7 & 6/35 & -33/35 \end{bmatrix}$$

Vimos que $R = Q^TA$ e portanto

$$R = Q^TA = \begin{bmatrix} 6/7 & 3/7 & -2/7 \\ -69/175 & 158/175 & 6/35 \\ -58/175 & 6/175 & -33/35 \end{bmatrix} \begin{bmatrix} 12 & -51 & 4 \\ 6 & 167 & -68 \\ -4 & 24 & -41 \end{bmatrix} = \begin{bmatrix} 14 & 21 & -14 \\ 0 & 175 & -70 \\ 0 & 0 & 35 \end{bmatrix}.$$

Finalmente podemos escrever

$$A = \begin{bmatrix} 12 & -51 & 4 \\ 6 & 167 & -68 \\ -4 & 24 & -41 \end{bmatrix} = \begin{bmatrix} 6/7 & -69/175 & -58/175 \\ 3/7 & 158/175 & 6/175 \\ -2/7 & 6/35 & -33/35 \end{bmatrix} \begin{bmatrix} 14 & 21 & -14 \\ 0 & 175 & -70 \\ 0 & 0 & 35 \end{bmatrix} = QR. \lhd$$

10.5 Matrizes Ortogonais preservam comprimento e ângulo

Vamos aproveitar para falar um pouco mais sobre matrizes ortogonais. Um exemplo clássico de matriz ortogonal de ordem 2×2 é dada pela matriz de rotação (esta no sentido anti-horário, ver Exemplo 20)

$$Q = \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix},$$

onde $0 \leq \theta \leq 2\pi$. Podemos ver que esta matriz define uma transformação linear no plano que rotaciona vetores em um ângulo θ no sentido anti-horário. Assim, tal transformação preserva comprimento e ângulo entre vetores. Estas propriedades não são particulares da matriz de rotação: todas as matrizes ortogonais se comportam da mesma forma. O teorema a seguir traz estas informações e mais algumas.

Teorema 100. *Seja* $Q = \begin{bmatrix} \vec{u}_1 & \vec{u}_2 & \cdots & \vec{u}_n \end{bmatrix}$ *uma matriz ortogonal de ordem* $n \times n$. *Então, para quaisquer vetores* $\vec{v}$ *e* $\vec{w}$ *em* $\mathbb{R}^n$ *temos*

(i) $Q\vec{v} \cdot Q\vec{w} = \vec{v} \cdot \vec{w}$

(ii) $\|Q\vec{v}\| = \|\vec{v}\|$

(iii) *o ângulo entre* $Q\vec{v}$ *e* $Q\vec{w}$ *é igual ao ângulo entre* $\vec{v}$ *e* $\vec{w}$

Demonstração. A prova do primeiro item é consequência direta da linearidade do produto interno e da hipótese das colunas formarem uma base ortonormal: se, na base canônica, temos $\vec{v} = (\alpha_1, ..., \alpha_n)$ e $\vec{w} = (\beta_1, ..., \beta_n)$, então

$$Q\vec{v} \cdot Q\vec{w} = \left(\sum_{i=1}^{n} \alpha_i \vec{u}_i \right) \cdot \left(\sum_{j=1}^{n} \beta_j \vec{u}_j \right) = \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \beta_j \; \vec{u}_i \cdot \vec{u}_j$$

devido a linearidade do produto interno (ou seja, a propriedade distributiva aplicada tanto na primeira quanto na segunda componente do produto interno). Agora, lembrando que $\vec{u}_i \cdot \vec{u}_j = 0$ se $i \neq j$, a ultima expressão iguala-se a

$$\sum_{i=1}^{n} \alpha_i \beta_i \; \vec{u}_i \cdot \vec{u}_i.$$

Finalmente, lembrando que cada coluna $\vec{u}_i$ de Q tem comprimento unitário, concluimos que

$$Q\vec{v} \cdot Q\vec{w} = \sum_{i=1}^{n} \alpha_i \beta_i = \vec{v} \cdot \vec{w}.$$

Os próximos dois itens seguem diretamente do primeiro item, visto que

$$\|Q\vec{v}\| = \sqrt{Q\vec{v} \cdot Q\vec{v}} = \sqrt{\vec{v} \cdot \vec{v}} = \|\vec{v}\|$$

e o ângulo entre Qv e Qw é definido como

$$\arccos \left(\frac{Qv \cdot Qw}{\|Qv\| \|Qw\|} \right) = \arccos \left(\frac{v \cdot w}{\|v\| \|w\|} \right),$$

que por definição é o ângulo entre v e w. Assim, finalizamos a prova do teorema. □

Outros exemplos de matrizes ortogonais são obtidos por matrizes de reflexão. Um caso particular é dado pela reflexão em torno do eixo horizontal:

$$Q = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

Exercício 9. Como é a matriz de ordem 2×2 que representa uma reflexão com respeito ao eixo vertical?

Exercício 10. Encontrar uma matriz de ordem 2×2 que representa uma reflexão com respeito à reta que passa pela origem de equação $ax + by = 0$.

Licença CC-BY-SA-3.0. Contato: livroscolaborativos@gmail.com

11

Semana 13

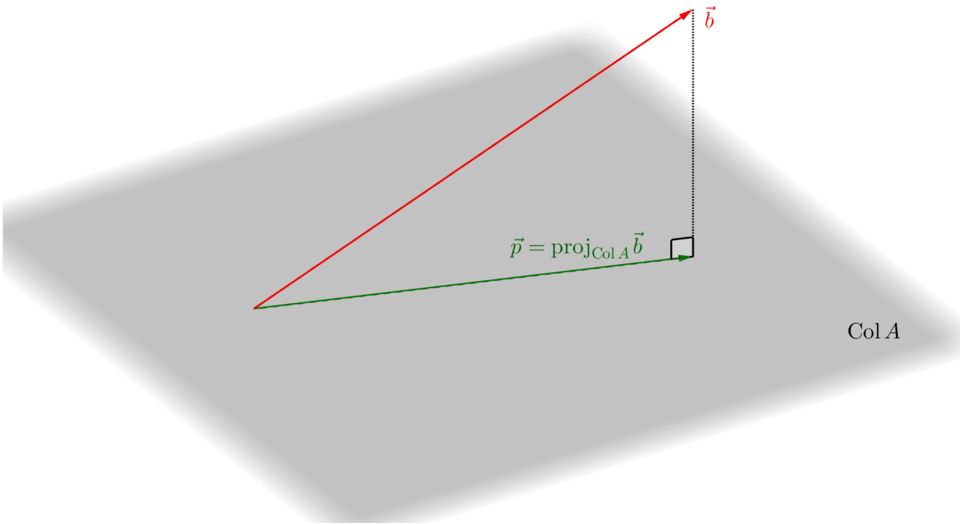
11.1 Método dos mínimos quadrados

Sistemas lineares aparecem como modelos matemáticos de vários fenômenos e em várias situações. Acontece que alguns sistemas simplesmente não possuem soluções e ficamos sem saber como proceder. O **método dos mínimos quadrados** é uma técnica que nos permite, de forma aproximada, retirar alguma informação desses sistemas impossíveis. A terminologia se deve ao fato de que, como veremos, este método minimiza a soma dos quadrados dos erros obtidos na aproximação.

Como sabemos, resolver o sistema linear

$$A\vec{x} = \vec{b}$$

consiste em encontrar um vetor $\vec{x}$ que satisfaça a esta equação. Na terminologia que construímos ao longo do curso, é equivalente dizer que $\vec{b}$ pertence ao espaço coluna da matriz A, isto é, $\vec{b} \in \operatorname{Col} A$. Desta forma, não é possível resolver o sistema quando $\vec{b} \notin \operatorname{Col} A$. Representamos esta situação na figura abaixo.



O método dos mínimos quadrados consiste em olhar para o vetor $\vec{p} = \operatorname{proj}_{\operatorname{Col} A} \vec{b}$ e resolver o sistema linear associado

$$A\vec{x} = \vec{p}.$$

Esta solução de $A\vec{x} = \vec{p}$ é chamada de **solução de mínimos quadrados**. A ideia é (ver figura) considerar o vetor pertencente a $\operatorname{Col} A$ que é o "mais próximo" possível de $\vec{b}$ e cujo sistema linear associado possua solução. Neste contexto, estamos pensando em mais próximo no sentido

$$\|\vec{b} - \operatorname{proj}_{\operatorname{Col} A} \vec{b}\| \leq \|\vec{b} - \vec{c}\|, \text{ para qualquer } \vec{c} \in \operatorname{Col} A,$$

isto é, no sentido de ser a projeção a melhor aproximação de $\vec{b}$ no espaço coluna de A. Escrevendo os vetores em coordenadas:

$$\vec{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} \quad \text{e} \quad \text{proj}_{\text{Col}\, A}\, \vec{b} = \begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_n \end{bmatrix},$$

podemos definir o valor

$$\|\vec{b} - \text{proj}_{\text{Col}\, A}\, \vec{b}\| = \sqrt{\sum_{i=1}^{n} (b_i - p_i)^2}$$

como o **erro da aproximação** ou **erro quadrático**. Logo, como anunciado, a soma dos quadrados dos erros obtidos cada componente é o mínimo possível.

Exemplo 101. Considere o sistema linear

$$\begin{cases} x_1 + x_2 - 2x_3 = 3 \\ x_1 - 2x_3 = 2 \\ x_2 + 2x_3 = 0 \\ x_1 + x_2 - 2x_3 = 0 \end{cases} \quad \leftrightsquigarrow \quad A\vec{x} = \begin{bmatrix} 1 & 1 & -2 \\ 1 & 0 & -2 \\ 0 & 1 & 2 \\ -1 & -1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \\ 0 \\ 0 \end{bmatrix} = \vec{b}.$$

Este sistema é claramente impossível, pois a primeira equação é inconsistente com a última. De forma mais automática, um passo no escalonamento da matriz aumentada associada revelaria a mesma conclusão:

$$[\, A \mid \vec{b}\,] = \begin{bmatrix} 1 & 1 & -2 & 3 \\ 1 & 0 & -2 & 2 \\ 0 & 1 & 2 & 0 \\ -1 & -1 & 2 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & -2 & 3 \\ 1 & 0 & -2 & 2 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}.$$

Vamos tentar encontrar, conforme descrito acima, uma solução de mínimos quadrados. Passos a serem realizados:

- Para calcular a projeção de $\vec{b}$ sobre $\text{Col}\, A$, precisamos, em primeiro lugar, obter uma base ortogonal do espaço $\text{Col}\, A$; isto pode ser feito pelo processo de Gram–Schmidt;
- Calcular a projeção $\vec{p} = \text{proj}_{\text{Col}\, A}\, \vec{b}$;
- Resolver o sistema linear $A\vec{x} = \vec{p}$.

Embora seja possível realizar estes três passos, temos de convir que seria muito trabalhoso. Por isto, vamos analisar um pouco mais da teoria para encontrar um método mais eficaz. $\lhd$

Suponhamos que:

- A é uma matriz $m \times n$,
- $\vec{x} \in \mathbb{R}^n$,
- $\vec{b} \in \mathbb{R}^m$ e
- o sistema $A\vec{x} = \vec{b}$ não possui solução.

A solução de mínimos quadrados é a solução de $A\vec{x} = \vec{p}$, onde $\vec{p} = \text{proj}_{\text{Col}\, A}\, \vec{b}$. Observamos, por outro lado, que, sendo $\vec{p}$ a projeção sobre o espaço coluna de A, o vetor $\vec{b} - \vec{p}$ deve ser ortogonal a todos os elementos de $\text{Col}\, A$. Em particular, se escrevermos A em termos de suas colunas

$$A = \begin{bmatrix} | & | & & | \\ \vec{a}_1 & \vec{a}_2 & \cdots & \vec{a}_n \\ | & | & & | \end{bmatrix}, \qquad \left(\text{que, em particular, implica } A^T = \begin{bmatrix} — & \vec{a}_1 & — \\ — & \vec{a}_2 & — \\ & \vdots & \\ — & \vec{a}_n & — \end{bmatrix} \right)$$

devemos ter

$$\vec{a}_j \cdot (\vec{b} - \vec{p}) = 0, \text{ para todo } j \quad \iff \quad A^T(\vec{b} - \vec{p}) = \vec{0} \quad \iff \quad A^T\vec{b} = A^T\vec{p} = A^TA\vec{x}.$$

Esta última linha é válida porque o produto escalar $\vec{a}_j \cdot (\vec{b} - \vec{p})$ é justamente a entrada j do vetor obtido ao multiplicar $A^T(\vec{b} - \vec{p})$. Além disso, $A\vec{x} = \vec{p}$.

