Curso: Métodos Numéricos II
Professor: Creto Augusto Vidal
Semestre: 2021.1
Aula # 17

1. Objetivo: Discutir os conceitos básicos de Autovalores e Autovetores de matrizes simétricas

2. Conceitos preliminares

Antes de discutirmos sobre matrizes simétricas e sobre o que são autovalores e autovetores, vamos revisar alguns conceitos de Álgebra Linear.

2.1 Espaço vetorial

Sugiro que assistam o seguinte vídeo bem descontraído e rápido em que espaço vetorial é definido: https://www.youtube.com/watch?v=ozwodzD5bJM

Vamos nos restringir ao espaço vetorial chamado $\mathbb{R}^n$ e vamos pensar num conjunto de escalares chamado $\mathbb{R}$ (o conjunto dos números reais).

$\mathbb{R}^n$ também é um conjunto, mas ele tem que ter algumas propriedades para ser chamado e espaço vetorial.

2.1.1 Elementos do "conjunto" $\mathbb{R}^n$

O nome dado a um elemento desse conjunto não importa. O que importa, é a composição desse elemento. Assim, um elemento $\mathbf{v} \in \mathbb{R}^n$ tem a seguinte composição

(1) $\mathbf{v} = (v_1, v_2, \cdots, v_n)$.

Os elementos do "conjunto" $\mathbb{R}^2$, por exemplo, serão escritos como um par ordenado, isto é,

$$\mathbf{v} = (v_1, v_2) = (3.5, 4.2) \text{ ou } \mathbf{w} = (w_1, w_2) = (4.2, -3.5).$$

Os elementos do "conjunto" $\mathbb{R}^3$, de maneira análoga, serão escritos como uma tripla, isto é,

$$\mathbf{v} = (v_1, v_2, v_3) = (2.0, -3.4, 5.2) \text{ ou } \mathbf{w} = (w_1, w_2, w_3) = (5.2, 5.0, -4.1).$$

2.1.2 Propriedades que darão ao $\mathbb{R}^n$ um título mais pomposo.

- **Operação de Adição entre os elementos de $\mathbb{R}^n$**

Considere dois elementos $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$. Assim

(2) $\mathbf{v} + \mathbf{w} = (v_1, v_2, \cdots, v_n) + (w_1, w_2, \cdots, w_n) = (v_1 + w_1, v_2 + w_2, \cdots, v_n + w_n) = \mathbf{x}$

Note que o resultado da adição mostrada na equação (2) tem a mesma estrutura de um elemento de $\mathbb{R}^n$. Portanto, podemos dizer que $\mathbf{x} \in \mathbb{R}^n$ e que o conjunto $\mathbb{R}^n$ tem a propriedade de **fechamento** com relação à operação de adição, ou seja, o resultado da adição de quaisquer dois elementos desse conjunto também é um elemento do conjunto.

Pela equação (2) podemos escrever

(3) $x_i = v_i + w_i = w_i + v_i, \; i = 1, 2, \cdots, n.$

Assim, podemos afirmar que

(4) $\mathbf{x} = \mathbf{v} + \mathbf{w} = \mathbf{w} + \mathbf{v}$ (a adição de elementos do $\mathbb{R}^n$ é **comutativa**)

O elemento cujas componentes são todas nulas é identificado por $\mathbf{0}$, isto é

(5) $\mathbf{0} = \left(\underbrace{0, 0, \cdots, 0}_{n}\right)$.

Esse elemento é chamado de **elemento neutro** da adição pois, se ele for adicionado a qualquer elemento, o resultado é o próprio elemento, isto é,

(6) $\mathbf{v} + \mathbf{0} = \mathbf{0} + \mathbf{v} = \mathbf{v}.$

Pelas equações (2) e (3), fica claro que, se as componentes de **w** forem todas iguais às componentes de **v** mas com sinais contrários,

(7) $x_i = v_i + w_i = v_i + (-v_i) = 0, \; i = 1, 2, \cdots, n.$

Assim, para qualquer **v**, existe um elemento $\mathbf{w} = -\mathbf{v}$ de modo que

(8) $\mathbf{v} + \mathbf{w} = \mathbf{v} + (-\mathbf{v}) = \mathbf{0}.$

Por todas as equações anteriores, podemos concluir que, dados quaisquer dois vetores $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$, existe um vetor único $\mathbf{x} \in \mathbb{R}^n$ tal que

