Capítulo 10

Cônicas, Matrizes Simétricas e Formas Quadráticas

Nesse capítulo vamos exibir dois métodos de mudança de coordenadas. No primeiro método mostraremos um resultado mais forte, o qual nos fala que: para toda equação do 2º grau, existe um sistema de coordenadas ortonormal no qual a equação se escreve como uma das equações padrões das cônicas. Para demonstrar esse resultado, introduziremos um tipo especial de operador linear, conhecido como operador autoadjunto, e mostraremos que todo operador desse tipo possui uma base ortonormal de autovetores.

O segundo método dar-nos-á um algoritmo que permite reduzir qualquer forma quadrática a uma forma diagonal, consistindo em uma técnica útil para cálculos em espaços de dimensão superior. Além disso, essa técnica é usada para classificar as formas bilineares simétricas. Ela tem o inconveniente, porém, de gerar um sistema de coordenadas final que não é ortogonal.

Ao final do capítulo aplicaremos tais técnicas para classificar as quadráticas em duas variáveis. Deve ficar claro que esse processo pode ser estendido para equações de segundo grau em mais de duas variáveis.

Além disso, neste capítulo, nós iremos introduzir três conceitos que estão relacionados: matrizes simétricas, formas bilineares simétricas e formas quadráticas. No texto, apresentaremos as conexões entre eles, mas não nos aprofundaremos no estudo das formas quadráticas e na classificação das formas bilineares. Gostaríamos apenas de chamar a atenção para o fato de que esses dois conceitos oferecem inúmeras aplicações em questões de otimização e de tratamento de dados.

10.1 Cônicas

As cônicas, ou secções cônicas, foram estudadas pelos gregos e desenvolvidas a partir de suas propriedades geométricas. Pouco se sabe a respeito dos precursores do estudo das cônicas. Restaram apenas os trabalhos de Apollonius de Perga (260-170 A. C.). Nesses trabalhos, não só foram compilados os resultados conhecidos na época como também apresentada de forma sistemática a dedução das propriedades das cônicas.

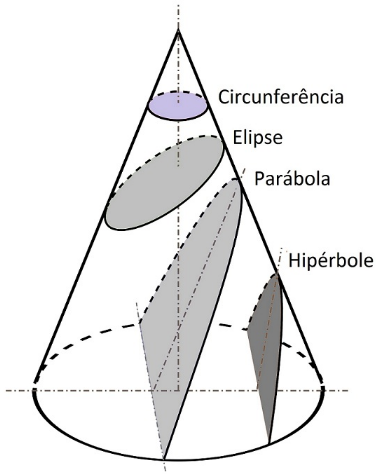


Figura 10.1: Seções de um Cone

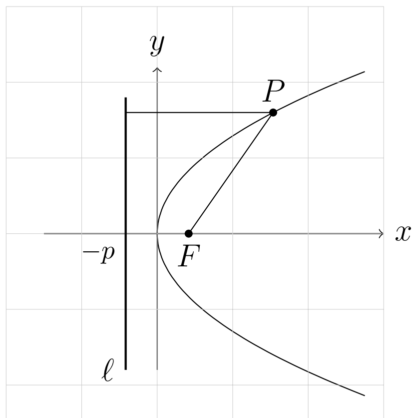
Abaixo daremos a definição das cônicas proposta por Apollonius.

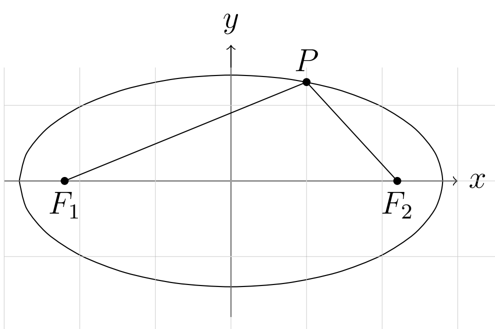
Definição 10.1

A **parábola** é obtida por considerar os pontos P do plano que satisfazem

$$\text{dist}(F, P) = \text{dist}(P, \ell),$$

onde ℓ é a **reta diretriz**, e sua equação é $x = -p$.





A **elipse** é obtida por considerar os pontos P do plano que satisfazem

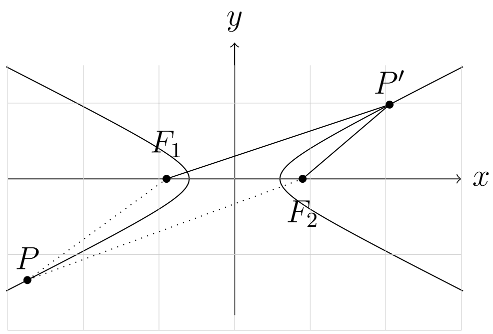
$$\text{dist}(P, F_1) + \text{dist}(P, F_2) = 2a. \tag{10.1}$$

A **hipérbole** é obtida por considerar os pontos P, P' do plano que satisfazem

$$\operatorname{dist}(P', F_1) - \operatorname{dist}(P', F_2) = 2a. \tag{10.2}$$

ou

$$\operatorname{dist}(P, F_1) - \operatorname{dist}(P, F_2) = -2a. \tag{10.3}$$



É possível, a partir das definições acima, obter as seguintes equações padrões:

$$y^2 = 4px, \quad \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \text{ ou } \frac{x^2}{a^2} - \frac{y^2}{b^2} = 1,$$

as quais descrevem uma parábola, uma elipse, ou uma hipérbole, respectivamente.

Para exemplificar como é feita a passagem da definição acima até a equação, vamos tratar de forma completa somente o caso da parábola, que, das três, é a mais simples de deduzir.

10.1.1 Equação da Parábola

Considere $a > 0$. Defina $p = a = c$ e introduza um sistema de coordenada xy tal que o eixo do x passe por cima dos focos e esse tenha coordenadas $F = \left[\begin{smallmatrix} p \\ 0 \end{smallmatrix}\right]$, e o eixo do y seja paralela à reta diretriz ℓ e satisfaça a equação $x = -p$. Da definição da parábola

$$\operatorname{dist}(F, P) = \operatorname{dist}(P, \ell),$$

se $P = \left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right]$, segue que $\operatorname{dist}(F, P) = \sqrt{(x-p)^2 + y^2}$ e $\operatorname{dist}(P, \ell) = |x + p|$. A equação fica

$$\sqrt{(x-p)^2 + y^2} = |x + p|.$$

Elevando ao quadrado ambos o membros dessa equação, e simplificando, obtemos:

$$y^2 = 4px.$$

10.1.2 Equação da Elipse

Considere $a > b > 0$. Defina $c = \sqrt{a^2 - b^2}$ que é conhecida como a **distância focal da elipse**. Vamos introduzir um sistema de coordenadas xy, de tal forma que os focos da elipse fiquem sobre o eixo x e tenham coordenadas $F_1 = \left[\begin{smallmatrix} -c \\ 0 \end{smallmatrix}\right]$ e $F_2 = \left[\begin{smallmatrix} c \\ 0 \end{smallmatrix}\right]$, denominados focos à esquerda e à direita, respectivamente. Escolha o eixo do y de tal forma que ele seja a mediatriz do segmento F_1F_2. Com esse sistema de coordenadas, a equação 10.1 torna-se

$$\sqrt{(x+c)^2 + y^2} + \sqrt{(x-c)^2 + y^2} = 2a.$$

Assim, passando o segundo radical para o outro lado da igualdade, elevando ao quadrado, simplificando, e usando a relação $c^2 = a^2 - b^2$, obtemos a equação padrão da elipse
$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1.$$

10.1.3 Equação da Hipérbole

Considere $a > b > 0$. Defina $c = \sqrt{a^2 + b^2}$ que é conhecido como a **distância focal da Hipérbole**. Aqui também introduza um sistema de coordenada xy tal que o eixo do x passe por cima dos focos que tenham coordenadas $F_1 = \left[\begin{smallmatrix} -c \\ 0 \end{smallmatrix}\right]$ e $F_2 = \left[\begin{smallmatrix} c \\ 0 \end{smallmatrix}\right]$, denominados focos à esquerda e à direita. Escolha o eixo de y de modo que ele seja a mediatriz do segmento F_1F_2. Com esse sistema, as equações 10.2 e 10.3 tornam-se, respectivamente,

$$\sqrt{(x+c)^2 + y^2} - \sqrt{(x-c)^2 + y^2} = 2a \text{ e } \sqrt{(x+c)^2 + y^2} - \sqrt{(x-c)^2 + y^2} = -2a.$$