Concluimos, desta forma, que, se $\vec{x}$ é uma solução de mínimos quadrados, ou seja, se $A\vec{x} = \vec{p}$, então necessariamente

$$\boxed{A^TA\vec{x} = A^T\vec{b}.} \tag{11.1}$$

Exercício 11 (Teórico)**.** Justifique que o conjunto de soluções do sistema (11.1) coincide com o conjunto das soluções de mínimos quadrados do sistema $A\vec{x} = \vec{b}$.

Tudo isto implica que podemos utilizar a equação (11.1) para encontrar as soluções de mínimos quadrados de $A\vec{x} = \vec{b}$.

Exemplo 102 (De volta ao Exemplo 101)**.** Calculamos

$$A^TA = \begin{bmatrix} 1 & 1 & 0 & -1 \\ 1 & 0 & 1 & -1 \\ -2 & -2 & 2 & 2 \end{bmatrix} \begin{bmatrix} 1 & 1 & -2 \\ 1 & 0 & -2 \\ 0 & 1 & 2 \\ -1 & -1 & 2 \end{bmatrix} = \begin{bmatrix} 3 & 2 & -6 \\ 2 & 3 & -2 \\ -6 & -2 & 16 \end{bmatrix}$$

e

$$A^T\vec{b} = \begin{bmatrix} 1 & 1 & 0 & -1 \\ 1 & 0 & 1 & -1 \\ -2 & -2 & 2 & 2 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 5 \\ 3 \\ -10 \end{bmatrix}$$

Para resolvermos o sistema $A^TA\vec{x} = A^T\vec{b}$, vamos escalonar a matriz aumentada associada:

$$\begin{bmatrix} 3 & 2 & -6 & 5 \\ 2 & 3 & -2 & 3 \\ -6 & -2 & 16 & -10 \end{bmatrix} \sim \begin{bmatrix} 3 & 2 & -6 & 5 \\ 0 & 5 & 6 & -1 \\ 0 & 2 & 4 & 0 \end{bmatrix} \sim \begin{bmatrix} 3 & 2 & -6 & 5 \\ 0 & 5 & 6 & -1 \\ 0 & 0 & 4 & 1 \end{bmatrix}$$

Logo,

$$\begin{cases} 3x_1 + 2x_2 - 6x_3 = 5 \\ 5x_2 + 6x_3 = -1 \\ 4x_3 = 1 \end{cases}.$$

Assim,

$$x_3 = \frac{1}{4} \implies 5x_2 + 6 \cdot \frac{1}{4} = -1 \implies x_2 = -\frac{1}{2} \implies 3x_1 - 1 - \frac{3}{2} = 5 \implies x_1 = \frac{3}{2}.$$

Portanto, uma solução de mínimos quadrados é

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 3/2 \\ -1/2 \\ 1/4 \end{bmatrix}.$$

Observe como isto é muito menos trabalhoso do que o método que havíamos esquematizado no Exemplo 101. $\triangleleft$

Exercício 12. Colocar em prática o método discutido no Exemplo 101 e comparar o resultado com o que obtivemos no Exemplo 102.

Observação 103. Caso a matriz A já seja uma matriz ortogonal, podemos calcular projeções diretamente, pois uma base ortogonal para $\operatorname{Col} A$ já esté disponível desde o começo. Neste caso, ambos os métodos exigiriam um trabalho parecido.

Nossa aproximação linear foi encontrada de modo a minimizar o erro quadrático. Para encontrar o erro, deveríamos calcular

$$\|\vec{b} - \text{proj}_{\text{Col}\, A}\, \vec{b}\|.$$

No entanto, note que não precisamos calcular a projeção. Observando que a solução de mínimos quadrados satisfaz $A\vec{x} = \text{proj}_{\text{Col}\, A}\, \vec{b}$, podemos calcular o erro por

$$\text{erro} \;= \|\vec{b} - A\vec{x}\|,$$

onde $\vec{x}$ é a solução de mínimos quadrados.

No Exemplo 102 acima, podemos calcular o erro desta forma. De fato,

$$A\vec{x} = \begin{bmatrix} 1 & 1 & -2 \\ 1 & 0 & -2 \\ 0 & 1 & 2 \\ -1 & -1 & 2 \end{bmatrix} \begin{bmatrix} 3/2 \\ -1/2 \\ 1/4 \end{bmatrix} = \begin{bmatrix} 1/2 \\ 1 \\ 0 \\ -1/2 \end{bmatrix} \implies \vec{b} - A\vec{x} = \begin{bmatrix} 5/2 \\ 1 \\ 0 \\ 1/2 \end{bmatrix}$$

$$\implies \text{erro} \;= \|\vec{b} - A\vec{x}\| = \sqrt{\frac{25}{4} + 1 + \frac{1}{4}} = \frac{\sqrt{30}}{2} \simeq 2,74.$$

11.2 Regressão Linear Simples

Vamos apresentar uma aplicação de algumas das técnicas do método de mínimos quadrados à Estatística ou Econometria. Uma **regressão linear simples** é uma equação, tipicamente da forma

$$y = a + bx,$$

para estimar os valores (x, y) apenas conhecendo alguns valores específicos $(x_1, y_1), (x_2, y_2), \ldots,$ (x_k, y_k). A ideia é tentar capturar como que mudanças na variável independente x afetam a variável dependente y (neste caso, supondo que esta dependência é linear).

Exemplo 104. [1] Com o intuito de analisar se é razoável supor que há um relação linear entre a idade de um motorista e quão longe ele consegue ver, uma empresa (Last Resource, Inc., Bellefonte, PA) coletou dados de 30 motoristas. Para simplificar as nossas contas, vamos listar abaixo *apenas alguns* destes dados.

Idade	Distância (em m)
20	590
32	410
41	460
49	380
66	350

Podemos pensar em y como a distância e em x como a idade. Gostaríamos de achar uma relação linear da forma

$$y = a + bx.$$

Desta forma, os dados obtidos implicam que

$$\begin{cases} b + 20a = 590 \\ b + 32a = 410 \\ b + 41a = 460 \\ b + 49a = 380 \\ b + 66a = 350 \end{cases} \quad \rightsquigarrow \quad \begin{bmatrix} 1 & 20 \\ 1 & 32 \\ 1 & 41 \\ 1 & 49 \\ 1 & 66 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 590 \\ 410 \\ 460 \\ 380 \\ 350 \end{bmatrix}$$

Ora, dificilmente um sistema linear com duas incógnitas e cinco equações terá solução (só terá solução se todos os pontos do conjunto de dados estiverem perfeitamente alinhados em uma reta!). Consequentemente, vamos procurar por uma solução de mínimos quadrados. Isto é **regressão linear simpes**.

[1] Exemplo adaptado de https://onlinecourses.science.psu.edu/stat501/node/257

Denotando

$$A = \begin{bmatrix} 1 & 20 \\ 1 & 32 \\ 1 & 41 \\ 1 & 49 \\ 1 & 66 \end{bmatrix} \quad \text{e} \quad \vec{b} = \begin{bmatrix} 590 \\ 410 \\ 460 \\ 380 \\ 350 \end{bmatrix}$$

precisamos calcular

$$A^T A = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 20 & 32 & 41 & 49 & 66 \end{bmatrix} \begin{bmatrix} 1 & 20 \\ 1 & 32 \\ 1 & 41 \\ 1 & 49 \\ 1 & 66 \end{bmatrix} = \begin{bmatrix} 5 & 208 \\ 208 & 9832 \end{bmatrix}$$

e

$$A^T \vec{b} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 20 & 32 & 41 & 49 & 66 \end{bmatrix} \begin{bmatrix} 590 \\ 410 \\ 460 \\ 380 \\ 350 \end{bmatrix} = \begin{bmatrix} 2190 \\ 85500 \end{bmatrix}.$$

Em seguida, a matriz associada aumentada pode ser reduzida por escalonamento:

$$[\,A^T A \mid \vec{b}\,] = \begin{bmatrix} 5 & 208 & 2190 \\ 208 & 9832 & 85500 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 468510/737 \\ 0 & 1 & -7005/1474 \end{bmatrix} \implies \begin{cases} a \simeq 635.7 \\ b \simeq -4.75 \end{cases}.$$

Os números são feios, mas as contas feitas foram as que sempre fizemos ao escalonar uma matriz até sua forma escalonada reduzida.

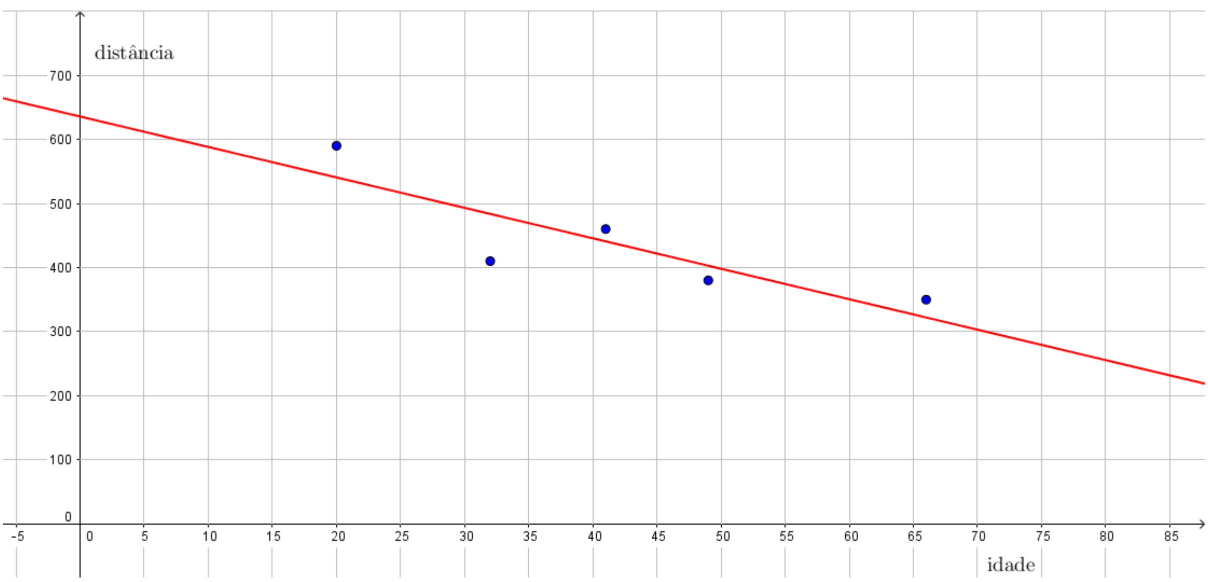
A conclusão é que a reta de mínimos quadrados que melhor aproxima os nossos dados é a reta

$$y = a + bx = 635.7 - 4.75x.$$

O erro de mínimos quadrados nesta aproximação (ou norma do resíduo) pode ser calculado como

$$A\vec{x} = \begin{bmatrix} 1 & 20 \\ 1 & 32 \\ 1 & 41 \\ 1 & 49 \\ 1 & 66 \end{bmatrix} \begin{bmatrix} 635.7 \\ 4.75 \end{bmatrix} \simeq \begin{bmatrix} 730.7 \\ 787.7 \\ 830.45 \\ 868.45 \\ 949.2 \end{bmatrix} \simeq \text{proj}_{\text{Col } A}\, \vec{b} \implies \text{erro } = \|\vec{b} - A\vec{x}\| \simeq 947.3$$