(9) $\mathbf{v} + \mathbf{x} = \mathbf{w}.$

- **Operação de multiplicação dos elementos de $\mathbb{R}^n$ por um escalar pertencente a $\mathbb{R}$**

Considere o elemento $\mathbf{v} \in \mathbb{R}^n$ e o escalar $\alpha \in \mathbb{R}$. Assim

(10) $\alpha\mathbf{v} = \alpha(v_1, v_2, \cdots, v_n) = (\alpha v_1, \alpha v_2, \cdots, \alpha v_n) = \mathbf{x}$

Note que o resultado da multiplicação de **v** pelo escalar α mostrada na equação (10) tem a mesma estrutura de um elemento de $\mathbb{R}^n$. Portanto, podemos dizer que $\mathbf{x} \in \mathbb{R}^n$ e que o conjunto $\mathbb{R}^n$ tem a propriedade de **fechamento** com relação à operação de multiplicação por um escalar, ou seja, o resultado da multiplicação de qualquer elemento desse conjunto por um escalar $\alpha \in \mathbb{R}$ também é um elemento do conjunto.

Pela equação (10) podemos escrever que

(11) $x_i = \alpha v_i = v_i \alpha, \; i = 1, 2, \cdots, n.$

Suponha que o escalar $\alpha \in \mathbb{R}$ seja o resultado da soma de dois escalares β e $\gamma \in \mathbb{R}$, isto é,

(12) $\alpha = \beta + \gamma,$

assim, substituindo (12) em (10) temos

(13)
$$\left.\begin{aligned} (\beta+\gamma)\mathbf{v} &= ((\beta+\gamma)v_1, (\beta+\gamma)v_2, \cdots, (\beta+\gamma)v_n) \\ &= \left((\beta v_1 + \gamma v_1), (\beta v_2 + \gamma v_2), \cdots (\beta v_n + \gamma v_n)\right) \\ &= (\beta v_1, \beta v_2, \cdots, \beta v_n) + (\gamma v_1, \gamma v_2, \cdots, \gamma v_n) \\ &= \beta\mathbf{v} + \gamma\mathbf{v} \end{aligned}\right\} \text{Propriedade}$$

distributiva da multiplicação de um vetor com relação à adição de dois escalares.

De maneira análoga, dados dois elementos $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$ e um escalar $\alpha \in \mathbb{R}$, podemos escrever

(14)
$$\left.\begin{aligned} \alpha(\mathbf{v}+\mathbf{w}) &= \left(\alpha(v_1+w_1), \alpha(v_2+w_2), \cdots, \alpha(v_n+w_n)\right) \\ &= (\alpha v_1, \alpha v_2, \cdots, \alpha v_n) + (\alpha w_1, \alpha w_2, \cdots, \alpha w_n) \\ &= \alpha\mathbf{v} + \alpha\mathbf{w} \end{aligned}\right\} \text{Propriedade } \textbf{distributiva}$$

da multiplicação de um escalar com relação à adição de dois vetores.

Também, dados dois escalares α e $\beta \in \mathbb{R}$ e $\mathbf{v} \in \mathbb{R}^n$, podemos escrever

$$\left.\begin{aligned} \alpha(\beta\mathbf{v}) &= \alpha\big((\beta v_1),(\beta v_2),\cdots,(\beta v_n)\big) \\ &= \big((\alpha\beta v_1),(\alpha\beta v_2),\cdots,(\alpha\beta v_n)\big) \\ &= (\alpha\beta)(v_1,v_2,\cdots,v_n) \\ &= (\alpha\beta)\mathbf{v} \end{aligned}\right\} \quad (15)$$

Propriedade **associativa do produto de dois escalares por um vetor.**

Pelas equações (10) e (11), podemos notar que, se $\alpha = 1$

(16) $\alpha\mathbf{v} = 1(v_1, v_2, \cdots, v_n) = \mathbf{v}$.

Finalmente, se todas as operações e propriedades expressas nas equações de (2) a (16) forem obedecidas, $\mathbb{R}^n$ é um **espaço vetorial**.