Escolhendo qualquer uma dessas equações, procedendo de forma idêntica ao que foi feito com a equação da elipse, e usando que $c^2 = a^2 + b^2$, obtemos a equação padrão da hipérbole
$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1.$$

10.1.4 Excentricidade e Propriedades das Cônicas

A seguir iremos definir alguns conceitos importantes a respeito das cônicas. Diremos que as diretrizes da elipse e da hipérbole são as retas:

$$y = -\frac{a}{e} \text{ e } x = \frac{a}{e}, \text{ respectivamente. A circunferência não tem diretriz.}$$

Observação 10.2
Se tivermos na elipse ou na hipérbole $b > a > 0$, a distância focal será, respectivamente, $c = \sqrt{b^2 - a^2}$, $c = \sqrt{a^2 + b^2}$. Devemos escolher o sistema de coordenadas de maneira que os focos tenham coordenadas $F_1 = \left[\begin{smallmatrix} 0 \\ -c \end{smallmatrix}\right]$ e $F_2 = \left[\begin{smallmatrix} 0 \\ c \end{smallmatrix}\right]$ e, nesse caso, as diretrizes serão $y = \pm\frac{b}{e}$.

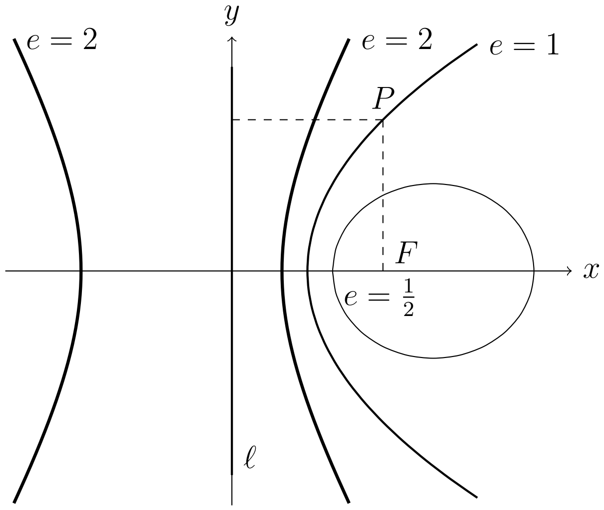
As cônicas possuem um invariante que nos permite defini-las mais sinteticamente e nos oferece uma outra visão sobre a natureza delas. Para esclarecer melhor isso, introduziremos a noção de **excentricidade** e, que é, por definição, $e = \frac{c}{a}$.

Segue da definição de excentricidade que, se a cônica for uma elipse, então, $0 \leq e < 1$ (na circunferência, e $= 0$); e se for uma hipérbole, então $e > 1$, e se for uma parábola, então, $e = 1$.

Proposição 10.3
Fixe um ponto F (Foco), uma linha ℓ (a diretriz), que não contenha F, e escolha um número Real positivo e (a excentricidade). Considere que os pontos P satisfaçam
$$\operatorname{dist}(P, F) = e \operatorname{dist}(P, \ell).$$

Mostre que: a) se $0 < e < 1$, obtemos uma elipse; b) se $e = 1$, obtemos uma parábola; c) se $e > 1$, obtemos uma hipérbole.

Demonstração: veja o exercício R10.1 e figura 10.1.4 que ilustra a propriedade. □



Nos exercícios, exibiremos outras propriedades interessantes das cônicas. No intuito de mostrar que toda equação do 2ª grau pode ser vista como uma cônica em um sistema de coordenadas ortonormal, vamos apresentar uma família de operadores lineares, conhecida como família de operadores autoadjuntos.

10.2 Operadores Autoadjuntos

Seja $\langle , \rangle$ um produto interno em $\mathbb{R}^n$. Diremos que um operador linear $T : \mathbb{R}^n \to \mathbb{R}^n$ será um operador autoadjunto se a matriz $[T]^\alpha_\alpha$, com respeito à uma base ortonormal α, for simétrica.

Suponha que $T : \mathbb{R}^2 \to \mathbb{R}^2$ seja dado por $\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right] \mapsto \left[\begin{smallmatrix} 3x+4y \\ 4x-7y \end{smallmatrix}\right]$. Considerando $\mathbb{R}^2$ com o produto interno usual e α a base canônica do $\mathbb{R}^2$, temos que a matriz

$$[T]^\alpha_\alpha = \begin{bmatrix} 3 & 4 \\ 4 & -7 \end{bmatrix}$$

é simétrica e, portanto, esse operador é um operador autoadjunto.

Teorema 10.4
Sejam $\langle\ ,\ \rangle$ um produto interno em $\mathbb{R}^n$ e $T : \mathbb{R}^n \to \mathbb{R}^n$ um operador linear. Então, T é um operador autoadjunto se, e somente se,

$$\langle T\mathbf{u}, \mathbf{v}\rangle = \langle \mathbf{u}, T\mathbf{v}\rangle\,,\ \textit{para todo } \mathbf{u}, \mathbf{v} \in \mathbb{R}^n.$$

Demonstração: veja o exercício R10.2 □

Vamos verificar tal propriedade com a transformação $T : \mathbb{R}^2 \to \mathbb{R}^2$ definida acima e com o produto interno usual do $\mathbb{R}^2$. Considere $\mathbf{u} = \left[\begin{smallmatrix} x_1 \\ x_2 \end{smallmatrix}\right]$ e $\mathbf{v} = \left[\begin{smallmatrix} y_1 \\ y_2 \end{smallmatrix}\right]$ então,

$$\begin{aligned}
\langle T\mathbf{u}, \mathbf{v} \rangle &= \left\langle \left[\begin{smallmatrix} 3x_1+4x_2 \\ 4x_1-7x_2 \end{smallmatrix}\right], \left[\begin{smallmatrix} y_1 \\ y_2 \end{smallmatrix}\right] \right\rangle \\
&= 3x_1y_1 + 4x_2y_1 + 4x_1y_2 - 7x_2y_2 \\
&= 3x_1y_1 + 4x_1y_2 + 4x_2y_1 - 7x_2y_2 \\
&= x_1(3y_1 + 4y_2) + x_2(4y_1 - 7y_2) \\
&= \left\langle \left[\begin{smallmatrix} x_1 \\ x_2 \end{smallmatrix}\right], \left[\begin{smallmatrix} 3y_1+4y_2 \\ 4y_1-7y_2 \end{smallmatrix}\right] \right\rangle = \langle \mathbf{u}, T\mathbf{v} \rangle .
\end{aligned}$$

Corolário 10.5
Suponha que $T : \mathbb{R}^n \to \mathbb{R}^n$ seja um operador autoadjunto; então a matriz de T será simétrica com respeito a qualquer base ortonormal.

Proposição 10.6
Seja $\langle\ ,\ \rangle$ um produto interno em $\mathbb{R}^n$. Se $T : \mathbb{R}^n \to \mathbb{R}^n$ é um operador autoadjunto então:

a) autovetores correspondentes a autovalores distintos são ortogonais;
b) T sempre possui pelo menos um autovalor.

Demonstração: veja o exercício R10.3 □

Lema 10.7
Seja $T : \mathbb{R}^n \to \mathbb{R}^n$ um operador autoadjunto. Se o subespaço $W \subset \mathbb{R}^n$ é T-invariante, então seu complemento ortogonal $W^\perp$ também o é.

Demonstração: considere $\mathbf{u} \in W$ e $\mathbf{v} \in W^\perp$, então $T\mathbf{u} \in W$ e

$$\langle \mathbf{u}, T\mathbf{v} \rangle = \langle T\mathbf{u}, \mathbf{v} \rangle = 0.$$

Mas isso implica que $T\mathbf{v} \in W^\perp$, logo $W^\perp$ também é T-invariante. □

Teorema 10.8 (Teorema Espectral)
Seja $\langle\ ,\ \rangle$ um produto interno em $\mathbb{R}^n$. Se $T : \mathbb{R}^n \to \mathbb{R}^n$ é um operador autoadjunto, então existe uma base ortogonal de $\mathbb{R}^n$ composta por autovetores de T.

