

1

Semana 13

1.1 Matrizes simétricas

Uma **matriz simétrica** é uma matriz que é igual à sua transposta. Para que isto faça sentido, podemos apenas considerar matrizes quadradas. De forma mais precisa, se $A = [a_{ij}]$ 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 $i, j \in \{1, 2, \dots, n\}$, temos $a_{ij} = a_{ji}$;
- (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 1. As matrizes

$$\begin{bmatrix} 1 & -3 \\ -3 & 2 \end{bmatrix}, \begin{bmatrix} 1 & -3 \\ -3 & 1 \end{bmatrix}, \begin{bmatrix} 1 & -3 & 0 \\ -3 & -1 & 4 \\ 0 & 4 & 6 \end{bmatrix} \checkmark$$

são simétricas, enquanto as matrizes

$$\begin{bmatrix} 1 & -3 \\ 3 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 3 \\ -3 & 2 \end{bmatrix}, \begin{bmatrix} 1 & -3 & 0 \\ 4 & -1 & -3 \\ 0 & 4 & 1 \end{bmatrix} \times$$

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 autovalores 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;
- (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 2. 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

$$A + 4I = \begin{bmatrix} 5 & -5 \\ -5 & 5 \end{bmatrix} \sim \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix} \Rightarrow \begin{cases} v_1 = v_2 \\ \text{uma variável livre} \end{cases} \Rightarrow \vec{v}_1 = v_2 \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Associado com 6:

$$A - 6I = \begin{bmatrix} -5 & -5 \\ -5 & -5 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \Rightarrow \begin{cases} v_2 = -v_1 \\ \text{uma variável livre} \end{cases} \Rightarrow \vec{v}_2 = v_2 \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

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:

$$\mathbb{R}^2 = \text{Span} \left\{ \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}, \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix} \right\}.$$

Lembrando do nosso processo de diagonalização, sabemos que

$$P = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix} \Rightarrow P^{-1}AP = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 1 & -5 \\ -5 & 1 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix} = \begin{bmatrix} -4 & 0 \\ 0 & 6 \end{bmatrix}.$$

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). <

Exemplo 3. 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

$$\begin{aligned} p(\lambda) &= \det(A - \lambda I) = \det \begin{bmatrix} 3-\lambda & -2 & 4 \\ -2 & 6-\lambda & 2 \\ 4 & 2 & 3-\lambda \end{bmatrix} \\ &= (3-\lambda)[(6-\lambda)(3-\lambda) - 4] + 2[-6 + 2\lambda - 8] + 4[-4 - 24 + 4\lambda] \\ &= (3-\lambda)[\lambda^2 - 9\lambda + 14] - 28 + 4\lambda - 112 + 16\lambda \\ &= 3\lambda^2 - 27\lambda + 42 - [\lambda^3 - 9\lambda^2 + 14\lambda] + 20\lambda - 140 \\ &= -\lambda^3 + 12\lambda^2 - 21\lambda - 98. \end{aligned}$$

Como sabemos, dificilmente conseguiríamos encontrar as raízes deste polinômio com precisão. De qualquer forma, raízes racionais (este exemplo foi construído 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 :

$$A + 2I = \begin{bmatrix} 5 & -2 & 4 \\ -2 & 8 & 2 \\ 4 & 2 & 5 \end{bmatrix} \sim \begin{bmatrix} -1 & 4 & 1 \\ 5 & -2 & 4 \\ 4 & 2 & 5 \end{bmatrix} \sim \begin{bmatrix} -1 & 4 & 1 \\ 0 & 18 & 9 \\ 0 & 18 & 9 \end{bmatrix} \sim \begin{bmatrix} -1 & 2 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 0 \end{bmatrix} \rightsquigarrow \begin{cases} v_1 = 2v_2 \\ v_3 = -2v_2 \\ 1 \text{ variável livre} \end{cases}$$

Assim, um autovetor associado é ($v_2 = 1$)

$$\vec{v}_1 = \begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix}.$$

- Autoespaço associado com o autovalor 7 :

$$A - 7I = \begin{bmatrix} -4 & -2 & 4 \\ -2 & -1 & 2 \\ 4 & 2 & -4 \end{bmatrix} \sim \begin{bmatrix} -2 & -1 & 2 \\ 2 & 1 & -2 \\ 0 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} -2 & -1 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \rightsquigarrow \begin{cases} v_2 = -2v_1 + 2v_3 \\ 2 \text{ variáveis livres} \end{cases}$$