2.2 Bases do espaço vetorial $\mathbb{R}^n$

Uma base de $\mathbb{R}^n$ é um conjunto de n elementos de $\mathbb{R}^n$ que podem ser combinados, usando as equações (2) e (10), para representar qualquer elemento de $\mathbb{R}^n$. Assim, o elemento $\mathbf{v} \in \mathbb{R}^n$ pode ser representado por

(16) $\mathbf{v} = \alpha_1\mathbf{b}_1 + \alpha_2\mathbf{b}_2 + \cdots + \alpha_n\mathbf{b}_n$,

onde os elementos $\mathbf{b}_i \in \mathbb{R}^n$, com $i = 1, 2, \cdots, n$, são os elementos que constituem a base do $\mathbb{R}^n$.

Obviamente, nenhum dos elementos que constituem a base pode ser representado como combinação dos demais elementos da base. Isso significa, que os elementos que constituem a base são **Linearmente Independentes** (L.I.). Um conjunto com menos de n elementos L.I. não consegue representar qualquer $\mathbf{v} \in \mathbb{R}^n$.

Vamos dar alguns exemplos no $\mathbb{R}^2$.

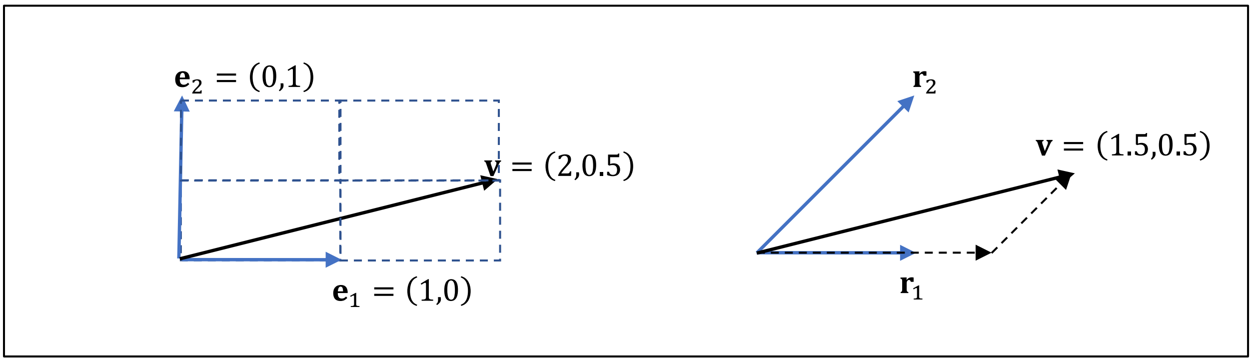

Figura 1. Representação do vetor **v** em duas bases distintas do $\mathbb{R}^2$.

Na Figura 1, o vetor $\mathbf{v} \in \mathbb{R}^2$ está representado por meio de duas bases distintas. Na primeira base, chamada base canônica, os elementos da base são $\mathbf{e}_1$ e $\mathbf{e}_2$. Esses dois vetores têm comprimento igual a 1 (vetores unitários), e são perpendiculares entre si. Na segunda base, os elementos da base são $\mathbf{r}_1$ e $\mathbf{r}_2$. Esses dois vetores são linearmente independentes, mas não são perpendiculares entre si. Além disso, o vetor $\mathbf{r}_1$ é um vetor unitário, mas o vetor $\mathbf{r}_2$ não é. Segundo a equação (16), na primeira base o vetor é representado como

(17) $\mathbf{v} = \alpha_1\mathbf{e}_1 + \alpha_2\mathbf{e}_2 = 2\mathbf{e}_1 + 0.5\mathbf{e}_2 = 2(1,0) + 0.5(0,1) = (2,0) + (0,0.5) = (2,0.5)_{\mathbf{b}}$.