Demonstração: veja o exercício R10.5. □

Exemplo 10.9
Considere $\mathbb{R}^2$ com o produto interno usual e $T\left(\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right]\right) = \left[\begin{smallmatrix} x+3y \\ 3x-7y \end{smallmatrix}\right]$. Seja $\mathcal{C}$ a base canônica do $\mathbb{R}^2$, chame

$$A = [T]_{\mathcal{C}}^{\mathcal{C}} = \begin{bmatrix} 1 & 3 \\ 3 & -7 \end{bmatrix}. \text{ Segue que } \Delta_A(x) = x^2 + 6x - 16 = (x-2)(x+8).$$

Como a matriz A é simétrica, segue que T é autoadjunto e, portanto, pelo teorema 10.8 (Teorema Espectral) este operador admite uma base de autovetores ortonormal. Vamos determiná-la.

(i) Por subtrair $\lambda = 2$ da diagonal da matriz A, obtemos

$$\begin{bmatrix} -1 & 3 \\ 3 & -9 \end{bmatrix} \left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right] = \left[\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right] \Leftrightarrow -x + 3y = 0.$$

E, portanto, um autovalor associado a $\lambda = 2$ é $\left[\begin{smallmatrix} 3y \\ y \end{smallmatrix}\right] = y \left[\begin{smallmatrix} 3 \\ 1 \end{smallmatrix}\right]$.
(ii) Ao invés de seguir o procedimento padrão, vamos usar a informação que o Teorema Espectral nos fornece. Sabemos que os autovetores associado ao autovalor $\lambda = 2$ são ortogonais à $\left[\begin{smallmatrix} 3 \\ 1 \end{smallmatrix}\right]$, e, como estamos em $\mathbb{R}^2$, um autovetor associado a $\lambda = -8$ deve ser $\left[\begin{smallmatrix} -1 \\ 3 \end{smallmatrix}\right]$. De fato, veja que

$$\begin{bmatrix} 1 & 3 \\ 3 & -7 \end{bmatrix} \left[\begin{smallmatrix} -1 \\ 3 \end{smallmatrix}\right] = \left[\begin{smallmatrix} 8 \\ -24 \end{smallmatrix}\right] = -8 \left[\begin{smallmatrix} -1 \\ 3 \end{smallmatrix}\right].$$

Normalizando esses vetores, obtemos uma base ortonormal $\beta = \left\{ \begin{bmatrix} 3/\sqrt{10} \\ 1/\sqrt{10} \end{bmatrix}, \begin{bmatrix} -1/\sqrt{10} \\ 3/\sqrt{10} \end{bmatrix} \right\}$. A matriz mudança de coordenadas

$$P = [I]_\alpha^\beta = \begin{bmatrix} \frac{3}{\sqrt{10}} & \frac{-1}{\sqrt{10}} \\ \frac{1}{\sqrt{10}} & \frac{3}{\sqrt{10}} \end{bmatrix}.$$

Como α e β são bases ortonormais, temos que P é ortogonal, isto é, $P^t = P^{-1}$. com base nisso, teremos

$$D = \begin{bmatrix} 2 & 0 \\ 0 & -8 \end{bmatrix} = P^t A P.$$

10.3 Formas Bilineares Simétricas

Diremos que uma aplicação $f : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ será bilinear se ela satisfizer as seguintes propriedades para $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^n$ e $\alpha \in \mathbb{R}$ quaisquer,

$$f(\mathbf{u} + \alpha\mathbf{w}, \mathbf{v}) = f(\mathbf{u}, \mathbf{v}) + \alpha f(\mathbf{w}, \mathbf{v}) \text{ e } f(\mathbf{u}, \mathbf{v} + \alpha\mathbf{w}) = f(\mathbf{u}, \mathbf{v}) + \alpha f(\mathbf{u}, \mathbf{w}).$$

Além disso, diremos que ela é simétrica se

$$f(\mathbf{u}, \mathbf{v}) = f(\mathbf{v}, \mathbf{u}) \text{ para quaisquer que sejam } \mathbf{u}, \mathbf{v} \in \mathbb{R}^n.$$

Como exemplos de formas bilineares simétrica, podemos citar os produtos internos sobre $\mathbb{R}^n$.

Às formas bilineares simétricas, podemos associar, para cada base, uma matriz simétrica A da seguinte forma: fixe uma base $\alpha = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ de $\mathbb{R}^n$, tome $\mathbf{u}$ e $\mathbf{v}$ vetores de $\mathbb{R}^n$. Calculando as suas coordenadas, temos $\mathbf{u} = \sum_{i=1}^n x_i \mathbf{v}_i$,

$\mathbf{v} = \sum_{j=1}^{n} y_j \mathbf{v}_j$ então,

$$\begin{aligned} f(\mathbf{u}, \mathbf{v}) &= f\left(\sum_{i=1}^{n} x_i \mathbf{v}_i, \sum_{j=1}^{n} y_j \mathbf{v}_j\right) \\ &= \sum_{i=1}^{n} x_i f\left(\mathbf{v}_i, \sum_{j=1}^{n} y_j \mathbf{v}_j\right) \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} x_i y_j f(\mathbf{v}_i, \mathbf{v}_j). \end{aligned}$$

Portanto, para calcular $f(\mathbf{u}, \mathbf{v})$, basta saber os valores $f(\mathbf{v}_i, \mathbf{v}_j)$, $1 \leq i, j \leq n$. Se criarmos a matriz (simétrica) $A = [f]_\alpha = [f(\mathbf{v}_i, \mathbf{v}_j)]$, isto é, a matriz que na posição ij tem o valor $f(\mathbf{v}_i, \mathbf{v}_j)$, então,

$$f(\mathbf{u}, \mathbf{v}) = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \begin{bmatrix} f(\mathbf{v}_1, \mathbf{v}_1) & f(\mathbf{v}_1, \mathbf{v}_2) & \cdots & f(\mathbf{v}_1, \mathbf{v}_n) \\ f(\mathbf{v}_2, \mathbf{v}_1) & f(\mathbf{v}_2, \mathbf{v}_2) & \cdots & f(\mathbf{v}_2, \mathbf{v}_n) \\ \vdots & \vdots & \ddots & \vdots \\ f(\mathbf{v}_n, \mathbf{v}_1) & f(\mathbf{v}_n, \mathbf{v}_2) & \cdots & f(\mathbf{v}_n, \mathbf{v}_n) \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}.$$

Chamamos a matriz $A = [f]_\alpha = [f(\mathbf{v}_i, \mathbf{v}_j)]$ de **matriz associada à forma bilinear simétrica na base** α.

Exemplo 10.10
Considere a forma bilinear f do $\mathbb{R}^3$ definida por

$$f\left(\left[\begin{smallmatrix} x_1 \\ x_2 \\ x_3 \end{smallmatrix}\right], \left[\begin{smallmatrix} y_1 \\ y_2 \\ y_3 \end{smallmatrix}\right]\right) = 2x_1y_1 - 4x_2y_1 - 8x_3y_1 - 4x_1y_2 + x_2y_2 + 7x_3y_2 - 8x_1y_3 + 7x_2y_3 + 5x_3y_3.$$

Seja $\mathcal{C} = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ a base canônica do $\mathbb{R}^3$. Calculando $f(\mathbf{e}_1, \mathbf{e}_1) = 2$, $f(\mathbf{e}_1, \mathbf{e}_2) = -4$ e todas as outras possibilidades, obtemos a matriz

$$[f]_{\mathcal{C}} = \begin{bmatrix} 2 & -4 & -8 \\ -4 & 1 & 7 \\ -8 & 7 & 5 \end{bmatrix}.$$

Definição 10.11
Dizemos que uma aplicação $Q : \mathbb{R}^n \to \mathbb{R}$ será uma forma quadrática, se $Q(\mathbf{u}) = f(\mathbf{u}, \mathbf{u})$ para alguma forma bilinear simétrica f de $\mathbb{R}^n$.

Há uma maneira de, considerando uma forma quadrática Q, obtermos a forma bilinear f necessária para a definição 10.11. Considere a aplicação $f : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ dada por

$$f(\mathbf{u}, \mathbf{v}) = \frac{1}{2}\left[Q(\mathbf{u} + \mathbf{v}) - Q(\mathbf{u}) - Q(\mathbf{v})\right] \text{ (Fórmula de Polarização).}$$

Para maiores detalhes, veja o exercício R10.6.