Assim, qualquer autovetor pode ser escrito parametricamente como

$$\vec{v} = \begin{bmatrix} v_1 \\ -2v_1 + 2v_3 \\ v_3 \end{bmatrix} = v_1 \begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix} + v_3 \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix}.$$

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

$$\left\{ \begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix}, \begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix}, \begin{bmatrix} 4 \\ 2 \\ 5 \end{bmatrix} \right\}.$$

¹Enfatizamos que isto não é verdade para qualquer matriz, apenas para as simétricas!

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, concluímos 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 (matriz simétrica), pode ser obtida ortogonal. Sendo P ortogonal, tem-se $P^{-1} = P^T$ e portanto $P^{-1}AP = D$, 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 $\text{Nul}(A - \lambda I)$. 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 .

1.1.1 Formas quadráticas

Nesta seção, apresentamos uma aplicação de matrizes simétricas. Uma **forma quadrática** em \mathbb{R}^n é uma função polinomial em que todos os termos são de grau 2.

Exemplo 4. A função $Q : \mathbb{R}^2 \rightarrow \mathbb{R}$, definida por

$$Q(x_1, x_2) = 3x_1^2 + x_1x_2$$

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

$$Q(x_1, x_2) = ax_1^2 + 2bx_1x_2 + cx_2^2,$$

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 (1.1) abaixo (observe que cada termo $x_i x_j$ aparece duas vezes na soma).

Exemplo 5. A função $Q : \mathbb{R}^3 \rightarrow \mathbb{R}$, definida por

$$Q(x_1, x_2, x_3) = x_2^2 - x_3^2 + 3x_1x_2 - 8x_1x_3$$

é uma forma quadrática em \mathbb{R}^3 . Qualquer forma quadrática em \mathbb{R}^3 pode ser escrita como

$$Q(x_1, x_2, x_3) = ax_1^2 + bx_2^2 + cx_3^2 + 2dx_1x_2 + 2ex_1x_3 + 2fx_2x_3,$$

onde $a, b, c, d, e, f \in \mathbb{R}$.

De forma geral, uma forma quadrática $Q : \mathbb{R}^n \rightarrow \mathbb{R}$ pode ser escrita como

$$Q(x_1, x_2, \dots, x_n) = \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j, \quad (1.1)$$

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

$$Q(x_1, x_2, \dots, x_n) = \sum_{i=1}^n \sum_{j=1}^n a_{ij}x_i x_j = \sum_{i=1}^n (A\vec{x})_i x_i = (A\vec{x}) \cdot \vec{x}.$$

E aí está a relação entre as formas quadráticas e matrizes simétricas. Acabamos de verificar que:

Qualquer forma quadrática em \mathbb{R}^n pode ser escrita na forma

$$Q(x_1, x_2, \dots, x_n) = A\vec{x} \cdot \vec{x},$$

onde A é uma matriz simétrica de ordem $n \times n$.

Uma outra forma de escrever a mesma coisa é

$$Q(x_1, x_2, \dots, x_n) = A\vec{x} \cdot \vec{x} = \vec{x}^T A \vec{x}.$$

Exemplo 6. Voltamos a analisar os Exemplos 4 e 5 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;

²Caso $Q(x_1, x_2, \dots, x_n) = \sum_{i=1}^n \sum_{j=1}^n b_{ij}x_i x_j$ com uma matriz $B = (b_{ij})$ não simétrica, podemos escrever

$$Q(x_1, x_2, \dots, x_n) = \sum_{i=1}^n b_{ii}x_i^2 + \sum_{i < j} (b_{ij} + b_{ji})x_i x_j. \text{ Definindo } a_{ij} = \frac{b_{ij} + b_{ji}}{2} \text{ temos também } Q(x_1, x_2, \dots, x_n) =$$

$\sum_{i=1}^n a_{ii}x_i^2 + 2 \sum_{i < j} a_{ij}x_i x_j$. 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 $\sum_{i=1}^n \sum_{j=1}^n b_{ij}x_i x_j =$

$$\sum_{i=1}^n a_{ii}x_i^2 + 2 \sum_{i < j} a_{ij}x_i x_j = \sum_{i=1}^n \sum_{j=1}^n a_{ij}x_i x_j.$$

- o coeficiente do termo misto x_1x_2 aparece (dividido por 2) nas posições a_{12} e a_{21} da matriz A ;
- o coeficiente do termo misto x_1x_3 aparece (dividido por 2) nas posições a_{13} e a_{31} da matriz A ;
- o coeficiente do termo misto x_2x_3 aparece (dividido por 2) nas posições a_{23} e a_{32} da matriz A .