Na segunda base, o vetor é representado como

(18) $\mathbf{v} = \beta_1\mathbf{r}_1 + \beta_2\mathbf{r}_2 = 1.5\mathbf{r}_1 + 0.5\mathbf{r}_2 = (1.5,0.5)_{\mathbf{r}}$

Isso significa que para chegarmos na ponta do vetor **v**, saindo da origem, devemos andar 1.5 vezes o tamanho de $\mathbf{r}_1$ na sua direção e andar 0.5 vezes o tamanho de $\mathbf{r}_2$ na sua direção. Se representarmos os próprios vetores $\mathbf{r}_1$ e $\mathbf{r}_2$ na base canônica teríamos

(19) $\mathbf{r}_1 = 1\mathbf{e}_1 + 0\mathbf{e}_2 = (1,0)_{\mathbf{e}} = \mathbf{e}_1$

(20) $\mathbf{r}_2 = 1\mathbf{e}_1 + 1\mathbf{e}_2 = (1,1)_{\mathbf{e}}$.

Assim, substituindo (19) e (20) em (18) temos

(21) $\mathbf{v} = 1.5\mathbf{r}_1 + 0.5\mathbf{r}_2 = 1.5(1,0)_{\mathbf{e}} + 0.5(1,1)_{\mathbf{e}} = (1.5,0)_{\mathbf{e}} + (0.5,0.5)_{\mathbf{e}} = (2,0.5)_{\mathbf{e}}$.

Note que, em (21) o vetor **v** novamente foi representado na base **e**.

2.3 Transformações lineares de $\mathbb{R}^n$ para $\mathbb{R}^n$.

Uma transformação linear de $\mathbb{R}^n$ para $\mathbb{R}^n$, é um mapeamento de um elemento do primeiro espaço vetorial para um elemento do segundo espaço vetorial que tem de respeitar algumas regras. Imagine uma transformação como operações, embutidas em uma caixa preta, sobre um elemento que entra na caixa preta, e que produzem um elemento de saída da caixa preta (Figura 2).

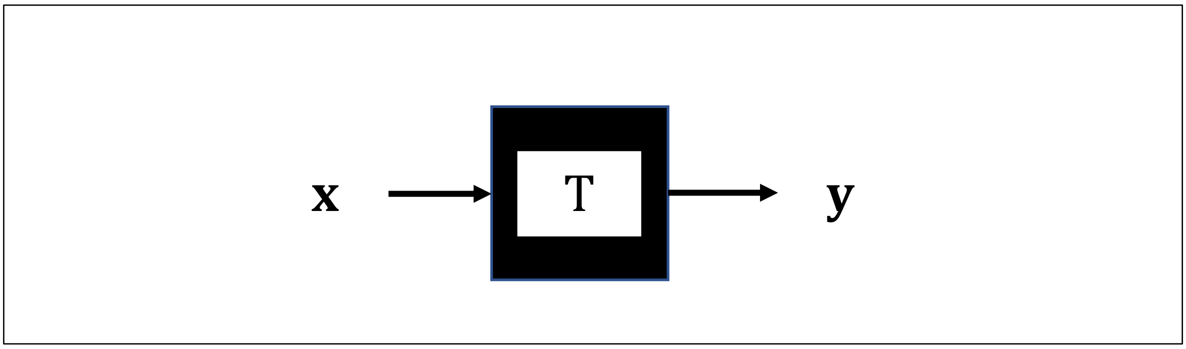

Figure 2. Transformação entre espaços vetoriais

A transformação T é linear se acontecer o seguinte comportamento com os elementos de entrada e saída da caixa preta:

(21)
$$\begin{aligned} T(\mathbf{x}) &= \mathbf{y} \\ T(\mathbf{w}) &= \mathbf{z} \\ T(\mathbf{x}+\mathbf{w}) &= \mathbf{y}+\mathbf{z} \end{aligned}$$

e

(22)
$$\begin{aligned} T(\mathbf{x}) &= \mathbf{y} \\ T(\alpha\mathbf{x}) &= \alpha\mathbf{y}. \end{aligned}$$

Exemplo de transformação linear do $\mathbb{R}^3$ para $\mathbb{R}^3$.