Portanto, dado uma forma quadrática $Q : \mathbb{R}^n \to \mathbb{R}$, podemos determinar uma forma bilinear simétrica f, através da fórmula de polarização. Além disso, se fixarmos uma base α de $\mathbb{R}^n$ podemos associar uma matriz $A = [f]_\alpha$ à forma bilinear f. Por isso, também, chamamos a matriz A de **matriz da forma quadrática Q com respeito à base α**.

É possível fazer o caminho inverso. Se tivermos uma matriz $A = [a_{ij}]_{n\times n}$, simétrica, poderemos determinar uma forma quadrática fazendo $Q(\mathbf{u}) = \mathbf{u}^t A\mathbf{u}$.

Observação 10.12
Seja f uma forma bilinear representada pela matriz simétrica $A = [a_{ij}]$. Considere $\mathbf{u} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$. Então a forma quadrática Q pode ser representada por

$$Q(\mathbf{u}) = f(\mathbf{u},\mathbf{u}) = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$
$$= \sum_{i,j} a_{ij}x_ix_j = \sum_i a_{ii}x_i^2 + \sum_{i<j} a_{ij}x_ix_j.$$

Segue que Q será uma forma quadrática se, e somente se, $Q(\mathbf{u})$ for um polinômio homogêneo de grau 2 nas entradas do vetor $\mathbf{u}$.

Pela observação 10.12, dada a aplicação $\mathbb{Q} : \mathbb{R}^2 \to \mathbb{R}$, definida por

$$Q\left(\left[\begin{smallmatrix} x_1 \\ x_2 \end{smallmatrix}\right]\right) = ax_1^2 + bx_1x_2 + bx_2^2, \text{ onde } a,\, b,\, c \in \mathbb{R},$$

temos que esta é uma forma quadrática, pois consiste de um polinômio homogêneo de grau 2. Além disso, é fácil verificar que

$$Q\left(\left[\begin{smallmatrix} x_1 \\ x_2 \end{smallmatrix}\right]\right) = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} a & \frac{b}{2} \\ \frac{b}{2} & c \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

Essa observação se aplica a qualquer $n \geq 3$, veja o exemplo 10.13.

Exemplo 10.13
Encontre a matriz simétrica A, associada à forma quadrática

$$Q\left(\left[\begin{smallmatrix} x \\ y \\ z \end{smallmatrix}\right]\right) = 2x^2 - 8xy + y^2 - 16xz + 14yz + 5z^2.$$

Veja que os coeficientes que aparecem na diagonal são os mesmos que acompanham os termos x^2, y^2 e z^2. Os coeficientes que acompanham os outros termos devem aparecer, divididos por 2, fora da diagonal principal. Por exemplo: -8 que é o coeficiente que acompanha xy deve aparecer, dividido por 2, na linha 1 coluna 2 e na linha 2 coluna 1. Então,

$$Q\left(\left[\begin{smallmatrix} x \\ y \\ z \end{smallmatrix}\right]\right) = \begin{bmatrix} x & y & z \end{bmatrix} \begin{bmatrix} 2 & -4 & -8 \\ -4 & 1 & 7 \\ -8 & 7 & 5 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}.$$

A matriz acima é matriz de Q com respeito à base canônica do $\mathbb{R}^3$ (veja exemplo 10.10).

10.3.1 Mudança de Coordenada e Formas Bilineares Simétricas

Suponha que $\mathbf{x}, \mathbf{y}$ são as coordenadas de dois vetores do $\mathbb{R}^n$ em relação à base $\alpha = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$. Digamos que $\mathbf{x}'$ e $\mathbf{y}'$ representam os mesmos os vetores em relação a outra base $\beta = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$. Dessa forma, existe uma matriz invertível $P = [I]_\alpha^\beta$, que satisfaz

$$\mathbf{x} = P\mathbf{x}' \text{ e, de outra forma, } \mathbf{y} = P\mathbf{y}'.$$

Se $f : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ é uma forma bilinear simétrica e $A = [f(\mathbf{u}_i, \mathbf{u}_j)]$, então:

$$f(\mathbf{x}, \mathbf{y}) = \mathbf{x}^t A\mathbf{y} = (P\mathbf{x}')^t AP\mathbf{y}' = \mathbf{x}'^t P^t AP\mathbf{y}' = f(\mathbf{x}', \mathbf{y}').$$

Segue que a matriz de f, com respeito a base β, isto é, $B = [f(\mathbf{v}_i, \mathbf{v}_j)]$, deve satisfazer $B = P^t AP$.

Definição 10.14
Diremos que uma matriz B será **congruente** a uma matriz A (e escrevemos $B \cong A$), se existir, uma matriz invertível P tal que $B = P^t AP$.

Observação 10.15
Não confunda o conceito de congruência (definição 10.14) entre matrizes com o de semelhança (definição 7.11) entre matrizes. Apesar disso, se a matriz P for ortogonal, os dois conceitos irão coincidir.

De acordo com a definição acima, podemos dizer que matrizes associadas a uma mesma forma quadrática, em relação a bases diferentes, serão sempre congruentes, e, reciprocamente, matrizes congruentes estarão associadas a uma mesma forma quadrática, só que relacionadas a outro sistema de coordenadas.

Exemplo 10.16
Considere $\mathbf{u} = \left[\begin{smallmatrix} x_1 \\ x_2 \end{smallmatrix}\right] \in \mathbb{R}^2$ e a forma quadrática $q : \mathbb{R}^2 \to \mathbb{R}$, definida por $q(\mathbf{u}) = x_1^2 + 6x_1x_2 - 7x_2^2$. Esta pode ser reescrita, em termos da matriz associada, da seguinte forma:

$$q(\left[\begin{smallmatrix} x_1 \\ x_2 \end{smallmatrix}\right]) = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ 3 & -7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

Se escolhemos outra base, digamos $\beta = \left\{ \begin{bmatrix} 3/\sqrt{10} \\ 1/\sqrt{10} \end{bmatrix}, \begin{bmatrix} -1/\sqrt{10} \\ 3/\sqrt{10} \end{bmatrix} \right\}$ (veja o exemplo 10.9). Sabemos que a base inicial era a canônica, portanto a matriz de mudança de coordenadas é

$$P = [I]_\alpha^\beta = \begin{bmatrix} \frac{3}{\sqrt{10}} & \frac{-1}{\sqrt{10}} \\ \frac{1}{\sqrt{10}} & \frac{3}{\sqrt{10}} \end{bmatrix}.$$

Se $[\mathbf{u}]_\beta = \left[\begin{smallmatrix} y_1 \\ y_2 \end{smallmatrix}\right]$ são as coordenadas do vetor $\mathbf{u}$ com respeito à base β, então:

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \frac{3}{\sqrt{10}} & \frac{-1}{\sqrt{10}} \\ \frac{1}{\sqrt{10}} & \frac{3}{\sqrt{10}} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \text{ e } \begin{bmatrix} x_1 & x_2 \end{bmatrix} = \begin{bmatrix} y_1 & y_2 \end{bmatrix} \begin{bmatrix} \frac{3}{\sqrt{10}} & \frac{1}{\sqrt{10}} \\ \frac{-1}{\sqrt{10}} & \frac{3}{\sqrt{10}} \end{bmatrix}.$$

Dessa forma, temos que

$$\begin{aligned} q(\mathbf{u}) = q(\left[\begin{smallmatrix} x_1 \\ x_2 \end{smallmatrix}\right]) &= \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ 3 & -7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \\ &= \begin{bmatrix} y_1 & y_2 \end{bmatrix} \begin{bmatrix} \frac{3}{\sqrt{10}} & \frac{-1}{\sqrt{10}} \\ \frac{1}{\sqrt{10}} & \frac{3}{\sqrt{10}} \end{bmatrix}^t \begin{bmatrix} 1 & 3 \\ 3 & -7 \end{bmatrix} \begin{bmatrix} \frac{3}{\sqrt{10}} & \frac{-1}{\sqrt{10}} \\ \frac{1}{\sqrt{10}} & \frac{3}{\sqrt{10}} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \\ &= \begin{bmatrix} y_1 & y_2 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & -8 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = 2y_1^2 - 8y_2^2 = q(\left[\begin{smallmatrix} y_1 \\ y_2 \end{smallmatrix}\right]). \end{aligned}$$

Quem já estudou equações do 2º grau ($Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0$) sabe que a principal dificuldade é retirar o fator xy, mas as contas acima nos sugerem uma técnica para resolvermos esse problema.