Faça as contas para verificar a igualdade acima! \triangleleft

1.1.2 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 dos autovetores associados é a mais natural de todas. Suponhamos que

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

e seja $\mathcal{B} = \{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n\}$ uma base de \mathbb{R}^n formada por autovetores da matriz A . Desta forma,

$$P = \begin{bmatrix} | & | & \cdots & | \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_n \\ | & | & \cdots & | \end{bmatrix} \text{ satisfaz } P^{-1}AP = D,$$

onde D é a matriz diagonal, com os respectivos autovalores na diagonal principal. Além disso, podemos encontrar uma matriz P com estas propriedades e que é ortogonal: $P^T = P^{-1}$. O vetor \vec{x} pode ser representado na base \mathcal{B} :

$$\vec{x} = y_1\vec{v}_1 + y_2\vec{v}_2 + \cdots + y_n\vec{v}_n = P\vec{y}, \quad \text{onde denotamos } \vec{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = [\vec{x}]_{\mathcal{B}}.$$

Agora, vamos escrever a forma quadrática na base \mathcal{B} , isto é, em termos do vetor \vec{y} (lembre que $\vec{x} = P\vec{y} \implies \vec{x}^T = \vec{y}^T P^T$):

$$\vec{x}^T A \vec{x} = (\vec{y}^T P^T) A (P\vec{y}) = \vec{y}^T (P^T A P) \vec{y} = \vec{y}^T D \vec{y}.$$

Isto significa que, ao mudar da base canônica (componentes x_i) para a base \mathcal{B} (componentes y_i), temos

$$\sum_{i,j=1}^n b_{ij} x_i x_j = \vec{x}^T A \vec{x} = \vec{y}^T D \vec{y} = \sum_{i=1}^n \lambda_i y_i^2,$$

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 7. Considere a forma quadrática

$$Q(x_1, x_2) = 6x_1^2 - 2x_1x_2 + 4x_2^2 \iff Q(\vec{x}) = \vec{x}^T A \vec{x}, \quad \text{onde } A = \begin{bmatrix} 6 & -2 \\ -2 & 3 \end{bmatrix}.$$

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 $\lambda_1 = 2$:

$$A - 2I = \begin{bmatrix} 4 & -2 \\ -2 & 1 \end{bmatrix} \sim \begin{bmatrix} 2 & -1 \\ 0 & 0 \end{bmatrix} \Rightarrow v_2 = 2v_1 \rightsquigarrow \begin{bmatrix} 1 \\ 2 \end{bmatrix} \rightsquigarrow \vec{v}_1 = \begin{bmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix}.$$

- Autovetor unitário associado ao autovalor $\lambda_2 = 7$:

$$A - 2I = \begin{bmatrix} -1 & -2 \\ -2 & -4 \end{bmatrix} \sim \begin{bmatrix} -1 & -2 \\ 0 & 0 \end{bmatrix} \Rightarrow v_1 = -2v_2 \rightsquigarrow \begin{bmatrix} -2 \\ 1 \end{bmatrix} \rightsquigarrow \vec{v}_1 = \begin{bmatrix} -2/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix}.$$

Portanto, encontramos uma base ortonormal (de modo que P abaixo é matriz ortogonal e satisfaz $P = P^T$) e podemos escrever

$$P = \begin{bmatrix} 1/\sqrt{5} & -2/\sqrt{5} \\ 2/\sqrt{5} & 1/\sqrt{5} \end{bmatrix} \Rightarrow P^T A P = D = \begin{bmatrix} 2 & 0 \\ 0 & 7 \end{bmatrix}$$

Assim,

$$Q(x_1, x_2) = Q(\vec{x}) = Q(P\vec{y}) = \vec{x}^T A \vec{x} = \vec{y}^T A \vec{y} = 2y_1^2 + 7y_2^2. \triangleleft$$

1.1.3 Aplicação à classificação de cônicas

Nesta subseção, nos restringiremos à 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