Na figura, mostramos um esboço da reta que melhor aproxima os dados deste exemplo. ◁



De uma maneira geral, para encontrar uma **reta de melhor ajuste** a uma quantidade de pontos (dados coletados de algum problema)

$$(x_1, y_1), \ (x_2, y_2), \ \ldots, (x_k, y_k),$$

devemos procurar pela solução (a, b) de mínimos quadrados do sistema linear

$$\begin{cases} a + bx_1 = y_1 \\ a + bx_2 = y_2 \\ \qquad \vdots \\ a + bx_k = y_k \end{cases} \quad \rightsquigarrow \quad \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_k \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_k \end{bmatrix}.$$

Vejamos um outro exemplo:

Exemplo 105. Encontrar a reta que melhor se ajusta aos pontos do plano

$$(1, 2), \quad (2, 3), \quad (2, 4), \quad (3, 2), \quad (4, 3), \quad (4, 4), \quad (5, 3), \quad (5, 5), \quad (6, 4).$$

Como vimos acima, podemos considerar

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \\ 1 & 4 \\ 1 & 5 \\ 1 & 5 \\ 1 & 6 \end{bmatrix} \quad \text{e} \quad \vec{b} = \begin{bmatrix} 2 \\ 3 \\ 4 \\ 2 \\ 3 \\ 4 \\ 3 \\ 5 \\ 4 \end{bmatrix}$$

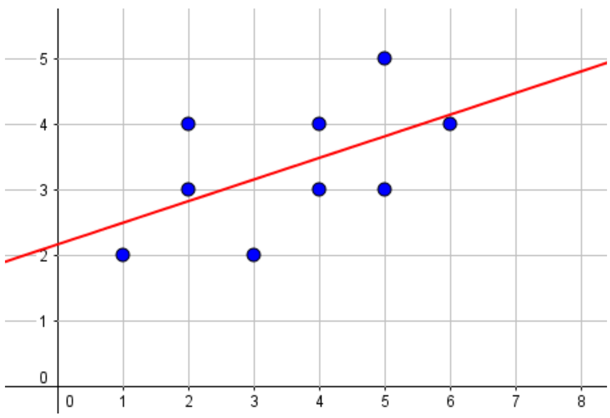
Assim,

$$A^T A = \begin{bmatrix} 9 & 32 \\ 32 & 136 \end{bmatrix} \quad \text{e} \quad A^T \vec{b} = \begin{bmatrix} 30 \\ 114 \end{bmatrix}.$$

A solução de mínimos quadrados, que é a solução de $A^T A\vec{x} = A^T \vec{b}$, é dada por

$$b = 54/25 = 2.16, a = 33/100 = 0.33.$$

A reta procurada é portanto $y = 0.33x + 2.16$.



11.3 Regressão não linear

Podemos também procurar (em contraste com a subseção anterior) por funções não lineares que se ajustem a um conjunto de pontos. Por exemplo, dado um certo conjunto de pontos

$$(x_1, y_1), \ (x_2, y_2), \ \ldots, (x_k, y_k),$$

vamos procurar por uma parábola

$$y = a + bx + cx^2$$

que melhor se ajuste a este conjunto de pontos, no sentido de que a soma dos quadrados dos erros seja a menor possível. Isto corrensponde a procurar pela solução (a, b, c) de mínimos quadrados do seguinte sistema linear

$$\begin{cases} a + bx_1 + cx_1^2 = y_1 \\ a + bx_2 + cx_2^2 = y_2 \\ \quad\vdots \\ a + bx_k + cx_k^2 = y_k \end{cases} \quad \leftrightsquigarrow \quad \begin{bmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ \vdots & \vdots & \\ 1 & x_k & x_k^2 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_k \end{bmatrix}.$$

Vejamos como ficaria a parábola que se ajusta aos dados do Exemplo 104:

Exemplo 106. Nossos pontos no Exemplo 104 são

$$(20, 590),\ (32, 410),\ (41, 460),\ (49, 380),\ (66, 350).$$

Para encontrar a parábola de melhor ajuste como acima, devemos procurar pela solução de mínimos quadrados do sistema

$$\begin{bmatrix} 1 & 20 & 400 \\ 1 & 32 & 1024 \\ 1 & 41 & 1681 \\ 1 & 49 & 2401 \\ 1 & 66 & 4356 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 590 \\ 410 \\ 460 \\ 380 \\ 350 \end{bmatrix} \quad \leftrightsquigarrow \quad A\vec{x} = \vec{b}.$$

Calculamos (com o auxílio de uma calculadora)

$$A^TA = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 20 & 32 & 41 & 49 & 66 \\ 400 & 1024 & 1681 & 2401 & 4356 \end{bmatrix} \begin{bmatrix} 1 & 20 & 400 \\ 1 & 32 & 1024 \\ 1 & 41 & 1681 \\ 1 & 49 & 2401 \\ 1 & 66 & 4356 \end{bmatrix} = \begin{bmatrix} 5 & 208 & 9862 \\ 208 & 9862 & 514834 \\ 9862 & 514834 & 28773874 \end{bmatrix},$$

$$A^TA = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 20 & 32 & 41 & 49 & 66 \\ 400 & 1024 & 1681 & 2401 & 4356 \end{bmatrix} \begin{bmatrix} 590 \\ 410 \\ 460 \\ 380 \\ 350 \end{bmatrix} = \begin{bmatrix} 2190 \\ 85500 \\ 3866080 \end{bmatrix}$$

e resolvemos (escalonando, por exemplo, com o auxílio de um computador)

$$\left[\begin{array}{ccc|c} 5 & 208 & 9862 & 2190 \\ 208 & 9862 & 514834 & 85500 \\ 9862 & 514834 & 28773874 & 3866080 \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & 0 & 0 & 28482529840/35036713 \\ 0 & 1 & 0 & -1505841055/105110139 \\ 0 & 0 & 1 & 11779375/105110139 \end{array}\right].$$

Logo, aproximando estas frações, obtemos

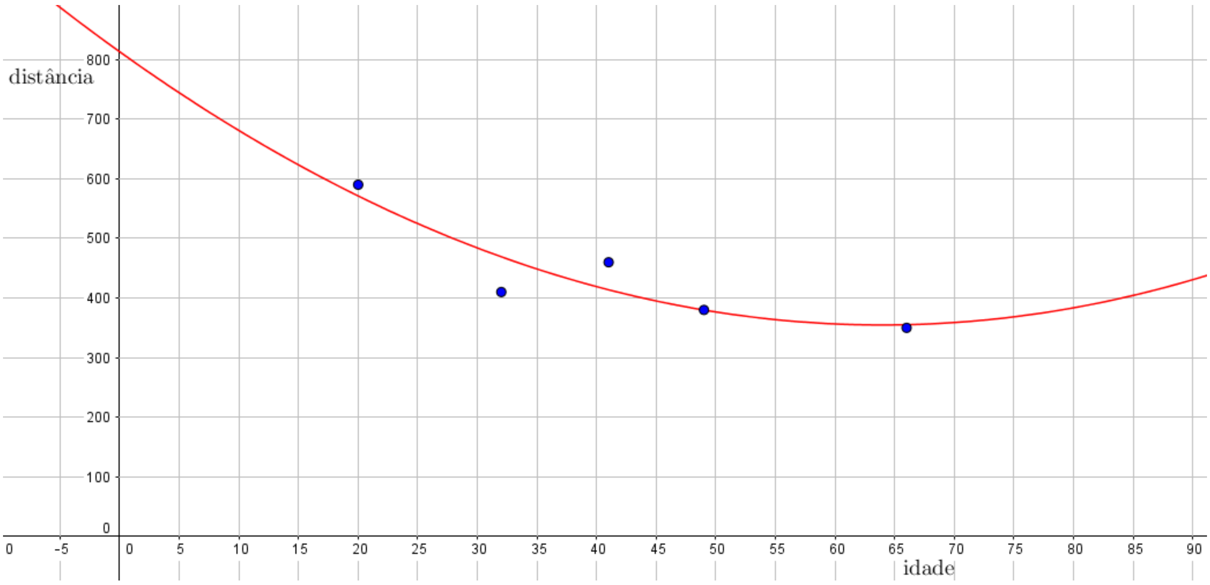
$$\begin{cases} a \simeq 812.934 \\ b \simeq -14.326 \\ c \simeq 0.112 \end{cases}$$

e a parábola desejada é, aproximadamente

$$y = 812.934 - 14.326x + 0.112x^2.$$

Observamos que, apesar de o erro quadrático ser provavelmente menor neste exemplo, esta parábola se torna crescente a partir de um certo momento, fazendo com que o modelo não seja tão razoável para idades maiores. Por exemplo, é pouco provável que (na média), pessoas com 95 anos vejam a uma distância maior do que pessoas de 65 anos, como sugere a parábola acima.

Licença CC-BY-SA-3.0. Contato: livroscolaborativos@gmail.com



11.4 Regressão linear múltipla

Uma **regressão linear múltipla** é o problema análogo à regressão linear simples no caso em que a variável dependente pode depender de mais fatores independentes. Tipicamente, queremos encontrar uma equação afim que melhor se ajusta a alguns dados conhecidos.

No caso especial em que y depende de apenas outros dois fatores, escreveremos

$$z = a + bx + cy,$$

que, geometricamente, representa um plano no espaço tridimensional $\mathbb{R}^3$. Seja agora uma conjunto de dados:

$$(x_1, y_1, z_1),\ (x_1, y_1, z_1), \ldots, (x_k, y_k, z_k).$$

Queremos encontrar coeficientes (a, b, c) que satisfaçam:

$$\begin{cases} a + bx_1 + cy_1 = z_1 \\ a + bx_2 + cy_2 = z_2 \\ \quad\vdots \\ a + bx_k + cy_k = z_k \end{cases} \quad \rightsquigarrow \quad \begin{bmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ \vdots & \vdots & \vdots \\ 1 & x_k & y_k \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_k \end{bmatrix}.$$

Quanto maior o número de dados, mais provável de o sistema ser impossível (podemos fazer a analogia geométrica de que por três pontos não colineares no espaço passa um único plano; se aumentarmos o número de pontos, mais difícil que haja um plano contendo todos eles). Por isto, procuramos por uma solução de mínimos quadrados.

Exemplo 107. Uma pesquisa com 214 mulheres em uma universidade americana[2] coletou informações sobre a altura das participantes, assim como a altura de seus pais. Abaixo, listamos *apenas alguns destes dados*, para que nossas contas não fiquem tão extensas. Fizemos também uma mudança de unidades nas alturas (de polegadas) para centímetros,

Altura	Altura Mãe	Altura Pai
152	155	165
162	155	160
165	170	173
170	163	183
173	168	183
183	165	183

[2]Referência: https://onlinecourses.science.psu.edu/stat501/node/292

Licença CC-BY-SA-3.0. Contato: livroscolaborativos@gmail.com

Queremos encontrar uma solução de mínimos quadrados para o sistema linear

$$\begin{bmatrix} 1 & 155 & 165 \\ 1 & 155 & 160 \\ 1 & 170 & 173 \\ 1 & 163 & 183 \\ 1 & 168 & 183 \\ 1 & 165 & 183 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 152 \\ 162 \\ 165 \\ 170 \\ 173 \\ 183 \end{bmatrix} \quad \leftrightsquigarrow \quad A\vec{x} = \vec{b}.$$

Calculamos

$$A^TA = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 155 & 155 & 170 & 163 & 168 & 165 \\ 165 & 160 & 173 & 183 & 183 & 183 \end{bmatrix} \begin{bmatrix} 1 & 155 & 165 \\ 1 & 155 & 160 \\ 1 & 170 & 173 \\ 1 & 163 & 183 \\ 1 & 168 & 183 \\ 1 & 165 & 183 \end{bmatrix} = \begin{bmatrix} 6 & 976 & 1047 \\ 976 & 158968 & 170553 \\ 1047 & 170553 & 183221 \end{bmatrix}$$

e

$$A^T\vec{b} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 155 & 155 & 170 & 163 & 168 & 165 \\ 165 & 160 & 173 & 183 & 183 & 183 \end{bmatrix} \begin{bmatrix} 152 \\ 162 \\ 165 \\ 170 \\ 173 \\ 183 \end{bmatrix} = \begin{bmatrix} 1005 \\ 163689 \\ 175803 \end{bmatrix}.$$