10.4 Algoritmo de Diagonalização

Apesar de já conhecermos um procedimento para diagonalizar matrizes simétricas, vamos apresentar um algoritmo que nos permite diagonalizar uma matriz simétrica sem precisar passar pelo processo de calcular o polinômio característico, os autovalores e autovetores. Esse algoritmo exige apenas o conhecimento de operações elementares. Contra esse procedimento, pesa o fato de que a base na qual a matriz simétrica se torna diagonal nem sempre é ortonormal.

Vamos exibir um algoritmo que permite obter uma matriz invertível P que diagonaliza a matriz A. Existe uma sequência de operações elementares aplicado às linhas e uma forma de aplicar, essas operações, às colunas, que transforma a matriz A na matriz diagonal D, além disso, $D = P^t AP$.

Antes de iniciar a leitura do procedimento, veja o exemplo ilustrativo a seguir.

Observação 10.17
Considere a matriz simétrica

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

Recordemos que se quisermos fazer uma operação elementar entre as linhas, por exemplo, $\ell_2 \to \ell_2 - 2\ell_1$, basta realizarmos a mesma operação na matriz identidade, obtendo uma matriz que chamamos de matriz elementar E. Então, se fizermos EA, obteremos a mesma matriz que havíamos obtido quando fizemos a operação elementar em A. Além disso, se tomarmos E^t e fizermos AE^t, obteremos a mesma matriz A que sofreu a operação $\mathbf{c}_2 \to \mathbf{c}_2 - 2\mathbf{c}_1$ nas colunas.

Veja o exemplo:

$$EA = \begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 2 & 5 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \text{ e } AE^t = \begin{bmatrix} 1 & 2 \\ 2 & 5 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix}^t = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}.$$

Tendo ele como parâmetro, vamos descrever o algoritmo:

Algoritmo
Passo 0 Dada a matriz simétrica $A = [a_{ij}]$ de ordem n.

Passo 1 Forme uma matriz $n \times 2n$ em blocos $M = [A_1 : I]$, em que $A_1 = A$ é a metade à esquerda de M e a metade à direita de M é a matriz identidade.

Passo 2 Ao examinar o termo a_{11} de M há três possibilidades:

Caso I Se $a_{11} \neq 0$, então faça as operações $\ell_i \to \ell_i - \frac{a_{i1}}{a_{11}}\ell_1$ com $i = 2, 3, \ldots, n$ e, depois, faça o mesmo com as colunas, isto é, $\mathbf{c}_i \to \mathbf{c}_i - \frac{a_{i1}}{a_{11}}\mathbf{c}_1$ com $i = 2, 3, \ldots, n$. Essas operações irão reduzir a matriz M à forma

$$M \approx \begin{bmatrix} a_{11} & 0 & * & * \\ 0 & A_2 & * & * \end{bmatrix}.$$

Caso II - Se $a_{11} = 0$, mas $a_{kk} \neq 0$, para algum $k > 1$. Nesse caso, faça $\ell_1 \leftrightarrow \ell_k$ e, também, $\mathbf{c}_1 \leftrightarrow \mathbf{c}_k$.
Essas operações reduzem a matriz ao Caso I.
Caso III - Todas as entradas diagonais $a_{ii} = 0$, mas algum $a_{ij} \neq 0$. Nesse caso, faça $\ell_i \to \ell_j + \ell_i$ e o mesmo na coluna $\mathbf{c}_i \to \mathbf{c}_j + \mathbf{c}_i$ (Essas operações levam $2a_{ij}$ para a i-ésima entrada diagonal). Assim M volta à situação do Caso II.

Passo 3 Repita o passo 2 com cada nova matriz A_k (suprimindo as primeiras linhas e colunas da matriz precedente) até que A esteja diagonalizada. Então, M torna-se $M' = [D : Q]$, onde D é diagonal.

Passo 4 Tome $P = Q^t$. Então, $D = P^t AP$.

Observação 10.18
No passo 2, as operações com as linhas modificam os dois lados de M, mas as operações com as colunas modificam apenas o lado esquerdo de M.

Exemplo 10.19
Seja $A = \begin{bmatrix} 1 & 2 & -3 \\ 2 & 5 & -4 \\ -3 & -4 & 8 \end{bmatrix}$. Utilizando, nas operações, o algoritmo anterior, vamos determinar uma matriz não singular P tal que $D = P^t AP$ seja diagonal.

Iniciamos contruindo a matriz $M = [A : I]$

$$M = [A : I] = \begin{bmatrix} 1 & 2 & -3 & 1 & 0 & 0 \\ 2 & 5 & -4 & 0 & 1 & 0 \\ -3 & -4 & 8 & 0 & 0 & 1 \end{bmatrix}.$$

Aplicando as operações $\ell_2 \to \ell_2 - 2\ell_1$ e $\ell_3 \to \ell_3 + 3\ell_1$ nas linhas de M, e depois, aplicando as operações correspondentes $\mathbf{c}_2 \to \mathbf{c}_2 - 2\mathbf{c}_1$ e $\mathbf{c}_3 \to \mathbf{c}_3 + 3\mathbf{c}_1$ nas colunas, obtemos

$$\begin{bmatrix} 1 & 2 & -3 & 1 & 0 & 0 \\ 0 & 1 & 2 & -2 & 1 & 0 \\ 0 & 2 & -1 & 3 & 0 & 1 \end{bmatrix} \text{ e depois, } \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 2 & -2 & 1 & 0 \\ 0 & 2 & -1 & 3 & 0 & 1 \end{bmatrix}.$$

Em seguida, faremos $\ell_3 \to \ell_3 - 2\ell_2$ e a operação correspondente $\mathbf{c}_3 \to \mathbf{c}_3 - 2\mathbf{c}_2$ obtendo, assim,

$$\begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 2 & -2 & 1 & 0 \\ 0 & 0 & -5 & 7 & -2 & 1 \end{bmatrix} \text{ e depois, } \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & -2 & 1 & 0 \\ 0 & 0 & -5 & 7 & -2 & 1 \end{bmatrix}.$$

Dessa forma, A foi diagonalizada. Escrevemos

$$P = \begin{bmatrix} 1 & -2 & 7 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{bmatrix} \text{ e, portanto, } D = P^t AP = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -5 \end{bmatrix}.$$

Observe que P é a transposta da metade direita da matriz final.

Por fim, vamos justificar o algoritmo. Considere a matriz em blocos $M = [A : I]$. O algoritmo aplica uma sequência de operações elementares nas linhas, seguida de uma sequência de operações nas colunas do lado esquerdo de M, que é a matriz A. Isso equivale a pré-multiplicar A por uma sequência de matrizes elementares, digamos $E_1, E_2, \ldots, E_k$ e, depois, a multiplicar A pelas transpostas das $E_{i's}$. Assim, quando o algoritmo terminar, a matriz diagonal D, à esquerda de M, será igual a

$$D = E_k \cdots E_2 E_1 A E_1^t E_2^t \cdots E_k^t, \text{ com } Q = E_k \cdots E_2 E_1.$$

Por outro lado, o algoritmo somente aplica as operações elementares com as linhas à matriz identidade I do lado direito de M. Assim, quando o algoritmo terminar, a matriz à direita de M, é igual a

$$E_k \cdots E_2 E_1 I = E_k \cdots E_2 E_1 = Q.$$

E por tomar $P = Q^t$, iremos obter $D = P^t AP$.