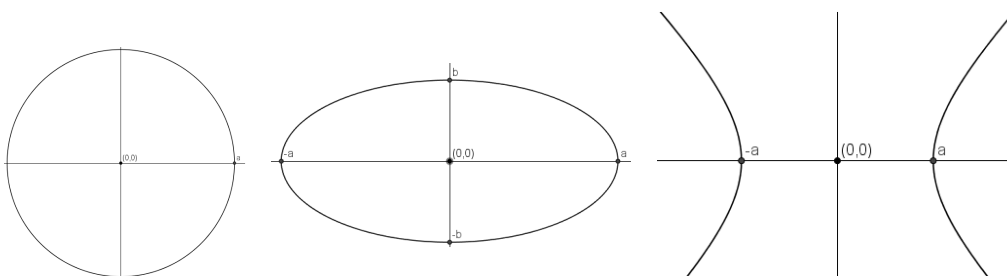
$$Q(x_1, x_2) = k \iff ax_1^2 + bx_1x_2 + cx_2^2 = k,$$

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 ;
- $\frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} = 1$ representa uma elipse;
- $\frac{x_1^2}{a^2} - \frac{x_2^2}{b^2} = 1$ 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 8. Vamos identificar a cônica cuja equação é

$$31x_1^2 - 10\sqrt{3}x_1x_2 + 21x_2^2 = 144.$$

Podemos escrever

$$\vec{x}^T A \vec{x} = 144, \quad \text{onde} \quad A = \begin{bmatrix} 31 & -5\sqrt{3} \\ -5\sqrt{3} & 21 \end{bmatrix}.$$

Autovalores de A :

$$\begin{aligned} p(\lambda) &= (31 - \lambda)(21 - \lambda) - 75 = 0 \implies \lambda^2 - 52\lambda + 576 = 0 \\ \implies \lambda &= \frac{52 \pm \sqrt{2704 - 2304}}{2} = \frac{52 \pm 20}{2} = 26 \pm 10 \implies \lambda_1 = 16, \lambda_2 = 36. \end{aligned}$$

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

$$P^{-1}AP = D = \begin{bmatrix} 16 & 0 \\ 0 & 26 \end{bmatrix}$$

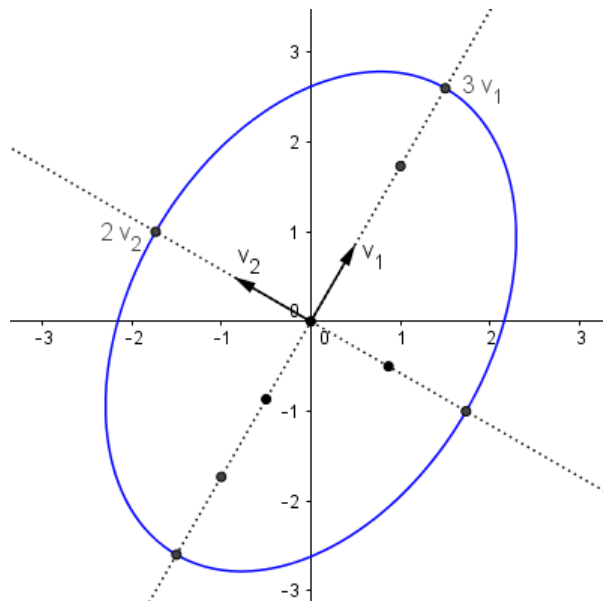
Assim, na base de \mathbb{R}^2 determinada por $\{\vec{v}_1, \vec{v}_2\}$, podemos escrever a equação da equação cônica como

$$\vec{x}^T A \vec{x} = \vec{y}^T D \vec{y} \implies \begin{bmatrix} y_1 & y_2 \end{bmatrix} \begin{bmatrix} 16 & 0 \\ 0 & 26 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = 144.$$

Isto pode ser reescrito como

$$16y_1^2 + 26y_2^2 = 144 \iff \frac{y_1^2}{(144/16)} + \frac{y_2^2}{(144/26)} = 1 \iff \frac{y_1^2}{3^2} + \frac{y_2^2}{2^2} = 1.$$

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 $\{\vec{v}_1, \vec{v}_2\}$

1.2 Problemas de otimização restritos à esfera

Nesta seção, vamos estudar alguns problemas de otimização bem específicos. Mais precisamente, vamos aprender a 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 $(x_1, x_2, \dots, x_n) \in \mathbb{R}^n$ 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 11 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 1. Fazer toda a análise acima para o problema de minimização. Dica: olhar para o menor autovalor.