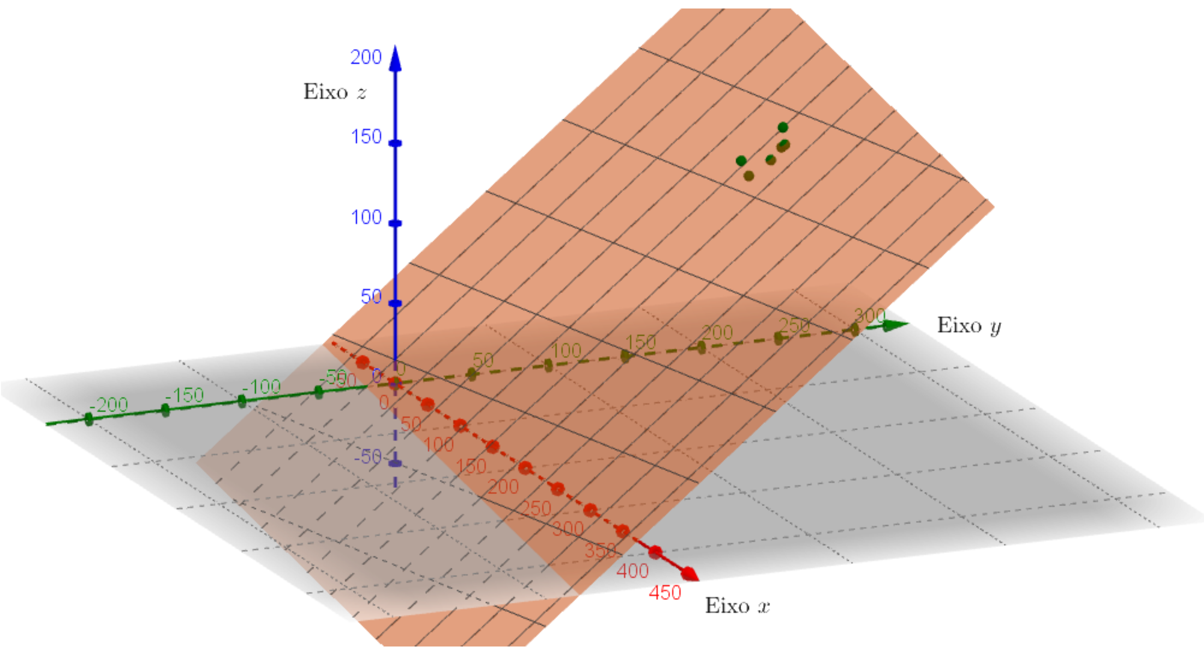
Por escalonamento,

$$\begin{bmatrix} 6 & 976 & 1047 & 1005 \\ 976 & 158968 & 170553 & 163689 \\ 1047 & 170553 & 183221 & 175803 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & 2154573/145769 \\ 0 & 1 & 0 & 14475/145769 \\ 0 & 0 & 1 & 114081/145769 \end{bmatrix} \implies \begin{cases} a \simeq 14.781 \\ b \simeq 0.099 \\ c \simeq 0.783 \end{cases}$$

A equação de melhor ajuste procurada é, portanto, aproximadamente,

$$z \simeq 14.781 + 0.099x + 0.783y.$$

Tente calcular sua altura z a partir da altura de sua mãe x e de seu pai y. O teste deve funcionar melhor para mulheres! Além disso, a aproximação linear deve ser melhor utilizando mais dados nos cálculos.



11.5 Fatoração QR e Mínimos Quadrados

Vimos que o processo de Gram-Schmidt nos permite escrever uma matriz A cujas linhas são linearmente independentes na forma $A = QR$, onde Q é uma matriz ortogonal, R é triangular superior invertível, e as colunas de Q formam uma base ortonormal para o espaço-coluna de A. Chamamos esta fatoração de A de fatoração QR.

A fatoração QR tem uma importante aplicação na solução de problemas em mínimos quadrados. Ocorre que, em muitos problemas, a matriz $A^T A$ é uma matriz *mal-condicionada*, ou seja, uma matriz que é muito sensível aos erros de aproximação cometidos em cálculos numéricos usualmente executados no tratamento de problemas aplicados. Os exemplos que virão a seguir nos mostram situações em que há uma grande variação de magnitude entre as entradas da matriz $A^T A$, e isto costuma ser fonte de instabilidade numérica.

Para isso, a fatoração QR oferece uma solução alternativa em mínimos quadrados para um sistema linear $A\vec{x} = \vec{b}$: se $A = QR$, então

$$A^T = R^T Q^T \quad \text{que implica em} \quad A^T A = R^T Q^T Q R = R^T R,$$

onde usamos o fato de que $Q^{-1} = Q^T$. Assim as equações normais $A^T A\vec{x} = A^T\vec{b}$ se reduzem a

$$R^T R\vec{x} = R^T Q^T \vec{b}.$$

Ainda, como R é invertível podemos eliminar (cancelar) R^T na equação acima, obtendo o seguinte:

Teorema 108. *Se $A = QR$ com Q ortogonal e R triangular superior invertível, então a solução em mínimos quadrados do sistema $A\vec{x} = \vec{b}$ é única e dada pela solução do sistema*

$$R\vec{x} = Q^T\vec{b}.$$

Note que o fato de R ser triangular superior torna a resolução do sistema acima pouco custosa do ponto de vista computacional. Alternativamente, temos uma expressão explícita para solução em mínimos quadrados:

$$\vec{x} = R^{-1} Q^T \vec{b}.$$

É importante notar que existem algoritmos para fatoração QR que são *numéricamente estáveis*, ou seja pouco suscetiveis a erros de aproximação ou arredondamento. Tais algorimos são diferentes do processo de Gram Schmidt usual.

Licença CC-BY-SA-3.0. Contato: livroscolaborativos@gmail.com

12

Semanas 14 e 15

12.1 Matrizes simétricas

Uma **matriz simétrica** é uma matriz que é igual à sua transposta. Para que esta definição faça sentido, apenas podemos considerar matrizes que são quadradas. De forma mais precisa, se $A = \begin{bmatrix} a_{ij} \end{bmatrix}$ é uma matriz de orden $n \times n$, nós dizemos que A é simétrica quando $A = A^T$.

Formas equivalentes de dizer que A é simétrica incluem:

(1) $A = A^T$;

(2) para todos os índices $i, j \in \{1, 2, \ldots, n\}$, temos que $a_{ij} = a_{ji}$;

(3) as colunas de A são iguais às linhas de A^T;

(4) as linhas de A são iguais às colunas de A^T.

Exemplo 109. As matrizes

$$\begin{bmatrix} 1 & -3 \\ -3 & 2 \end{bmatrix}, \quad \begin{bmatrix} 1 & -3 \\ -3 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & -3 & 0 \\ -3 & -1 & 4 \\ 0 & 4 & 6 \end{bmatrix} \checkmark$$

são simétricas, enquanto as matrizes

$$\begin{bmatrix} 1 & -3 \\ 3 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 3 \\ -3 & 2 \end{bmatrix}, \quad \begin{bmatrix} 1 & -3 & 0 \\ 4 & -1 & -3 \\ 0 & 4 & 1 \end{bmatrix} \text{x}$$

não são. Deve-se tomar o cuidado de verificar a simetria *com relação à diagonal principal.* Qualquer outro tipo de simetria não deve interferir na nossa análise. $\triangleleft$

Um dos resultados mais importantes de Álgebra Linear é o Teorema Espectral, que diz que toda matriz simétrica é diagonalizável por uma base de autovetores que é ortonormal. De forma mais precisa:

Teorema Espectral. *Suponhamos que A é uma matriz simétrica com coeficientes reais. Então:*

(i) *todos os autovalores de A são reais;*

(ii) *autovetores associados a autovalores distintos são ortogonais.*

(iii) *a dimensão do autoespaço associado a um autovalor é igual à multiplicidade deste autovalor como raiz do polinômio característico;*

(iv) *existe uma base ortonormal de $\mathbb{R}^n$ formada por autovetores de A;*

(v) *A é diagonalizável por uma matriz de mudança de coordenadas ortonormal P: a matriz P cujas colunas são os autovetores (ortonormais) de A é tal que*

$$P^T A P = P^{-1} A P = D.$$

Nós vamos apresentar uma prova deste resultado no final destas notas. Antes de mais nada, faremos alguns exemplos e aplicações que justificam a importância do Teorema Espectral. No entanto, é fortemente aconselhado que o leitor acompanhe os elementos da demonstração, para efetivamente entender as propriedades essenciais das matrizes simétricas.

Exemplo 110. Considere

$$A = \begin{bmatrix} 1 & -5 \\ -5 & 1 \end{bmatrix}.$$

O polinômio característico é

$$\det \begin{bmatrix} 1-\lambda & -4 \\ -4 & 5-\lambda \end{bmatrix} = (1-\lambda)^2 - 25 = \lambda^2 - 2\lambda - 24$$

cujas raízes são 6 e -4. Os autovetores são calculados como

$$A + 4I = \begin{bmatrix} 5 & -5 \\ -5 & 5 \end{bmatrix} \sim \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix} \implies \begin{cases} v_1 = v_2 \\ \text{uma variável livre} \end{cases} \implies \vec{v}_1 = v_2 \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Associado com 6:

$$A - 6I = \begin{bmatrix} -5 & -5 \\ -5 & -5 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \implies \begin{cases} v_2 = -v_1 \\ \text{uma variável livre} \end{cases} \implies \vec{v}_2 = v_2 \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

Uma base de autovetores é

$$\left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right\}.$$

Observe que são de fato reais os autovalores e que os autovetores associados são ortogonais (faça a conta!). Se quisermos uma base ortonormal para $\mathbb{R}^2$ formada por autovetores, normalizamos os que já obtivemos:

$$\mathbb{R}^2 = \operatorname{Span} \left\{ \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}, \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix} \right\}.$$

Lembrando do nosso processo de diagonalização, sabemos que

$$P = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix} \implies P^{-1}AP = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 1 & -5 \\ -5 & 1 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix} = \begin{bmatrix} -4 & 0 \\ 0 & 6 \end{bmatrix}.$$

A matriz P que montamos acima é uma matriz ortogonal (de modo que $P^{-1} = P^T$), pois sabemos que autovetores associados com autovalores distintos são ortogonais e fizemos a normalização dos autovetores (lembra que uma matriz ortogonal deve ter colunas ortonormais). $\lhd$

Exemplo 111. Vamos aplicar o procedimento de diagonalização para encontrar uma base ortonormal de autovetores de uma matriz de ordem 3×3. Considere a matriz simétrica

$$A = \begin{bmatrix} 3 & -2 & 4 \\ -2 & 6 & 2 \\ 4 & 2 & 3 \end{bmatrix}.$$

O polinômio característico de A é

$$\begin{aligned} p(\lambda) = \det(A - \lambda I) &= \det \begin{bmatrix} 3-\lambda & -2 & 4 \\ -2 & 6-\lambda & 2 \\ 4 & 2 & 3-\lambda \end{bmatrix} \\ &= (3-\lambda)\left[(6-\lambda)(3-\lambda) - 4\right] + 2\left[-6 + 2\lambda - 8\right] + 4\left[-4 - 24 + 4\lambda\right] \\ &= (3-\lambda)\left[\lambda^2 - 9\lambda + 14\right] - 28 + 4\lambda - 112 + 16\lambda \\ &= 3\lambda^2 - 27\lambda + 42 - \left[\lambda^3 - 9\lambda^2 + 14\lambda\right] + 20\lambda - 140 \\ &= -\lambda^3 + 12\lambda^2 - 21\lambda - 98. \end{aligned}$$

Como sabemos, se os coeficientes da matriz simétrica acima fossem escolhidos de maneira aleatória, dificilmente conseguiríamos encontrar as raízes deste polinômio com precisão. De qualquer forma, quaisquer raízes racionais (este exemplo foi construido artificialmente) deste polinômio só podem ser divisores de 98 dividido pelos divisores de -1, isto é: $\pm 1, \pm 2, \pm 7, \pm 14, \pm 21$. Testando os primeiros, vemos que -2 é uma raiz. Pela divisão de polinômios, fatoramos

$$p(\lambda) = -(\lambda + 2)(\lambda^2 - 14\lambda + 49) = -(\lambda + 2)(\lambda - 7)^2.$$

Logo, os autovalores de A são -2 (simples) e 7 (duplo).