(23)
$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}_{\mathbf{b}} = \begin{pmatrix} 3x_1 + 2x_2 + x_3 \\ x_1 - 2x_2 \\ 4x_2 - 2x_3 \end{pmatrix}_{\mathbf{b}}$$

Se entrarmos as componentes do vetor **w**

(24)
$$\begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix}_{\mathbf{b}} = \begin{pmatrix} 3w_1 + 2w_2 + w_3 \\ w_1 - 2w_2 \\ 4w_2 - 2w_3 \end{pmatrix}_{\mathbf{b}}$$

Agora vamos entrar com a componentes de um vetor $\mathbf{s} = \mathbf{x} + \mathbf{w}$. Assim

$$
(25)\quad \begin{aligned}
\begin{pmatrix} t_1 \\ t_2 \\ t_3 \end{pmatrix}_{\mathbf{b}} &= \begin{pmatrix} 3(x_1 + w_1) + 2(x_2 + w_2) + (x_3 + w_3) \\ (x_1 + w_1) - 2(x_2 + w_2) \\ 4(x_2 + w_2) - 2(x_3 + w_3) \end{pmatrix}_{\mathbf{b}} \\
&= \begin{pmatrix} 3x_1 + 2x_2 + x_3 \\ x_1 - 2x_2 \\ 4x_2 - 2x_3 \end{pmatrix}_{\mathbf{b}} + \begin{pmatrix} 3w_1 + 2w_2 + w_3 \\ w_1 - 2w_2 \\ 4w_2 - 2w_3 \end{pmatrix}_{\mathbf{b}} \\
&= \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}_{\mathbf{b}} + \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix}_{\mathbf{b}}
\end{aligned}
$$

Parte da exigência para ser uma transformação linear foi satisfeita. Vamos a segunda exigência, isto é, vamos entrar com o vetor $\mathbf{m} = \alpha\mathbf{x}$. Assim

$$
(26)\quad \begin{aligned}
\begin{pmatrix} t_1 \\ t_2 \\ t_3 \end{pmatrix}_{\mathbf{b}} &= \begin{pmatrix} 3(\alpha x_1) + 2(\alpha x_2) + (\alpha x_3) \\ (\alpha x_1) - 2(\alpha x_2) \\ 4(\alpha x_2) - 2(\alpha x_3) \end{pmatrix}_{\mathbf{b}} \\
&= \alpha \begin{pmatrix} 3x_1 + 2x_2 + x_3 \\ x_1 - 2x_2 \\ 4x_2 - 2x_3 \end{pmatrix}_{\mathbf{b}} \\
&= \alpha \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}_{\mathbf{b}} .
\end{aligned}
$$

Os resultados das equações (25) e (26) comprovam as exigências (21) e (22) e a transformação é linear.

2.3 Matriz de uma Transformação linear de $\mathbb{R}^n$ para $\mathbb{R}^n$.

Nos exemplos de transformação linear mostrados nas equações (23) a (26), nós indicamos que as componentes dos vetores se referem à base **b**.

Vamos representar um vetor $\mathbf{x} = \alpha_1\mathbf{b}_1 + \alpha_2\mathbf{b}_2 + \cdots + \alpha_n\mathbf{b}_n$ na base **b** e passá-lo como entrada para a caixa preta da transformação linear. Vamos assumir que o resultado é um outro vetor também representado na base **b**, isto é, $\mathbf{y} = \beta_1\mathbf{b}_1 + \beta_2\mathbf{b}_2 + \cdots + \beta_n\mathbf{b}_n$. Assim,

$$
(27)\quad \begin{aligned}
\mathbf{y} &= \beta_1\mathbf{b}_1 + \beta_2\mathbf{b}_2 + \cdots + \beta_n\mathbf{b}_n \\
&= T(\mathbf{x}) \\
&= T(\alpha_1\mathbf{b}_1 + \alpha_2\mathbf{b}_2 + \cdots + \alpha_n\mathbf{b}_n).
\end{aligned}
$$

Porém, como T é uma transformação linear, pelas propriedades (21) e (22), podemos escrever a equação (27) como

$$
(28)\quad \begin{aligned}
\mathbf{y} &= \beta_1\mathbf{b}_1 + \beta_2\mathbf{b}_2 + \cdots + \beta_n\mathbf{b}_n \\
&= T(\alpha_1\mathbf{b}_1 + \alpha_2\mathbf{b}_2 + \cdots + \alpha_n\mathbf{b}_n) \\
&= \alpha_1T(\mathbf{b}_1) + \alpha_2T(\mathbf{b}_2) + \cdots + \alpha_nT(\mathbf{b}_n).
\end{aligned}
$$

Chamando de **T** a matriz da transformação linear, podemos escrever a equação (28) como

$$
(29)\quad \begin{aligned}
\mathbf{y} &= \mathbf{Tx} \\
\begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix} &= [T(\mathbf{b}_1)|T(\mathbf{b}_2)|\cdots|T(\mathbf{b}_n)] \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix}.
\end{aligned}
$$

Cada coluna $\mathbf{c}_i$ da matriz $\mathbf{T}$ é a saída da caixa preta da transformação linear $T(\,)$ colocando as componentes do vetor base $\mathbf{b}_i$ como entrada.