10.5 Classificação das Equações do 2º grau

Agora, vamos cumprir o prometido no início do capítulo e aplicar a teoria para classificar todas as formas quadráticas. Para isso, iniciemos com uma equação do 2º grau completa nas variáveis x, y da forma:

$$g(x,y) = b_{11}x^2 + b_{12}xy + b_{22}y^2 + a_1 x + a_2 y + a = 0.$$

Se fizermos

$$\mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix}, A = \begin{bmatrix} b_{11} & \frac{b_{12}}{2} \\ \frac{b_{12}}{2} & b_{22} \end{bmatrix} \text{ e } A_1 = \begin{bmatrix} a_1 & a_2 \end{bmatrix},$$

então, poderemos escrever

$$\begin{aligned} g(x,y) &= \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} b_{11} & \frac{b_{12}}{2} \\ \frac{b_{12}}{2} & b_{22} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} a_1 & a_2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + a \\ &= \mathbf{x}^t A \mathbf{x} + A_1 \mathbf{x} + a = 0. \end{aligned}$$

Como A é uma matriz simétrica, pelo teorema 10.8, existe uma matriz invertível e ortogonal P tal que

$$\begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} = P^t AP \text{ e seja } \mathbf{y} = \begin{bmatrix} x' \\ y' \end{bmatrix} \text{ tal que } \mathbf{x} = P\mathbf{y}.$$

E a equação inicial tomará a forma

$$\begin{aligned} g(x', y') &= \mathbf{y}^t \left(P^t A P\right) \mathbf{y} + A_1 P \mathbf{y} + a \\ &= \begin{bmatrix} x' & y' \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix} + \begin{bmatrix} d_1 & d_2 \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix} + a \\ &= \lambda_1 x'^2 + \lambda_2 y'^2 + d_1 x' + d_2 y' + a = 0, \end{aligned}$$

onde d_1 e d_2 serão os coeficientes obtidos ao fazer $A_1 P$. Voltando à equação, temos duas situações:

Se $\lambda_1\lambda_2 \neq 0$ (Se $\lambda_1\lambda_2 < 0$, temos uma hipérbole, e se $\lambda_1\lambda_2 > 0$, temos a elipse) podemos completar os quadrados, se necessário. E, com isso, obtemos

$$q(x', y') = \lambda_1 \left(x' + \frac{d_1}{2\lambda_1}\right)^2 + \lambda_2 \left(y' + \frac{d_2}{2\lambda_2}\right)^2 + b = 0,$$

onde $b = a - \left(\frac{d_1}{2\lambda_1}\right)^2 - \left(\frac{d_2}{2\lambda_2}\right)^2$. Assim, basta realizarmos uma translação no $\mathbb{R}^2$, dada por

$$x'' = x' + \frac{d_1}{2\lambda_1}, \qquad y'' = y' + \frac{d_2}{2\lambda_2},$$

para que a equação se torne

$$q(x'', y'') = \lambda_1 x''^2 + \lambda_2 y''^2 + b = 0.$$

Se $\lambda_1\lambda_2 = 0$ (parábola) Pelo menos um dos dois autovalores é não nulo, e digamos que $\lambda_1 \neq 0$ então, podemos completar o quadrado com respeito a x' e obtemos

$$q(x', y') = \lambda_1 \left(x' + \frac{d_1}{2\lambda_1}\right)^2 + d_2 y' + b' = 0,$$

onde $b' = a - \left(\frac{d_1}{2\lambda_1}\right)^2$. Fazendo a mudança de coordenadas

$$x'' = x' + \frac{d_1}{2\lambda_1}, \qquad y'' = y',$$

a equação se torna

$$q(x'', y'') = \lambda_1 x''^2 + d_2 y''^2 + b' = 0.$$

Em ambos os casos, é fácil classificar a equação. Além disso, podemos determinar quais mudanças de coordenadas foram necessárias para levá-la a essa configuração.

Percebemos, também, que esse procedimento se generaliza para qualquer número de variáveis ≥ 2.

Observação 10.20
Observe que no procedimento anterior, passamos da matriz

$$A = \begin{bmatrix} b_{11} & \frac{b_{12}}{2} \\ \frac{b_{12}}{2} & b_{22} \end{bmatrix} \text{ para a matriz } \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$$

por uma mudança de coordenadas. Utilizamos a matriz P, que é ortogonal, e obtemos

$$\det \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} = \det(A) = \frac{1}{4}\left(4b_{11}b_{22} - b_{12}^2\right).$$

Com exceção dos casos nos quais a equação representa figuras degeneradas, a classificação dessa equação em elipse, parábola ou hipérbole depende, apenas, do sinal das constantes λ_1 e λ_2, e podemos determinar esses sinais se conhecemos o produto $\lambda_1\lambda_2 = \det(A)$. Esse sinal pode ser obtido calculando $4b_{11}b_{22} - b_{12}^2$. Por isso, chamamos esse número de **discriminante da equação do 2º grau**.

Exemplo 10.21
Vamos aplicar o procedimento descrito acima para classificar a equação dada por

$$4x^2 - 24yx + 56x + 11y^2 - 58y + 95 = 0.$$

Inicialmente, podemos escrever

$$\begin{aligned} g(x,y) &= 4x^2 - 24yx + 11y^2 + 56x - 58y + 95 \\ &= \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} 4 & -12 \\ -12 & 11 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} 56 & -58 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + 95 \\ &= \mathbf{x}^t A\mathbf{x} + A_1\mathbf{x} + 95 = 0. \end{aligned}$$

Depois, precisamos calcular a base de autovetores de

$$A = \begin{bmatrix} 4 & -12 \\ -12 & 11 \end{bmatrix} \Rightarrow \Delta_A(x) = x^2 - 15x - 100 = (x+5)(x-20).$$

i) Calculando o autovetor associado a $\lambda_1 = -5$, obtemos $\left[\begin{smallmatrix} 4 \\ 3 \end{smallmatrix}\right]$.
ii) O autovetor associado a $\lambda_2 = 20$ é $\left[\begin{smallmatrix} -3 \\ 4 \end{smallmatrix}\right]$.

Portanto, em relação à base $\beta = \left\{ \left[\begin{smallmatrix} 4/5 \\ 3/5 \end{smallmatrix}\right], \left[\begin{smallmatrix} -3/5 \\ 4/5 \end{smallmatrix}\right] \right\}$, as coordenadas dos vetores serão escritas por ser $\mathbf{u} = \left[\begin{smallmatrix} x' \\ y' \end{smallmatrix}\right]$ e considerando a matriz

$$P = \begin{bmatrix} 4/5 & -3/5 \\ 3/5 & 4/5 \end{bmatrix}, \text{ então } \begin{bmatrix} x \\ y \end{bmatrix} = P \begin{bmatrix} x' \\ y' \end{bmatrix}.$$

E, no sistema de coordenadas β, a equação se torna

$$\begin{aligned} g(x',y') &= \begin{bmatrix} x' & y' \end{bmatrix} \begin{bmatrix} -5 & 0 \\ 0 & 20 \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix} + \begin{bmatrix} 10 & -80 \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix} + 95 \\ &= -5x'^2 + 20y'^2 + 10x' - 80y' + 95 \\ &= -5(x'^2 - 2x') + 20(y'^2 - 4y') + 95 \\ &= -5(x'^2 - 2x' + 1) + 20(y'^2 - 4y' + 4) + (95 + 5 - 80) \\ &= -5(x'-1)^2 + 20(y'-2)^2 + 20. \end{aligned}$$

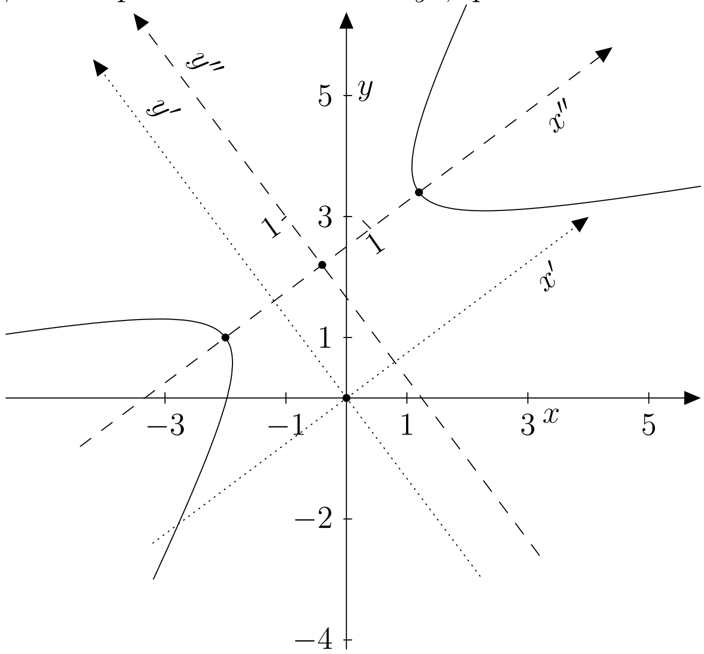
Fazendo a mudança de coordenadas

$$x'' = x' - 1, \quad y'' = y' - 2,$$

obtemos

$$g(x'', y'') = -5x''^2 + 20y''^2 + 20 = 0 \Leftrightarrow \frac{x''^2}{4} - y''^2 = 1.$$

Que é, portanto, uma hipérbole no sistema $x''y''$, que não cruza o eixo y''.