- Autoespaço associado com o autovalor -2:

$$A+2I = \begin{bmatrix} 5 & -2 & 4 \\ -2 & 8 & 2 \\ 4 & 2 & 5 \end{bmatrix} \sim \begin{bmatrix} -1 & 4 & 1 \\ 5 & -2 & 4 \\ 4 & 2 & 5 \end{bmatrix} \sim \begin{bmatrix} -1 & 4 & 1 \\ 0 & 18 & 9 \\ 0 & 18 & 9 \end{bmatrix} \sim \begin{bmatrix} -1 & 2 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 0 \end{bmatrix} \leftrightsquigarrow \begin{cases} v_1 = 2v_2 \\ v_3 = -2v_2 \\ \text{1 variável livre} \end{cases}$$

  Assim, um autovetor associado é ($v_2 = 1$)

$$\vec{v}_1 = \begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix}.$$

- Autoespaço associado com o autovalor 7:

$$A - 7I = \begin{bmatrix} -4 & -2 & 4 \\ -2 & -1 & 2 \\ 4 & 2 & -4 \end{bmatrix} \sim \begin{bmatrix} -2 & -1 & 2 \\ 2 & 1 & -2 \\ 0 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} -2 & -1 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \leftrightsquigarrow \begin{cases} v_2 = -2v_1 + 2v_3 \\ \text{2 variáveis livres} \end{cases}$$

  Assim, qualquer autovetor pode ser escrito parametricamente como

$$\vec{v} = \begin{bmatrix} v_1 \\ -2v_1 + 2v_3 \\ v_3 \end{bmatrix} = v_1 \begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix} + v_2 \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix}.$$

  Já era de se esperar que encontraríamos duas variáveis livres, porque sabemos (item (iii) do Teorema Espectral) que deve ser dois a dimensão do autoespaço associado com um autovalor de multiplicidade dois.[1] Concluimos que uma base para o autoespaço $\text{Nul}(A - 7I)$ é

$$\left\{ \vec{v}_2 = \begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix}, \vec{v}_3 = \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix} \right\}.$$

Observamos que, em concordância com o item (ii) do Teorema Espectral, autovetores associados a autovalores distintos são ortogonais: temos que $\vec{v}_1$ é ortogonal a cada um dos vetores $\vec{v}_2$ e $\vec{v}_3$ (faça as contas!). No entanto, estes três vetores não são ortogonais, pois os dois autovetores que estão associados com o mesmo autovalor satisfazem $\vec{v}_2 \cdot \vec{v}_3 = -4 \neq 0$.

O fato de que os três autovetores que encontramos não são ortogonais não contradiz o Teorema Espectral. Pelo que vimos, o autoespaço associado com o autovalor 7 tem dimensão dois. Encontramos uma base da maneira usual, mas ocorre que esta base não é ortogonal. Para obter uma base ortogonal, aplicamos o procedimento de Gram-Schmidt:

$$\vec{u}_2 = \vec{v}_2 = \begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix}$$

$$\vec{u}_3 = \vec{v}_3 - \text{proj}_{\vec{u}_2} \vec{v}_3 = \vec{v}_3 - \frac{\vec{v}_3 \cdot \vec{u}_2}{\vec{u}_2 \cdot \vec{u}_2} \vec{u}_2 = \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix} - \frac{-4}{5} \begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix} = \begin{bmatrix} 4/5 \\ 2/5 \\ 1 \end{bmatrix} \xrightarrow{\text{podemos considerar } \times 5} \begin{bmatrix} 4 \\ 2 \\ 5 \end{bmatrix}$$

Agora

$$\left\{ \begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix}, \begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix}, \begin{bmatrix} 4 \\ 2 \\ 5 \end{bmatrix} \right\}.$$

[1] Enfatizamos que isto não é verdade para qualquer matriz, apenas para as simétricas!

Licença CC-BY-SA-3.0. Contato: livroscolaborativos@gmail.com

é uma base ortogonal de $\mathbb{R}^3$ formada por autovalores da matriz A. Para obter uma base ortonormal, basta normalizar cada um dos vetores da base:

$$\left\{ \begin{bmatrix} 2/3 \\ 1/3 \\ -2/3 \end{bmatrix}, \begin{bmatrix} 1/\sqrt{5} \\ -2/\sqrt{5} \\ 0 \end{bmatrix}, \begin{bmatrix} 4/(3\sqrt{5}) \\ 2/(3\sqrt{5}) \\ 5/(3\sqrt{5}) \end{bmatrix} \right\}.$$

Pelo método de diagonalização de matrizes, concluimos que A pode ser diagonalizada pela matriz

$$P = \begin{bmatrix} 2/3 & 1/\sqrt{5} & 4/(3\sqrt{5}) \\ 1/3 & -2/\sqrt{5} & 2/(3\sqrt{5}) \\ -2/3 & 0 & 5/(3\sqrt{5}) \end{bmatrix},$$

que, neste caso (em que A é matriz simétrica), pode ser obtida ortogonal. Sendo P ortogonal, tem-se $P^{-1} = P^T$ e portanto $P^T A P = D$, isto é,

$$\begin{bmatrix} 2/3 & 1/\sqrt{5} & 4/(3\sqrt{5}) \\ 1/3 & -2/\sqrt{5} & 2/(3\sqrt{5}) \\ -2/3 & 0 & 5/(3\sqrt{5}) \end{bmatrix} \begin{bmatrix} 5 & -2 & 4 \\ -2 & 8 & 2 \\ 4 & 2 & 5 \end{bmatrix} \begin{bmatrix} 2/3 & 1/\sqrt{5} & 4/(3\sqrt{5}) \\ 1/3 & -2/\sqrt{5} & 2/(3\sqrt{5}) \\ -2/3 & 0 & 5/(3\sqrt{5}) \end{bmatrix} = \begin{bmatrix} -2 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & 7 \end{bmatrix}. \triangleleft$$

Enfatizamos ainda mais uma vez a aplicação do Teorema Espectral: ele afirma que, no caso de uma matriz simétrica A, o procedimento que aplicamos nos exemplos acima sempre vai funcionar! Teremos todos os autovalores reais. Sendo um autovalor λ de multiplicidade m, seremos capazes de encontrar m autovetores linearmente independentes, que geram o autoespaço $\mathrm{Nul}(A - \lambda I)$. Estes podem ser ortonormalizados pelo processo de Gram-Schmidt. Além disso, associados a autovalores distintos, temos autovetores ortogonais, de modo que seremos capazes de obter uma base ortonormal de autovetores de A.

12.1.1 Formas quadráticas

Nesta seção, apresentamos uma aplicação matemática de matrizes simétricas. Uma **forma quadrática** em $\mathbb{R}^n$ é uma função polinomial em que todos os termos são de grau 2.

Exemplo 112. A função $Q : \mathbb{R}^2 \to \mathbb{R}$, definida por

$$Q(x_1, x_2) = 3x_1^2 + x_1 x_2$$

é uma forma quadrática em $\mathbb{R}^2$, já que cada um dos termos tem grau 2.

Mais geralmente, qualquer forma quadrática em $\mathbb{R}^2$ pode ser escrita como

$$Q(x_1, x_2) = a x_1^2 + 2b x_1 x_2 + c x_2^2,$$

onde $a, b, c \in \mathbb{R}$, já que os três termos acima são todos os possíveis monômios de grau dois em duas variáveis x_1 e x_2. O coeficiente 2 que aparece no segundo termo é motivado pela fórmula mais geral (12.1) abaixo (observe que cada termo $x_i x_j$ aparece duas vezes na soma).

Exemplo 113. A função $Q : \mathbb{R}^3 \to \mathbb{R}$, definida por

$$Q(x_1, x_2, x_3) = x_2^2 - x_3^2 + 3x_1 x_2 - 8x_1 x_3$$

é uma forma quadrática em $\mathbb{R}^3$. Qualquer forma quadrática em $\mathbb{R}^3$ pode ser escrita como

$$Q(x_1, x_2) = a x_1^2 + b x_2^2 + c x_3^2 + 2d x_1 x_2 + 2e x_1 x_3 + 2f x_2 x_3,$$

onde $a, b, c, d, e, f \in \mathbb{R}$.

De forma geral, uma forma quadrática $Q : \mathbb{R}^n \to \mathbb{R}$ pode ser escrita como

$$Q(x_1, x_2, \ldots, x_n) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j, \tag{12.1}$$

Licença CC-BY-SA-3.0. Contato: livroscolaborativos@gmail.com

onde cada um dos coeficientes a_{ij} são números reais. Podemos também supor que a matriz $A = (a_{ij})$ é simétrica.[2]

Notamos o seguinte: se A é uma matriz de ordem $m \times n$ e $\vec{x} \in \mathbb{R}^n$ é um vetor com n componentes, então o vetor $A\vec{x}$, que é o produto de A por $\vec{x}$, tem na sua i-ésima componente o produto escalar entre a linha i de A e o vetor $\vec{x}$:

$$(A\vec{x})_i = \text{componente } i \text{ de } A\vec{x} = \sum_{j=1}^{n} a_{ij}x_j.$$

Dessa forma, temos

$$Q(x_1, x_2, \ldots, x_n) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_ix_j = \sum_{i=1}^{n} (A\vec{x})_i\, x_i = (A\vec{x}) \cdot \vec{x}.$$

E aí está a relação entre as formas quadráticas e matrizes simétricas. Acabamos de verificar que:

$$\boxed{\begin{array}{c}\text{Qualquer forma quadrática em } \mathbb{R}^n \text{ pode ser escrita na forma} \\ Q(x_1, x_2, \ldots, x_n) = A\vec{x} \cdot \vec{x}, \\ \text{onde } A \text{ é uma matriz simétrica de ordem } n \times n.\end{array}}$$

Uma outra forma de escrever a mesma coisa (verifique a última igualdade como exercício) é

$$Q(x_1, x_2, \ldots, x_n) = A\vec{x} \cdot \vec{x} = \vec{x}^T A\vec{x}.$$

Exemplo 114. Voltamos a analisar os Exemplos 112 e 113 acima. Uma forma quadrática em $\mathbb{R}^2$ pode ser escrita na forma

$$Q(x_1, x_2) = ax_1^2 + 2bx_1x_2 + cx_2^2 = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} a & b \\ b & c \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Assim,

$$Q(x_1, x_2) = 3x_1^2 + x_1x_2 = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 3 & 1/2 \\ 1/2 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

Observe como o coeficiente do termo misto x_1x_2 aparece (dividido por 2) nas posições a_{12} e a_{21} da matriz A. Os coeficientes de x_1^2 e de x_2^2 aparecem na diagonal principal. Antes de prosseguir, faremos uma conta para verificar que de fato tudo está de acordo:

$$\begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 3 & 1/2 \\ 1/2 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 3x_1 + x_2/2 \\ x_1/2 \end{bmatrix} = 3x_1^2 + \frac{x_1x_2}{2} + \frac{x_1x_2}{2}.$$

No caso 3×3, aplicamos a mesma ideia

$$Q(x_1, x_2, x_3) = x_2^2 - x_3^2 + 3x_1x_2 - 8x_1x_3 = \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \begin{bmatrix} 0 & 3/2 & -4 \\ 3/2 & 1 & 0 \\ -4 & 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

Assim como anteriormente, observe que

- os coeficientes de x_1^2, de x_2^2 e de x_3^2 aparecem na diagonal principal;

[2] Caso $Q(x_1, x_2, \ldots, x_n) = \sum_{i=1}^{n}\sum_{j=1}^{n} b_{ij}x_ix_j$ com uma matriz $B = (b_{ij})$ não simétrica, podemos escrever $Q(x_1, x_2, \ldots, x_n) = \sum_{i=1}^{n} b_{ii}x_i^2 + \sum_{i<j}(b_{ij} + b_{ji})x_ix_j$. Definindo $a_{ij} = \dfrac{b_{ij} + b_{ji}}{2}$ temos também $Q(x_1, x_2, \ldots, x_n) = \sum_{i=1}^{n} a_{ii}x_i^2 + 2\sum_{i<j} a_{ij}x_ix_j$. Esta forma é mais semelhante com as dos exemplos acima. Note como aparece o coeficiente 2 nos termos com $i \neq j$. Finalmente, observamos que $A = (a_{ij})$ é uma matriz simétrica e que $\sum_{i=1}^{n}\sum_{j=1}^{n} b_{ij}x_ix_j = \sum_{i=1}^{n} a_{ii}x_i^2 + 2\sum_{i<j} a_{ij}x_ix_j = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_ix_j$.