No exemplo da equação (23), assumindo que $\mathbf{b}$ é a base canônica, temos que $\mathbf{b}_1 = (1, 0, 0)$, $\mathbf{b}_2 = (0, 1, 0)$ e $\mathbf{b}_3 = (0, 0, 1)$. Assim

(30) $$(\mathbf{c_1})_\mathbf{b} = \begin{pmatrix} 3 \times 1 + 2 \times 0 + 0 \\ 1 - 2 \times 0 \\ 4 \times 0 - 2 \times 0 \end{pmatrix}_\mathbf{b} = \begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix}_\mathbf{b}$$

(31) $$(\mathbf{c_1})_\mathbf{b} = \begin{pmatrix} 3 \times 0 + 2 \times 1 + 0 \\ 0 - 2 \times 1 \\ 4 \times 1 - 2 \times 0 \end{pmatrix}_\mathbf{b} = \begin{pmatrix} 2 \\ -2 \\ 4 \end{pmatrix}_\mathbf{b}$$

(32) $$(\mathbf{c_1})_\mathbf{b} = \begin{pmatrix} 3 \times 0 + 2 \times 0 + 1 \\ 0 - 2 \times 0 \\ 4 \times 0 - 2 \times 1 \end{pmatrix}_\mathbf{b} = \begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix}_\mathbf{b}.$$

Então, a matriz $\mathbf{T}$ é dada por

(33) $$\mathbf{T} = \begin{bmatrix} 3 & 2 & 1 \\ 1 & -2 & 0 \\ 0 & 4 & -2 \end{bmatrix}.$$

É possível representar a matriz de transformação da caixa preta para que o vetor de entrada esteja na base $\mathbf{b}$ e o vetor de saída esteja na base $\mathbf{r}$. Vamos mostrar como fazer isso com um exemplo. Suponha que a matriz $\mathbf{T}$ seja uma matriz de rotação de 30 graus em torno do eixo z dada por

(34) $$\mathbf{T} = \begin{bmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 \\ \frac{1}{2} & \frac{\sqrt{3}}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Se o vetor de entrada for

(35) $$(\mathbf{x})_\mathbf{b} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}_\mathbf{b},$$

o vetor de saída é

(36) $$(\mathbf{y})_\mathbf{b} = \begin{bmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 \\ \frac{1}{2} & \frac{\sqrt{3}}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}_\mathbf{b} = \begin{pmatrix} \frac{\sqrt{3}}{2} - \frac{1}{2} \\ \frac{\sqrt{3}}{2} + \frac{1}{2} \\ 0 \end{pmatrix}_\mathbf{b}.$$

Mudança de coordenadas também é uma transformação linear. Assim na Figura 1, por exemplo, a matriz de transformação da base $\mathbf{b} \equiv \mathbf{e}$ para a base $\mathbf{r}$ é obtida pela inversa da matriz que transforma da base $\mathbf{r}$ para a base $\mathbf{b}$. Portanto,

(37) $$(\mathbf{v})_\mathbf{b} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} (\mathbf{v})_\mathbf{r}$$

e o mapeamento inverso

$$(38)\qquad (\mathbf{v})_{\mathbf{r}} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}^{-1} (\mathbf{v})_{\mathbf{b}} = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} (\mathbf{v})_{\mathbf{b}}$$

Multiplicando a matriz da equação (38) nos dois lados da equação (36) obtemos

$$\begin{aligned} (\mathbf{y})_{\mathbf{r}} &= \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} (\mathbf{y})_{\mathbf{b}} \\ (39)\qquad &= \begin{bmatrix} -\frac{1}{2}+\frac{\sqrt{3}}{2} & -\frac{1}{2}-\frac{\sqrt{3}}{2} & 0 \\ \frac{1}{2} & \frac{\sqrt{3}}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}_{\mathbf{b}} \\ &= \begin{pmatrix} -1 \\ \frac{1}{2}+\frac{\sqrt{3}}{2} \\ 0 \end{pmatrix}_{\mathbf{r}}. \end{aligned}$$