Vamos utilizar o segundo método (o algoritmo) para eliminar o fator xy na equação

$$4x^2 - 24yx + 56x + 11y^2 - 58y + 95 = 0$$

$$\begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} 4 & -12 \\ -12 & 11 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} 56 & -58 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + 95 = 0.$$

Considere a matriz

$$M = \begin{bmatrix} 4 & -12 & 1 & 0 \\ -12 & 11 & 0 & 1 \end{bmatrix}.$$

Para escaloná-la, faça $\ell_2 \to \ell_2 + 3\ell_1$ e $\mathbf{c}_2 \to \mathbf{c}_2 + 3\mathbf{c}_1$ e obtemos

$$\begin{bmatrix} 4 & 0 & 1 & 0 \\ 0 & -25 & 3 & 1 \end{bmatrix}.$$

Tomando

$$P = \begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix}, \text{ temos que } \begin{bmatrix} 4 & 0 \\ 0 & -25 \end{bmatrix} = P^t AP.$$

Portanto, se as coordenadas de $\mathbf{u}$ são $\left[\begin{smallmatrix} x' \\ y' \end{smallmatrix}\right]$, com respeito à base $\beta = \{\left[\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right], \left[\begin{smallmatrix} 3 \\ 1 \end{smallmatrix}\right]\}$, então

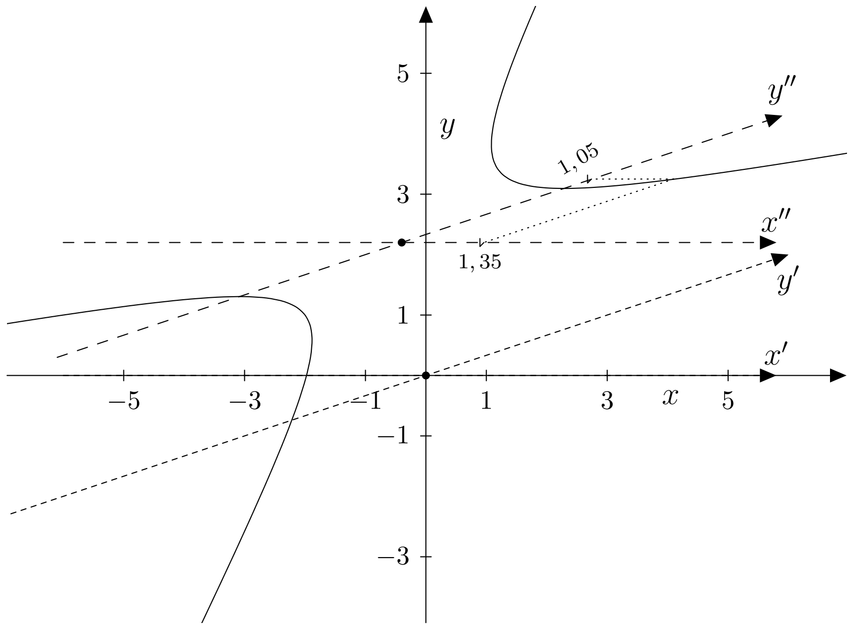
$$P = \begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix}, \text{ e sabemos que } \begin{bmatrix} x \\ y \end{bmatrix} = P \begin{bmatrix} x' \\ y' \end{bmatrix}.$$

Então,

$$\begin{aligned}
q(x', y') &= \begin{bmatrix} x' & y' \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 0 & -25 \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix} + \begin{bmatrix} 56 & 110 \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix} + 95 \\
&= 4x'^2 - 25y'^2 + 56x' + 110y' + 95 \\
&= 4(x'^2 + 14x' + 49) - 25\left(y'^2 - \frac{22}{5}y' + \left(\frac{22}{5}\right)^2\right)^2 - 4 \times 49 + 25\left(\frac{22}{5}\right)^2 + 95 \\
&= 4(x' + 7)^2 - 25(y'^2 - \frac{11}{5})^2 + 20.
\end{aligned}$$

Fazendo $x'' = x' + 7$ e $y'' = y' - \frac{11}{5}$, temos

$$4x''^2 - 25y''^2 + 20 = 0 \Leftrightarrow -\frac{x''^2}{5} + \frac{y''^2}{\frac{4}{5}} = 1.$$



Veja que o sistema de coordenadas $x''y''$ não é ortogonal nem preserva as medidas.

Com as técnicas propostas aqui, é possível classificar qualquer equação do segundo grau como uma das situações: degeneradas ou como

$$y^2 = 4px, \ \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1, \ \frac{x^2}{a^2} - \frac{y^2}{b^2} = 1.$$

Exercícios resolvidos

R10.1. Demonstre a proposição 10.3. Fixe um ponto F (Foco), uma linha ℓ (a diretriz) que não contém F, escolha, um número Real não-negativo e (a excentricidade). Considere os pontos P, satisfazendo

$$\operatorname{dist}(P, F) = e \operatorname{dist}(P, \ell).$$

Mostre que, se: $0 \leq e < 1$ obtemos uma elipse; se $e = 1$, obtemos uma parábola; e se $e > 1$, obtemos uma hipérbole.

Solução: Observe, que, se $e = 0$, claramente obtemos uma circunferência de raio 0 e centro P. Suponha, que $P = \left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right]$ e, $F = \left[\begin{smallmatrix} p \\ 0 \end{smallmatrix}\right]$ e ℓ seja o eixo do y, isto é, a reta $x = 0$. Pela definição, temos

$$\sqrt{(x-p)^2 + y^2} = e|x| \Leftrightarrow (1-e^2)x^2 - 2px + y^2 + p^2 = 0.$$

Se $e = 1$, então, podemos reescrever a equação acima como

$$y^2 = 2p(x - \frac{p}{2}),$$

que é uma parábola.

Se $e \neq 1$, podemos dividir por $1 - e^2 \neq 0$. Completando o quadrado e dividindo por $-\frac{p^2}{1-e^2}$, obtemos

$$\frac{\left(x - \frac{p}{1-e^2}\right)^2}{\frac{p^2e^2}{(1-e^2)^2}} + \frac{y^2}{\frac{p^2e^2}{1-e^2}} = 1.$$

Se $e < 1$, então $1 - e^2 > 0$ e, com isso, temos uma elipse com

$$a^2 = \frac{p^2e^2}{(1-e^2)^2} \text{ e } b^2 = \frac{p^2e^2}{1-e^2}.$$

Se $e > 1$, então $1 - e^2 < 0$ e, com isso, temos uma hipérbole. $\square$

R10.2. Prove o teorema 10.4. Sejam $\langle\, , \,\rangle$ um produto interno em $\mathbb{R}^n$ e $T : \mathbb{R}^n \to \mathbb{R}^n$ um operador linear. Então, T será um operador autoadjunto se, e somente se,

$$\langle T\mathbf{u}, \mathbf{v}\rangle = \langle \mathbf{u}, T\mathbf{v}\rangle\,, \text{ para todo } \mathbf{u}, \mathbf{v} \in \mathbb{R}^n.$$

Solução: Seja $\alpha = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ a base ortonormal de $\mathbb{R}^n$, na qual a matriz $[T]_\alpha^\alpha = [a_{ij}]$ é simétrica. Os coeficientes a_{ij} são determinados por

$$T(\mathbf{u}_j) = \sum_{i=1}^{n} a_{ij}\mathbf{u}_i, \text{ para todo } 1 \leq j \leq n.$$