Licença CC-BY-SA-3.0. Contato: livroscolaborativos@gmail.com

- o coeficiente do termo misto x_1x_2 aparece (dividido por 2) nas posições a_{12} e a_{21} de A;
- o coeficiente do termo misto x_1x_3 aparece (dividido por 2) nas posições a_{13} e a_{31} de A;
- o coeficiente do termo misto x_2x_3 aparece (dividido por 2) nas posições a_{23} e a_{32} de A.

Faça as contas para verificar a igualdade acima! $\triangleleft$

12.1.2 Diagonalização de formas quadráticas

Embora não tenhamos dito explicitamente, estávamos trabalhando na base canônica de $\mathbb{R}^n$. Da forma usual, é possível fazer mudança de coordenadas. Do ponto de vista da matriz A, que é simétrica e portanto diagonalizável, a base (ortonormal) de autovetores associados é a mais natural de todas. Suponhamos que

$$Q(\vec{x}) = \vec{x}^T A\vec{x}$$

e que $\mathcal{B} = \{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n\}$ é base ortonormal de $\mathbb{R}^n$ formada por autovetores de A. Desta forma,

$$P = \begin{bmatrix} | & | & & | \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_n \\ | & | & & | \end{bmatrix} \text{ satisfaz } P^{-1}AP = D,$$

onde D é a matriz diagonal, com os respectivos autovalores na diagonal principal. Além disso, P é ortogonal: $P^T = P^{-1}$. O vetor $\vec{x}$ pode ser representado na base $\mathcal{B}$ por:

$$\vec{x} = y_1\vec{v}_1 + y_2\vec{v}_2 + \cdots + y_n\vec{v}_n = P\vec{y}, \quad \text{onde denotamos} \quad \vec{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = [\vec{x}]_{\mathcal{B}}.$$

Agora, vamos escrever a forma quadrática na base $\mathcal{B}$, isto é, em termos do vetor $\vec{y}$ (lembre que $\vec{x} = P\vec{y} \implies \vec{x}^T = \vec{y}^T P^T$):

$$\vec{x}^T A\vec{x} = (\vec{y}^T P^T)A(P\vec{y}) = \vec{y}^T (P^T AP)\vec{y} = \vec{y}^T D\vec{y}.$$

Isto significa que, ao mudar da base canônica (componentes x_i) para a base $\mathcal{B}$ (componentes y_i), temos

$$\sum_{i,j=1}^{n} a_{ij}x_ix_j = \vec{x}^T A\vec{x} = \vec{y}^T D\vec{y} = \sum_{i=1}^{n} \lambda_i y_i^2,$$

onde os números $\lambda_i \in \mathbb{R}$ são os autovalores da matriz A (são reais porque a matriz é simétrica). Em outras palavras, a base (ortonormal) de autovetores de A deixou transformou a nossa forma quadrática original em uma que não possui mais termos mistos, muito mais simples de trabalhar.

Exemplo 115. Considere a forma quadrática

$$Q(x_1, x_2) = 6x_1^2 - 2x_1x_2 + 4x_2^2 \quad \rightsquigarrow \quad Q(\vec{x}) = \vec{x}^T A\vec{x}, \quad \text{onde} \quad A = \begin{bmatrix} 6 & -2 \\ -2 & 3 \end{bmatrix}.$$

Vamos fazer uma mudança de coordenadas que transforma a forma quadrática Q em uma forma sem termos mistos. Os autovalores de A são as raízes de

$$p(\lambda) = \det(A - \lambda I) = (6 - \lambda)(3 - \lambda) - 4 = \lambda^2 - 9\lambda + 14.$$

Assim,

$$\lambda = \frac{9 \pm \sqrt{81 - 56}}{2} = \frac{9 \pm 5}{2} \implies \lambda_1 = 2, \lambda_2 = 7.$$

Já sabemos que, ao obter uma base ortonormal de autovetores e escrevermos $\vec{x}$ nesta base, a saber $\vec{x} = P\vec{y}$, teremos

$$Q(\vec{x}) = Q(P\vec{y}) = 2y_1^2 + 7y_2^2.$$

Vamos verificar neste exemplo que isto de fato acontece.

- Autovetor unitário associado ao autovalor $\lambda_1 = 2$:

$$A - 2I = \begin{bmatrix} 4 & -2 \\ -2 & 1 \end{bmatrix} \sim \begin{bmatrix} 2 & -1 \\ 0 & 0 \end{bmatrix} \implies v_2 = 2v_1 \rightsquigarrow \begin{bmatrix} 1 \\ 2 \end{bmatrix} \rightsquigarrow \vec{v}_1 = \begin{bmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix}.$$

- Autovetor unitário associado ao autovalor $\lambda_2 = 7$:

$$A - 2I = \begin{bmatrix} -1 & -2 \\ -2 & -4 \end{bmatrix} \sim \begin{bmatrix} -1 & -2 \\ 0 & 0 \end{bmatrix} \implies v_1 = -2v_2 \rightsquigarrow \begin{bmatrix} -2 \\ 1 \end{bmatrix} \rightsquigarrow \vec{v}_2 = \begin{bmatrix} -2/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix}.$$

Portanto, encontramos uma base ortonormal (de modo que P abaixo é matriz ortogonal e satisfaz $P = P^T$) e podemos escrever

$$P = \begin{bmatrix} 1/\sqrt{5} & -2/\sqrt{5} \\ 2/\sqrt{5} & 1/\sqrt{5} \end{bmatrix} \implies P^TAP = D = \begin{bmatrix} 2 & 0 \\ 0 & 7 \end{bmatrix}$$

Assim,

$$Q(x_1, x_2) = Q(\vec{x}) = Q(P\vec{y}) = \vec{x}^T A\vec{x} = \vec{y}^T A\vec{y} = 2y_1^2 + 7y_2^2. \lhd$$

12.1.3 Aplicação à classificação de cônicas

Nesta subseção, nos restringimos a formas quadráticas em $\mathbb{R}^2$. Uma **curva de nível** de uma forma quadrática, também conhecida como uma **cônica**, é um conjunto definido implicitamente pela equação

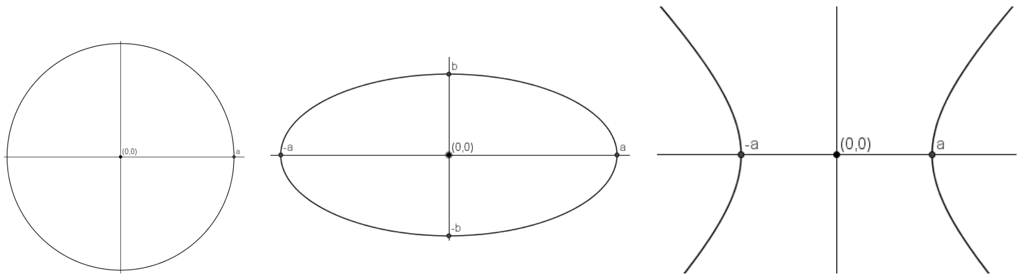
$$Q(x_1, x_2) = k \quad \longleftrightarrow \quad ax_1^2 + bx_1x_2 + cx_2^2 = k,$$

onde k é uma constante. Em outras palavras, uma curva de nível é o conjunto de todos os pontos (x_1, x_2) que satisfazem $Q(x_1, x_2) = k$.

Já conhecemos alguns exemplos:

- $x_1^2 + x_2^2 = r^2$ representa um círculo de raio r;
- $\dfrac{x_1^2}{a^2} + \dfrac{x_2^2}{b^2} = 1$ representa uma elipse;
- $\dfrac{x_1^2}{a^2} - \dfrac{x_2^2}{b^2} = 1$ representa uma hipérbole.

Graficamente, vemos alguns exemplos na figura abaixo.



Vamos ver como que a diagonalização de formas simétricas se aplica para conseguirmos identificar cônicas de formas quadráticas que possuem termos mistos.

Exemplo 116. Vamos identificar a cônica cuja equação é

$$31x_1^2 - 10\sqrt{3}x_1x_2 + 21x_2^2 = 144.$$

Podemos escrever

$$\vec{x}^T A\vec{x} = 144, \quad \text{onde} \quad A = \begin{bmatrix} 31 & -5\sqrt{3} \\ -5\sqrt{3} & 21 \end{bmatrix}.$$

Licença CC-BY-SA-3.0. Contato: livroscolaborativos@gmail.com

Autovalores de A:

$$p(\lambda) = (31-\lambda)(21-\lambda) - 75 = 0 \implies \lambda^2 - 52\lambda + 576 = 0$$

$$\implies \lambda = \frac{52 \pm \sqrt{2704 - 2304}}{2} = \frac{52 \pm 20}{2} = 26 \pm 10 \implies \lambda_1 = 16,\ \lambda_2 = 36.$$

Autovetores associados:

- ao autovalor $\lambda_1 = 16$:

$$\begin{bmatrix} 15 & -5\sqrt{3} \\ -5\sqrt{3} & 5 \end{bmatrix} \sim \begin{bmatrix} 3 & -\sqrt{3} \\ 0 & 0 \end{bmatrix} \implies 3v_1 = \sqrt{3}v_2 \xrightarrow[\text{e já normalizando}]{\text{podemos escolher}} \vec{v}_1 = \begin{bmatrix} 1/2 \\ \sqrt{3}/2 \end{bmatrix}.$$

- ao autovalor $\lambda_2 = 36$:

$$\begin{bmatrix} -5 & -5\sqrt{3} \\ -5\sqrt{3} & -15 \end{bmatrix} \sim \begin{bmatrix} 1 & \sqrt{3} \\ 0 & 0 \end{bmatrix} \implies v_1 = -\sqrt{3}v_2 \xrightarrow[\text{e já normalizando}]{\text{podemos escolher}} \vec{v}_2 = \begin{bmatrix} -\sqrt{3}/2 \\ 1/2 \end{bmatrix}.$$

Temos

$$P^{-1}AP = D = \begin{bmatrix} 16 & 0 \\ 0 & 26 \end{bmatrix}$$

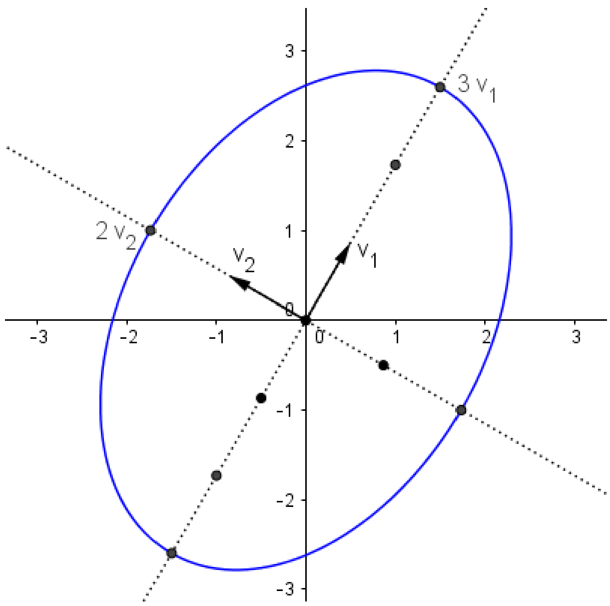
Assim, na base de $\mathbb{R}^2$ determinada por $\{\vec{v}_1, \vec{v}_2\}$, podemos escrever a equação da equação cônica como

$$\vec{x}^T A\vec{x} = \vec{y}^T D\vec{y} \implies \begin{bmatrix} y_1 & y_2 \end{bmatrix} \begin{bmatrix} 16 & 0 \\ 0 & 26 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = 144.$$

Isto pode ser reescrito como

$$16y_1^2 + 36y_2^2 = 144 \rightsquigarrow \frac{y_1^2}{(144/16)} + \frac{y_1^2}{(144/36)} = 1 \rightsquigarrow \frac{y_1^2}{3^2} + \frac{y_1^2}{2^2} = 1.$$

Esta equação é pode ser agora identificada como uma elipse nos "novos eixos" $\vec{v}_1$ e $\vec{v}_2$.