Exemplo 9. 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

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \implies Q(\vec{x}) = \vec{x}^T \begin{bmatrix} 7 & -4 & -2 \\ -4 & 1 & -4 \\ -2 & -4 & 7 \end{bmatrix} \vec{x}.$$

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

$$A - 9I = \begin{bmatrix} -2 & -4 & -2 \\ -4 & -8 & -4 \\ -2 & -4 & -2 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \Rightarrow \vec{v} = \begin{bmatrix} -2v_2 - v_3 \\ v_2 \\ v_3 \end{bmatrix} = v_2 \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} + v_3 \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}.$$

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

$$Q\left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right) = \frac{7}{2} + 0 + \frac{7}{2} - 0 + \frac{4}{2} - 0 = 9. \triangleleft$$

Observação 10 (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

$$\text{minimizar } f(x_1, x_2, \dots, x_n) \text{ com a restrição } g(x_1, x_2, \dots, x_n) = 0.$$

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 \vec{x} e possíveis valores de λ , podemos testar para ver quais são os de mínimo e quais são os de máximo.

Vejamos como esta técnica se aplicaria no Exemplo 9 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

$$\begin{array}{l} 14x_1 - 8x_2 - 4x_3 = 2\lambda x_1 \\ -8x_1 + 2x_2 - 8x_3 = 2\lambda x_2 \\ -4x_1 - 8x_2 + 14x_3 = 2\lambda x_3 \end{array} \xrightarrow[\text{em cada eq.}]{\text{Simplifica um 2}} \begin{bmatrix} 7 & -4 & -2 \\ -4 & 1 & -4 \\ -2 & -4 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \lambda \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}.$$

Chegamos, por outro caminho, à mesma conclusão: λ deve ser um autovalor e \vec{x} um autovetor (unitário) associado. \triangleleft

1.3 Matrizes simétricas e demonstração 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} (quando um deles está sendo multiplicado por uma matriz).

Observação 11. 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 seria (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 12. 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 11 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 13. 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

$$\langle k\vec{u}, \vec{v} \rangle = k \langle \vec{u}, \vec{v} \rangle, \quad \text{enquanto que} \quad \langle \vec{u}, k\vec{v} \rangle = \bar{k} \langle \vec{u}, \vec{v} \rangle.$$

□

Proposição 14. 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!)

$$\lambda_1 \langle \vec{v}_1, \vec{v}_2 \rangle = \langle \lambda_1 \vec{v}_1, \vec{v}_2 \rangle = \langle A\vec{v}_1, \vec{v}_2 \rangle = \langle \vec{v}_1, A\vec{v}_2 \rangle = \bar{\lambda}_2 \langle \vec{v}_1, \vec{v}_2 \rangle = \lambda_2 \langle \vec{v}_1, \vec{v}_2 \rangle,$$

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, $\lambda_1 \neq \lambda_2$, concluímos que $\langle \vec{v}_1, \vec{v}_2 \rangle = 0$. \square

Proposição 15. Se A é 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

$$\{\vec{v}_1, \vec{y}_2, \vec{y}_3, \dots, \vec{y}_n\}$$

é uma base ortonormal de \mathbb{R}^n .

Considere a matriz Q cujas colunas são os elementos desta base:

$$Q = \begin{bmatrix} | & | & | & \cdots & | \\ \vec{v}_1 & \vec{y}_2 & \vec{y}_3 & \cdots & \vec{y}_n \\ | & | & | & \cdots & | \end{bmatrix},$$

que é, consequentemente, ortogonal. Vamos calcular $Q^{-1}AQ = Q^T AQ$. O produto AQ pode ser interpretado como o produto de A por cada uma das colunas de Q :

$$AQ = \begin{bmatrix} | & | & | & \cdots & | \\ A\vec{v}_1 & A\vec{y}_2 & A\vec{y}_3 & \cdots & A\vec{y}_n \\ | & | & | & \cdots & | \end{bmatrix} = \begin{bmatrix} | & | & | & \cdots & | \\ \lambda_1 \vec{v}_1 & A\vec{y}_2 & A\vec{y}_3 & \cdots & A\vec{y}_n \\ | & | & | & \cdots & | \end{bmatrix}.$$