Finalmente, a matriz de transformação incluindo mudança de base pode ser escrita como uma generalização da equação (29) como

$$\begin{aligned} \mathbf{y}_{\boldsymbol{r}} &= \overline{\mathbf{T}}\mathbf{x}_{\boldsymbol{b}} \\ (40)\qquad \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix}_{\mathbf{r}} &= [\boldsymbol{M}_{\mathbf{b}\to\mathbf{r}}][T(\mathbf{b}_1)|T(\mathbf{b}_2)|\cdots|T(\mathbf{b}_n)] \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix}_{\mathbf{b}}. \end{aligned}$$

3. Autovalores e Autovetores de Matrizes Simétricas com Coeficientes Reais

Uma matriz simétrica $\boldsymbol{A}$ é aquela que é igual à sua transposta, ou seja,

$$(41)\qquad \boldsymbol{A} = \boldsymbol{A}^{\mathrm{T}}.$$

Assim, uma matriz simétrica tem que ter o número de linhas iguais ao número de colunas (matriz quadrada).

Como a matriz $\boldsymbol{A}$ representa uma transformação linear de $\mathbb{R}^n$ para $\mathbb{R}^n$, podemos definir três tipos de problemas envolvendo essa transformação:

1) Multiplicação de matriz por vetor

$$(42)\qquad \mathbf{y} = \mathbf{A}\mathbf{x}$$

onde $\mathbf{A}$ e $\mathbf{x}$ são dados e temos que obter $\mathbf{y}$. Esse problema é trivial e foi utilizado na disciplina de Fundamentos de Programação para ilustrar o conceito de loop aninhado.

2) Solução de sistemas de equações algébricas lineares

$$(43)\qquad \mathbf{y} = \mathbf{A}\mathbf{x}$$

onde $\mathbf{A}$ e $\mathbf{y}$ são dados e temos que obter $\mathbf{x}$. Esse problema foi discutido na disciplina de Métodos Numéricos I.

3) Problema de Autovalores e Autovetores

$$(44)\qquad \mathbf{y} = \mathbf{A}\mathbf{x}$$

onde $\mathbf{A}$ é dada e temos de obter $\mathbf{x}$ tal que o vetor $\mathbf{y}$ seja o próprio vetor $\mathbf{x}$ multiplicado por um escalar λ, isto é

(45) $\mathbf{Ax} = \lambda\mathbf{x}$.

Antes de discutir os métodos numéricos para resolver esse problema para matrizes grandes, vamos discutir conceitos preliminares, usando uma matriz não simétrica que transforma de $\mathbb{R}^3$ para $\mathbb{R}^3$. (https://www.youtube.com/watch?v=gJhlkEBZsfI)

3.1 Discussões preliminares

Considere o seguinte problema de autovalores e autovetores

(46) $\begin{bmatrix} 4 & 2 & 1 \\ 1 & 6 & 2 \\ 0 & 1 & 5 \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \lambda \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$

O vetor **x** e o escalar λ que satisfazem a equação (46) ou, de forma mais geral, a equação (45), são chamados, respectivamente de autovetor e autovalor.

Vamos trabalhar simultaneamente com as equações (45) (mais geral) e (46) (caso particular), e trazer o termo do lado direito para o lado esquerdo da equação. Assim,

(47) $\mathbf{Ax} - \lambda\mathbf{x} = \mathbf{0}$.

(48) $\begin{bmatrix} 4 & 2 & 1 \\ 1 & 6 & 2 \\ 0 & 1 & 5 \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} - \lambda \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$

Se tentássemos, nessa forma, colocar o vetor **x** em evidência, obteríamos uma operação inválida de subtração da matriz **A** menos o escalar λ (isso não é possível). Porém, podemos reescrever as equações (47) e (48) de modo a permitir que o vetor **x** seja posto em evidência, isto é

(49) $\mathbf{Ax} - \lambda\mathbf{Ix} = \mathbf{0} \rightarrow [\mathbf{A} - \lambda\mathbf{I}]\mathbf{x} = \mathbf{0}$.