Observamos que quando a coordenada $y_1 = 0$ devemos ter que $y_2^2 = 2^2$, de modo que deve ser $\vec{x} = y_1\vec{v}_1 + y_2\vec{v}_2 = \pm 2\vec{v}_2$, como na figura. Este é o análogo da análise que normalmente fazemos, só que agora na base ortonormal $\{\vec{v}_1, \vec{v}_2\}$

Licença CC-BY-SA-3.0. Contato: livroscolaborativos@gmail.com

12.2 Problemas de otimização restritos à esfera

Nesta seção, estudamos alguns problemas de otimização bem específicos. Mais precisamente, podemos encontrar valores de máximo e mínimo de formas quadráticas

$$Q(\vec{x}) = \vec{x}^T A\vec{x}$$

sujeitos à restrição

$$x_1^2 + x_2^2 + \cdots + x_n^2 = 1,$$

que caracteriza os pontos $(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ que tem comprimento unitário, isto é, a esfera de raio 1 e centro na origem em $\mathbb{R}^n$. Em outras palavras, nosso problema de otimização restrito consiste em achar, na esfera unitária, os vetores em que a forma quadrática Q assume seu maior valor e seu menor valor.

Mais resumidamente, queremos:

$$\text{maximizar/minimizar } Q(\vec{x}) \text{ sujeito à restrição } \|\vec{x}\| = 1.$$

Ideia por trás do método. Sendo A uma matriz simétrica, todos os seus autovalores são reais. Vamos ordená-los em forma crescente

$$\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n.$$

Podemos também encontrar uma base ortonormal de autovetores e, com eles, formar uma matriz ortogonal que diagonaliza A:

$$P = \begin{bmatrix} | & | & & | \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_n \\ | & | & & | \end{bmatrix} \implies P^{-1}AP = D = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}.$$

Nota que P, sendo ortogonal, preserva comprimentos (ver Observação 119 na seção seguinte). Assim, escrevendo $\vec{x} = P\vec{y}$, temos que $\|\vec{x}\| = 1$ se, e somente se, $\|\vec{y}\| = 1$. Além disso, esta matriz P é a matriz ortogonal que diagonaliza A, de modo que, pelo que vimos na seção anterior,

$$\vec{x}^T A\vec{x} = \vec{y}^T D\vec{y}.$$

Mas a forma quadrática $\vec{y}^T D\vec{y}$ é mais fácil de maximizar: sendo λ_n a maior das entradas da diagonal principal de D, temos (já que cada $\lambda_j \leq \lambda_n$ e $\|\vec{y}\| = 1$)

$$\vec{y}^T D\vec{y} = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2 \leq \lambda_n(y_1^2 + y_2^2 + \cdots + y_n^2) = \lambda_n.$$

Logo, λ_n é uma cota superior para $\vec{y}^T D\vec{y}$. Além disso, já que D é uma matriz diagonal, esta cota superior é atingida pelo vetor $\vec{y} = \vec{e}_n$, o último vetor da base canônica de $\mathbb{R}^n$. Traduzindo estas informações para a variável $\vec{x}$, temos que λ_n é o valor máximo de $\vec{x}^T A\vec{x}$ (pois é igual a $\vec{y}^T D\vec{y}$) na esfera unitária. Além disso, este valor máximo é atingido quando $\vec{x} = P\vec{e}_n = \vec{v}_n$, um autovetor unitário associado com o maior autovalor λ_n.

Exercício 13. Fazer toda a análise acima para o problema de minimização. *Dica*: olhar para o menor autovalor.

Exemplo 117. Vamos maximizar a forma quadrática

$$Q(x_1, x_2, x_3) = 7x_1^2 + x_2^2 + 7x_3^2 - 8x_1x_2 - 4x_1x_3 - 8x_2x_3,$$

sujeito à restrição $x_1^2 + x_2^2 + x_3^2 = 1$.

Note que podemos escrever

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \implies Q(\vec{x}) = \vec{x}^T \begin{bmatrix} 7 & -4 & -2 \\ -4 & 1 & -4 \\ -2 & -4 & 7 \end{bmatrix} \vec{x}.$$

Licença CC-BY-SA-3.0. Contato: livroscolaborativos@gmail.com

Precisamos determinar o maior autovalor e um autovetor unitário associado. O polinômio característico de A é dado por

$$\det \begin{bmatrix} 7-\lambda & -4 & -2 \\ -4 & 1-\lambda & -4 \\ -2 & -4 & 7-\lambda \end{bmatrix} = -\lambda^3 + 15\lambda^2 - 27\lambda - 243,$$

cujas raízes são $\lambda_1 = -3$ e $\lambda_2 = 9$. Portanto, o valor máximo que a forma quadrática Q assume na esfera unitária é 9.

Além disso, sabemos que qualquer um autovetor unitário associado com o autovalor $\lambda_2 = 9$ assume este valor máximo. Resolvemos

$$A - 9I = \begin{bmatrix} -2 & -4 & -2 \\ -4 & -8 & -4 \\ -2 & -4 & -2 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \implies \vec{v} = \begin{bmatrix} -2v_2 - v_3 \\ v_2 \\ v_3 \end{bmatrix} = v_2 \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} + v_3 \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}.$$

Um autovetor (unitário!) possível é

$$\frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{bmatrix}.$$

Conferimos que

$$Q\left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right) = \frac{7}{2} + 0 + \frac{7}{2} - 0 + \frac{4}{2} - 0 = 9. \ \lhd$$

Observação 118 (Opcional)**.** Em cursos de cálculo com várias variáveis, estudamos multiplicadores de Lagrange, uma ferramenta para estudar problemas de otimização com restrição. De uma maneira geral, queremos

$$\text{minimizar } f(x_1, x_2, \cdots, x_n) \text{ com a restrição } g(x_1, x_2, \cdots, x_n) = 0.$$

O método dos multiplicadores de Lagrange nos dá uma condição necessária: se um ponto é de máximo/mínimo com restrição, então existe uma constante λ tal que, neste ponto,

$$\vec{\nabla} f = \lambda \vec{\nabla} g.$$

Esta constante λ é conhecida como um **multiplicador de Lagrange**. Como mencionamos, esta é uma condição necessária, de modo que não garante que o ponto é de mínimo/máximo. Na prática, depois de encontrarmos os possíveis valores de $\vec{x}$ e de λ, podemos substituir na função para ver quais são os pontos de mínimo e quais são os de máximo.

Vejamos como esta técnica se aplicaria no Exemplo 117 que já resolvemos acima. Vamos considerar $f(x_1, x_2, x_3) = Q(x_1, x_2, x_3)$ e $g(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2 - 1$ Se estivermos em um ponto de máximo, então existe a constante λ como acima. Calculamos:

$$\vec{\nabla} Q = (14x_1 - 8x_2 - 4x_3, 2x_2 - 8x_1 - 8x_3, 14x_3 - 4x_1 - 8x_2)$$

e

$$\vec{\nabla} g = (2x_1, 2x_2, 2x_3).$$

A condição $\vec{\nabla} Q = \lambda \vec{\nabla} g$ nos diz que devemos resolver

$$\begin{array}{r} 14x_1 - 8x_2 - 4x_3 = 2\lambda x_1 \\ -8x_1 + 2x_2 - 8x_3 = 2\lambda x_2 \\ -4x_1 - 8x_2 + 14x_3 = 2\lambda x_3 \end{array} \xrightarrow[\text{em cada eq.}]{\text{Simplifica um 2}} \begin{bmatrix} 7 & -4 & -2 \\ -4 & 1 & -4 \\ -2 & -4 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \lambda \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}.$$

Chegamos, por outro caminho, à mesma conclusão: λ deve ser um autovalor e $\vec{x}$ um autovetor (unitário, já que $g(x_1, x_2, x_3) = 0$) associado. $\lhd$

Licença CC-BY-SA-3.0. Contato: livroscolaborativos@gmail.com

12.3 Matrizes simétricas e prova do Teorema Espectral

Nesta seção vamos utilizar a notação $\langle \vec{u}, \vec{v} \rangle$ para o produto escalar, ao invés de $\vec{u} \cdot \vec{v}$, para enfatizar qual é o vetor $\vec{u}$ e qual é o vetor $\vec{v}$ (em especial quando um deles está sendo multiplicado por uma matriz).

Observação 119. Uma matriz quadrada A, não necessariamente simétrica, cujos coeficientes são reais, satisfaz

$$\langle A\vec{x}, \vec{y} \rangle = \langle \vec{x}, A^T\vec{y} \rangle, \quad \text{para todos } \vec{x}, \vec{y} \in \mathbb{R}^n.$$

Isto se pode verificar ao analisar os coeficientes de $A\vec{x}$ e de $A^T\vec{y}$. Assim, uma matriz A de ordem $n \times n$ é **simétrica** se, e somente se, satisfaz

$$\langle A\vec{x}, \vec{y} \rangle = \langle \vec{x}, A\vec{y} \rangle, \quad \text{para todos } \vec{x}, \vec{y} \in \mathbb{R}^n.$$

Outras formas de escrever esta mesma coisa (concorda?) são

$$A\vec{x} \cdot \vec{y} = \vec{x} \cdot A\vec{y}, \quad \text{para todos } \vec{x}, \vec{y} \in \mathbb{R}^n$$

e

$$\vec{y}^T A\vec{x} = \vec{x}^T A\vec{y}, \quad \text{para todos } \vec{x}, \vec{y} \in \mathbb{R}^n.$$

Observação 120. A observação anterior nos permite estender a definição de matriz simétrica para transformações lineares, pois é independente do sistema de coordenadas utilizado (depende apenas de uma escolha de produto interno/escalar).

A equivalência que foi apresentada na Observação 119 é muito útil na análise de matrizes simétricas. Vamos ver como se aplica diversas vezes na demonstração do Teorema Espectral que será dividida em três proposições preliminares. Cada uma delas apresenta, em ordem, o procedimento usual que viemos utilizando para diagonalizar matrizes: análise dos autovalores, construção dos autoespaços e verificação de que é possível formar uma base de autovetores. Neste caso de uma matriz simétrica, veremos que, além de tudo, é possível encontrar uma base ortonormal de autovetores.