Logo,

$$\begin{aligned} Q^T AQ &= \begin{bmatrix} - & \vec{v}_1 & - \\ - & \vec{y}_2 & - \\ - & \vec{y}_3 & - \\ & \vdots & \\ - & \vec{y}_n & - \end{bmatrix} \begin{bmatrix} | & | & | & \cdots & | \\ \lambda_1 \vec{v}_1 & A\vec{y}_2 & A\vec{y}_3 & \cdots & A\vec{y}_n \\ | & | & | & \cdots & | \end{bmatrix} \\ \Rightarrow Q^T AQ &= \begin{bmatrix} \lambda_1 \langle \vec{v}_1, \vec{v}_1 \rangle & \langle \vec{v}_1, A\vec{y}_2 \rangle & \langle \vec{v}_1, A\vec{y}_3 \rangle & \cdots & \langle \vec{v}_1, A\vec{y}_n \rangle \\ \lambda_1 \langle \vec{y}_2, \vec{v}_1 \rangle & \langle \vec{y}_2, A\vec{y}_2 \rangle & \langle \vec{y}_2, A\vec{y}_3 \rangle & \cdots & \langle \vec{y}_2, A\vec{y}_n \rangle \\ \lambda_1 \langle \vec{y}_3, \vec{v}_1 \rangle & \langle \vec{y}_3, A\vec{y}_2 \rangle & \langle \vec{y}_3, A\vec{y}_3 \rangle & \cdots & \langle \vec{y}_3, A\vec{y}_n \rangle \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \lambda_1 \langle \vec{y}_n, \vec{v}_1 \rangle & \langle \vec{y}_n, A\vec{y}_2 \rangle & \langle \vec{y}_n, A\vec{y}_3 \rangle & \cdots & \langle \vec{y}_n, A\vec{y}_n \rangle \end{bmatrix} \end{aligned}$$

Observamos que, pela simetria de A e por ser $\{\vec{v}_1, \vec{y}_2, \vec{y}_3, \dots, \vec{y}_n\}$ 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,

$$\Rightarrow Q^T AQ = \begin{bmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & b_{22} & b_{23} & \cdots & b_{2n} \\ 0 & b_{32} & b_{33} & \cdots & b_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & b_{n2} & b_{n3} & \cdots & b_{nn} \end{bmatrix} \stackrel{\text{def}}{=} \hat{A}.$$

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

$$\det(A - \lambda I) = \det(Q^{-1}(A - \lambda I)Q) = \det(\hat{A} - \lambda I) = (\lambda_1 - \lambda) \cdot \det(B - \lambda I),$$

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, A possui um autovetor unitário \vec{v}_2 associado com λ_1 que está contido em $\text{Span}\{\vec{y}_2, \vec{y}_3, \dots, \vec{y}_n\}$ e, portanto, é distinto de (e ortogonal a) \vec{v}_1 .

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 maneira até encontrar k autovetores ortogonais e unitários associados ao autovalor λ_1 . Assim, o autoespaço $\text{Nul}(A - \lambda_1 I)$ tem dimensão k , que é igual à multiplicidade do autovalor λ_1 . \square

1.3.1 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 13 garante que todas estes n autovalores são 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 $m_1 + m_2 + \dots + m_k = n$.

Pela Proposição 15, a dimensão do autoespaço associado a λ_1 é m_1 . Podemos encontrar (da maneira usual, por escalonamento) uma base de autovetores

$$\text{Nul}(A - \lambda_1 I) = \text{Span}\{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_{m_1}\}.$$

Estes autovetores, possivelmente não são ortogonais, mas, aplicando o Processo de Gram-Schmidt, podemos obter uma base ortonormal de $\text{Nul}(A - \lambda_1 I)$:

$$\text{Nul}(A - \lambda_1 I) = \text{Span}\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_{m_1}\}.$$

Fazendo este mesmo procedimento para cada um dos autovalores, obtemos n autovetores, que, conseqüentemente, formam uma base para \mathbb{R}^n . Logo, A é diagonalizável.

Além disso, a Proposição 14 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

$$P = \begin{bmatrix} | & | & \cdots & | \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_n \\ | & | & \cdots & | \end{bmatrix}$$

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 $P^{-1} = P^T$.