(50) $\begin{bmatrix} 4 & 2 & 1 \\ 1 & 6 & 2 \\ 0 & 1 & 5 \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} - \lambda \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \rightarrow \left[\begin{bmatrix} 4 & 2 & 1 \\ 1 & 6 & 2 \\ 0 & 1 & 5 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\right] \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$

A matriz $[\mathbf{A} - \lambda\mathbf{I}]$ é obtida a partir da matriz **A**, subtraindo o escalar λ de todos os elementos da diagonal da matriz **A**. Assim, a equação (50) pode ser reescrita como

(51) $\begin{bmatrix} 4-\lambda & 2 & 1 \\ 1 & 6-\lambda & 2 \\ 0 & 1 & 5-\lambda \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$

As equações (49) e (51) representam um sistema de equações algébricas linear homogêneo. Neste caso, se o determinante da matriz do sistema for diferente de zero só existe uma solução para o sistema. Porém o vetor $\mathbf{x} = \mathbf{0}$ sempre é uma solução desse sistema. Para que existam outras possibilidades de solução, o determinante tem de ser igual a zero. Assim, vamos escrever

(52) $\det([\mathbf{A} - \lambda\mathbf{I}]) = 0 \rightarrow p(\lambda) = \det([\mathbf{A} - \lambda\mathbf{I}]) = 0$

(53) $\det\left(\begin{bmatrix} 4-\lambda & 2 & 1 \\ 1 & 6-\lambda & 2 \\ 0 & 1 & 5-\lambda \end{bmatrix}\right) = 0 \rightarrow -\lambda^3 + 15\lambda^2 - 103 = 0$

Note que essa condição de determinante da matriz do sistema homogêneo ser igual a zero leva a uma equação chamada "**Equação Característica**" que, ao ser resolvida, nos **fornece os autovalores do problema**. Porém, se a matriz for grande, por exemplo 10×10, o polinômio seria de grau 10 e teria 10 raízes (os autovalores). No entanto, achar a equação característica para matrizes com dimensões maiores do que 3 não é prático. A importância dessa discussão é apenas teórica, pois nos permite saber que uma matriz de dimensão $n \times n$ tem uma Equação Característica de grau n e podemos esperar encontrar n autovalores. Para o caso da equação (53), a solução nos dá os seguintes autovalores

$$(54)\quad \begin{cases} \lambda_1 & = & 7.49086 \\ \lambda_2 & = & 4.34338 \\ \lambda_3 & = & 3.16576. \end{cases}$$

Resolvendo o sistema homogêneo da equação (51) para cada um dos três valores de λ, nos dará três pares $(\lambda_i, \mathbf{x}_i), i = 1,2,3$ autovalor-autovetor correspondente, isto é

$$(55)\quad) \begin{cases} (\lambda_1, \mathbf{x}_1) & = & \big(7.49086, (\;1.71354 \quad 2.49086 \quad 1)\big) \\ (\lambda_2, \mathbf{x}_2) & = & \big(4.34338, (-0.91222 \quad -0.65662 \quad 1)\big) \\ (\lambda_3, \mathbf{x}_3) & = & \big(3.16576, (\;3.19869 \quad -1.83424 \quad 1)\big). \end{cases}$$

Observações.

1) Nos casos gerais, as raízes do polinômio podem ser números complexos, podem ser repetidas (multiplicidade algébrica) e podem ser nulas.
2) O número de autovalores não nulos determina o posto da matriz.
3) O número de autovetores pode não ser igual ao número de autovalores. Quando os autovalores são repetidos (multiplicidade algébrica) nem sempre existem autovetores em igual número (multiplicidade geométrica).
4) Quando as matrizes são simétricas e seus elementos são números reais, todos os autovalores são números reais e os autovetores correspondentes a autovalores distintos são perpendiculares entre si. No caso de autovalores repetidos, a multiplicidade algébrica é igual à multiplicidade geométrica. Assim, uma matriz simétrica $n \times n$ tem n autovalores e n autovetores linearmente independentes.
5) Como o sistema é homogêneo, existem variáveis livres e variáveis dependentes. Assim, um autovetor multiplicado por um fator de escala é o mesmo autovetor. O que importa é a direção.