Proposição 121. *Os autovalores de uma matriz simétrica A, de entradas reais, são números reais.*

Demonstração. Seja λ uma raiz do polinômio característico de A, que pode ser um número real ou um número complexo. Podemos encontrar um vetor $\vec{v}$, de entradas possivelmente complexas, tal que $A\vec{v} = \lambda\vec{v}$. A menos de um passo de normalização, podemos pensar que $\|\vec{v}\| = 1$. Vamos verificar que de fato λ é um número real. Para isto, vamos mostrar que λ é igual ao seu conjugado, $\lambda = \bar{\lambda}$, de forma que a parte imaginária de λ deve ser igual a zero. Sendo A simétrica, temos que

$$\bar{\lambda} = \overline{\lambda\|\vec{v}\|^2} = \overline{\langle \lambda\vec{v}, \vec{v} \rangle} = \overline{\langle A\vec{v}, \vec{v} \rangle} = \overline{\langle \vec{v}, A\vec{v} \rangle} = \overline{\langle \vec{v}, \lambda\vec{v} \rangle} = \lambda\overline{\langle \vec{v}, \vec{v} \rangle} = \lambda.$$

A simetria de A foi necessária (apenas) no quarto sinal de igualdade acima. Ainda, na penúltima igualdade, utilizamos sem muitos comentários uma propriedade do produto interno em espaços vetoriais com escalares complexos, a saber, que

$$\langle \vec{u}, \vec{v} \rangle \stackrel{\text{def}}{=} \sum_{i=1}^{n} u_i \bar{v}_i.$$

Os conjugados das componentes do segundo vetor aparecem para que a fórmula usual para o comprimento continue sendo válida:

$$\|\vec{u}\|^2 = \langle \vec{u}, \vec{u} \rangle = \sum_{i=1}^{n} u_i \bar{u}_i = \sum_{i=1}^{n} |u_i|^2.$$

Desta forma, continua valendo que

$$\langle k\vec{u}, \vec{v} \rangle = k\langle \vec{u}, \vec{v} \rangle, \quad \text{enquanto que} \quad \langle \vec{u}, k\vec{v} \rangle = \bar{k}\langle \vec{u}, \vec{v} \rangle. \qquad \square$$

Licença CC-BY-SA-3.0. Contato: livroscolaborativos@gmail.com

Proposição 122. *Se A é uma matriz simétrica, então autovetores associados a autovalores distintos são ortogonais.*

Demonstração. Sejam λ_1 e λ_2 autovalores distintos de A e $\vec{v}_1, \vec{v}_2$ respectivos autovetores. Então (observe que é necessário que A seja simétrica na segunda e na última igualdade!)

$$\lambda_1\langle\vec{v}_1, \vec{v}_2\rangle = \langle\lambda_1\vec{v}_1, \vec{v}_2\rangle = \langle A\vec{v}_1, \vec{v}_2\rangle = \langle\vec{v}_1, A\vec{v}_2\rangle = \overline{\lambda}_2\langle\vec{v}_1, \vec{v}_2\rangle = \lambda_2\langle\vec{v}_1, \vec{v}_2\rangle,$$

de modo que $(\lambda_1 - \lambda_2)\langle\vec{v}_1, \vec{v}_2\rangle = 0$. Já que estamos supondo que os autovalores são distintos, $\lambda_1 \neq \lambda_2$, concluimos que $\langle\vec{v}_1, \vec{v}_2\rangle = 0$. □

Proposição 123. *Seja A uma matriz simétrica. Então, a dimensão do autoespaço associado a um autovalor é igual à multiplicidade deste autovalor.*

Demonstração. Suponhamos que λ_1 é um autovalor de A de multiplicidade k. Seja $\vec{v}_1$ um autovetor unitário associado. Podemos encontrar uma base ortonormal de $\mathbb{R}^n$ da qual $\vec{v}_1$ faz parte, isto é, podemos encontrar vetores unitários $\vec{y}_j$ de modo que

$$\{\vec{v}_1, \vec{y}_2, \vec{y}_3, \ldots, \vec{y}_n\}$$

é uma base ortonormal de $\mathbb{R}^n$.

Considere a matriz Q cujas colunas são os elementos desta base:

$$Q = \begin{bmatrix} | & | & | & & | \\ \vec{v}_1 & \vec{y}_2 & \vec{y}_3 & \cdots & \vec{y}_n \\ | & | & | & & | \end{bmatrix},$$

que é, consequentemente, ortogonal. Vamos calcular $Q^{-1}AQ = Q^TAQ$. O produto AQ pode ser interpretado como o produto de A por cada uma das colunas de Q:

$$AQ = \begin{bmatrix} | & | & | & & | \\ A\vec{v}_1 & A\vec{y}_2 & A\vec{y}_3 & \cdots & A\vec{y}_n \\ | & | & | & & | \end{bmatrix} = \begin{bmatrix} | & | & | & & | \\ \lambda_1\vec{v}_1 & A\vec{y}_2 & A\vec{y}_3 & \cdots & A\vec{y}_n \\ | & | & | & & | \end{bmatrix}.$$

Logo,

$$Q^TAQ = \begin{bmatrix} — & \vec{v}_1 & — \\ — & \vec{y}_2 & — \\ — & \vec{y}_3 & — \\ & \vdots & \\ — & \vec{y}_n & — \end{bmatrix} \begin{bmatrix} | & | & | & & | \\ \lambda_1\vec{v}_1 & A\vec{y}_2 & A\vec{y}_3 & \cdots & A\vec{y}_n \\ | & | & | & & | \end{bmatrix}$$

$$\implies Q^TAQ = \begin{bmatrix} \lambda_1\langle\vec{v}_1, \vec{v}_1\rangle & \langle\vec{v}_1, A\vec{y}_2\rangle & \langle\vec{v}_1, A\vec{y}_3\rangle & \cdots & \langle\vec{v}_1, A\vec{y}_n\rangle \\ \lambda_1\langle\vec{y}_2, \vec{v}_1\rangle & \langle\vec{y}_2, A\vec{y}_2\rangle & \langle\vec{y}_2, A\vec{y}_3\rangle & \cdots & \langle\vec{y}_2, A\vec{y}_n\rangle \\ \lambda_1\langle\vec{y}_3, \vec{v}_1\rangle & \langle\vec{y}_3, A\vec{y}_2\rangle & \langle\vec{y}_3, A\vec{y}_3\rangle & \cdots & \langle\vec{y}_3, A\vec{y}_n\rangle \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \lambda_1\langle\vec{y}_n, \vec{v}_1\rangle & \langle\vec{y}_n, A\vec{y}_2\rangle & \langle\vec{y}_n, A\vec{y}_3\rangle & \cdots & \langle\vec{y}_n, A\vec{y}_n\rangle \end{bmatrix}$$

Observamos que, pela simetria de A e por ser $\{\vec{v}_1, \vec{y}_2, \vec{y}_3, \ldots, \vec{y}_n\}$ um conjunto ortonormal, temos

$$\langle\vec{v}_1, A\vec{y}_j\rangle = \langle A\vec{v}_1, \vec{y}_j\rangle = \lambda_1\langle\vec{v}_1, \vec{y}_j\rangle = 0, \quad \text{para qualquer índice } j.$$

Logo,

$$\implies Q^TAQ = \begin{bmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & b_{22} & b_{23} & \cdots & b_{2n} \\ 0 & b_{32} & b_{33} & \cdots & b_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & b_{n2} & b_{n3} & \cdots & b_{nn} \end{bmatrix} \stackrel{\text{def}}{=} \hat{A}.$$

Chamamos os coeficientes restantes de b_{ij} simplesmente porque não nos interessa muito quanto valem (notamos, no entanto, que a simetria de A implica que $b_{ij} = b_{ji}$). Observamos que

$$\det(A - \lambda I) = \det\left(Q^{-1}(A - \lambda I)Q\right) = \det\left(\hat{A} - \lambda I\right) = (\lambda_1 - \lambda)\cdot\det(B - \lambda I),$$

Licença CC-BY-SA-3.0. Contato: livroscolaborativos@gmail.com

onde B é a matriz de ordem $(n-1)\times(n-1)$ dada por (b_{ij}). Além disso, se a multiplicidade de λ_1 como autovalor de A for maior do que 1, devemos ter, pela igualdade acima, que λ_1 também é um autovalor de B. Desta forma, $\hat{A}$ possui um autovetor unitário $\vec{v}_2$ associado com λ_1 que está contido em $\mathrm{Span}\{\vec{y}_2,\vec{y}_3,\dots,\vec{y}_n\}$ e, portanto, é distinto e linearmente independente a $\vec{v}_1$. O processo de Gram-Schmidt permite obter $\{\vec{v}_1,\vec{v}_2\}$ ortonormais.

A partir daí, repetimos o processo completando o conjunto $\{\vec{v}_1,\vec{v}_2\}$ a uma base ortonormal $\{\vec{v}_1,\vec{v}_2,\vec{y}_3,\dots,\vec{y}_n\}$ de $\mathbb{R}^n$. Daí vamos obter um terceiro autovetor ortogonal aos anteriores e unitário. Procedemos desta meneira até encontrar k autovetores ortogonais e unitários associados ao autovalor λ_1. Assim, o autoespaço $\mathrm{Nul}(A-\lambda_1 I)$ tem dimensão k, que é igual à multiplicidade do autovalor λ_1. □

12.3.1 Prova do Teorema Espectral

O Teorema Fundamental da Álgebra afirma que todo polinômio não constante de grau n possui exatamente n raízes complexas. Sendo assim, qualquer matriz A de ordem $n\times n$ possui n autovalores complexos. No entanto, sendo A uma matriz simétrica, a Proposição 121 garante que todas estes n autovalores são todos reais. Vamos denotar por

$$\lambda_1,\lambda_2,\dots,\lambda_k\in\mathbb{R}$$

os autovalores distintos de A e por $m_1,m_2,\dots,m_k$ suas respectivas multiplicidades. Temos, em particular, que $m_1+m_2+\cdots+m_k=n$.

Pela Proposição 123, a dimensão do autoespaço associado a λ_1 é m_1. Podemos encontrar (da maneira usual, por escalonamento) uma base de autovetores

$$\mathrm{Nul}(A-\lambda_1 I)=\mathrm{Span}\{\vec{u}_1,\vec{u}_2,\dots,\vec{u}_{m_1}\}.$$

Estes autovetores, possivelmente não são ortogonais, mas, aplicando o Processo de Gram-Schmidt, podemo obter uma base ortonormal de $\mathrm{Nul}(A-\lambda_1 I)$:

$$\mathrm{Nul}(A-\lambda_1 I)=\mathrm{Span}\{\vec{v}_1,\vec{v}_2,\dots,\vec{v}_{m_1}\}.$$

Fazendo este mesmo procedimento para cada um dos autovalores, obtemos n autovetores, que, consequentemente, formam uma base para $\mathbb{R}^n$. Logo, A é diagonalizável.

Além disso, a Proposição 122 garante que todo autovetor de um autoespaço é ortogonal a todo autovetor de outro autoespaço (associado a um autovalor diferente). Temos então k autoespaços, cujas bases são ortonormais e elementos da base de um autoespaço são ortogonais aos elementos das outras bases. Segue daí que a união das bases dos autoespaços, como consideramos, é uma base *ortonormal* para $\mathbb{R}^n$.

Explicitamente, a matriz

$$P=\begin{bmatrix} | & | & & | \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_n \\ | & | & & | \end{bmatrix}$$

formada pelos autovetores é ortogonal e

$$P^{-1}AP=D=\begin{bmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ 0 & 0 & \lambda_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_n \end{bmatrix}.$$

Sendo P ortogonal, tem-se ainda que $P^{-1}=P^T$.

Licença CC-BY-SA-3.0. Contato: livroscolaborativos@gmail.com

Colaboradores

Aqui você encontra a lista de colaboradores do livro. Esta lista contém somente aqueles que explicitamente se manifestaram a favor de terem seus nomes registrados aqui. A lista completa de colaborações pode ser obtida no repositório GitHub do livro:

https://github.com/livroscolaborativos/AlgebraLinear

Além das colaborações via GitHub, o livro também recebe colaborações via discussões, sugestões e avisos deixados em nossa lista de e-mails:

livro_colaborativo@googlegroups.com

Estas colaborações não estão listadas aqui, mas podem ser vistas no site do grupo de e-mails.

Caso encontre algum equívoco ou veja seu nome listado aqui por engano, por favor, entre em contato conosco por e-mail:

livroscolaborativos@gmail.com

ou via o repositório GitHub.

Tabela 12.1: Lista de colaboradores

Nome	Afiliação	E-Mail	1ł Contribuição