Por calcular

$$\begin{aligned} \langle T(\mathbf{u}_j), \mathbf{u}_k\rangle &= \left\langle \sum_{i=1}^{n} a_{ij}\mathbf{u}_i, \mathbf{u}_k \right\rangle \\ &= a_{kj} = a_{jk} \\ &= \left\langle \mathbf{u}_j, \sum_{i=1}^{n} a_{ik}\mathbf{u}_i \right\rangle = \langle \mathbf{u}_j, T(\mathbf{u}_k)\rangle\,. \end{aligned}$$

E, com isso, vemos que essa propriedade é verdadeira se substituírmos os vetores $\mathbf{u}, \mathbf{v}$ pelos vetores da base α. Seja agora $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ dois vetores quaisquer, e por escrevê-los em termos da base α, obtemos $\mathbf{u} = \sum_{i=1}^{n} x_i \mathbf{u}_i$ e $\mathbf{v} = \sum_{j=1}^{k} y_j \mathbf{u}_j$ e, portanto,

$$
\begin{aligned}
\langle T\mathbf{u}, T\mathbf{v} \rangle &= \left\langle T(\sum_{i=1}^{n} x_i \mathbf{u}_i), T(\sum_{j=1}^{k} y_j \mathbf{u}_j) \right\rangle \\
&= \sum_{i=1}^{n} x_i \left\langle T(\mathbf{u}_i), T(\sum_{j=1}^{k} y_j \mathbf{u}_j) \right\rangle \\
&= \sum_{i=1}^{n} \sum_{j=1}^{k} x_i y_j \left\langle T(\mathbf{u}_i), T(\mathbf{u}_j) \right\rangle \\
&= \sum_{i=1}^{n} \sum_{j=1}^{k} x_i y_j \left\langle \mathbf{u}_i, \mathbf{u}_j \right\rangle \\
&= \sum_{i=1}^{n} x_i \left\langle \mathbf{u}_i, \sum_{j=1}^{k} y_j \mathbf{u}_j \right\rangle \\
&= \left\langle \sum_{i=1}^{n} x_i \mathbf{u}_i, \sum_{j=1}^{k} y_j \mathbf{u}_j \right\rangle = \langle \mathbf{u}, \mathbf{v} \rangle .
\end{aligned}
$$

Reciprocamente, suponha que:

$$\langle T\mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{u}, T\mathbf{v} \rangle \text{ para todo } \mathbf{u}, \mathbf{v} \in \mathbb{R}^n.$$

E seja $\beta = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ uma base ortonormal qualquer de $\mathbb{R}^n$, então, a matriz $[T]_\beta^\beta = [b_{ij}]$ será determinada por satisfazer

$$T(\mathbf{v}_j) = \sum_{i=1}^{n} b_{ij} \mathbf{v}_i.$$

E como β é uma base ortonormal temos que

$$b_{lj} = \langle T(\mathbf{v}_j), \mathbf{v}_l \rangle = \langle \mathbf{v}_j, T(\mathbf{v}_l) \rangle = b_{jl}.$$

Portanto, $[T]_\beta^\beta$ é simétrica. Daí, segue que, se a matriz de uma transformação é simétrica, então ela é simétrica, também, com relação a qualquer base ortonormal. □

R10.3. Prove a proposição 10.6. Seja $\langle\, , \,\rangle$ um produto interno em $\mathbb{R}^n$. Se $T : \mathbb{R}^n \to \mathbb{R}^n$ é um operador autoadjunto. Então:

a) Autovetores correspondentes a autovalores distintos são ortogonais.

b) T sempre possui pelo menos um autovalor.

Solução: a) Sejam $i \neq j$ quaisquer:

$$\begin{aligned}(\lambda_i - \lambda_j)\langle \mathbf{u}_i, \mathbf{u}_j\rangle &= \langle \lambda_i \mathbf{u}_i, \mathbf{u}_j\rangle - \langle \mathbf{u}_i, \lambda_j \mathbf{u}_j\rangle \\ &= \langle T\mathbf{u}_i, \mathbf{u}_j\rangle - \langle \mathbf{u}_i, T\mathbf{u}_j\rangle = \langle T\mathbf{u}_i, \mathbf{u}_j\rangle - \langle T\mathbf{u}_i, \mathbf{u}_j\rangle = 0,\end{aligned}$$

uma vez que T é autoadjunto. Como $\lambda_i - \lambda_j \neq 0$ e $(\lambda_i - \lambda_j)\langle \mathbf{u}_i, \mathbf{u}_j\rangle = 0$, isso implica que $\langle \mathbf{u}_i, \mathbf{u}_j\rangle = 0$.

b) Considere a aplicação $\phi : \mathbb{R}^n \to \mathbb{R}$, definida por $\phi(\mathbf{u}) = \langle T\mathbf{u}, \mathbf{u}\rangle$, que é contínua e deverá assumir um máximo em $K : \{\mathbf{u} \in \mathbb{R}^n : \|\mathbf{u}\| = 1\}$, uma vez que o mesmo é compacto. Seja $\mathbf{u}_1$ esse ponto, e λ_1 o valor de ϕ nesse ponto. Então, para todo $\mathbf{u} \in \mathbb{R}^n$, vale

$$\begin{aligned}\langle (T - \lambda_1 I)\mathbf{u}, \mathbf{u}\rangle &= \|\mathbf{u}\|^2 \left\langle (T - \lambda_1 I)\frac{1}{\|\mathbf{u}\|}\mathbf{u}, \frac{1}{\|\mathbf{u}\|}\mathbf{u}\right\rangle \\ &= \|\mathbf{u}\|^2 \left[\left\langle T\left(\frac{1}{\|\mathbf{u}\|}\mathbf{u}\right), \frac{1}{\|\mathbf{u}\|}\mathbf{u}\right\rangle - \lambda_1\right] \leq 0,\end{aligned}$$

mostrando que

$$\langle (T - \lambda_1 I)\mathbf{u}, \mathbf{u}\rangle \leq 0, \ \forall \mathbf{u} \in \mathbb{R}^n.$$

Então, se escolhermos $\mathbf{u} = \mathbf{u}_1 + t\mathbf{v}$, com $t \in \mathbb{R}$ e $\mathbf{v} \in \mathbb{R}^n$, teremos

$$\begin{aligned}&\langle (T - \lambda_1 I)(\mathbf{u}_1 + t\mathbf{v}), \mathbf{u}_1 + t\mathbf{v}\rangle \\ &\qquad = t^2 \langle (T - \lambda_1 I)(\mathbf{v}), \mathbf{v}\rangle + t\left[\langle (T - \lambda_1 I)(\mathbf{u}_1, \mathbf{v}\rangle + \langle (T - \lambda_1 I)\mathbf{v}, \mathbf{u}_1\rangle\right] \\ &\qquad\qquad + \langle (T - \lambda_1 I)\mathbf{u}_1, \mathbf{u}_1\rangle \\ &\qquad = t^2 \langle (T - \lambda_1 I)(\mathbf{v}), \mathbf{v}\rangle + 2t\langle (T - \lambda_1 I)\mathbf{u}_1, \mathbf{v}\rangle \leq 0.\end{aligned}$$

Temos uma equação do 2º grau com o coeficiente que acompanha t^2 negativo e, portanto, o discriminante precisa ser menor ou igual a zero, isto é, $\langle (T - \lambda_1 I)\mathbf{u}_1, \mathbf{v}\rangle = 0$ para todo $\mathbf{v} \in \mathbb{R}^n$. Então $(T - \lambda_1 I)\mathbf{u}_1 = 0$, mostrando, assim, a existência de um autovalor λ_1 e de um autovetor $\mathbf{u}_1$. □

R10.4. Mostre que a equação da reta tangente à parábola $y^2 = 4px$ em $P = \left[\begin{smallmatrix} x_0 \\ y_0 \end{smallmatrix}\right]$ é

$$y_0 y = 2p(x + x_0).$$

Solução: Queremos determinar a equação da reta tangente

$$y - y_0 = m(x - x_0)$$

à parábola passando por $P = \left[\begin{smallmatrix} x_0 \\ y_0 \end{smallmatrix}\right]$. Precisamos determinar m. Para isto, faça a substituição $y = mx - mx_0 + y_0$ na equação da parábola e teremos

$$(mx - mx_0 + y_0)^2 = 4px \Leftrightarrow m^2x^2 + \left(2my_0 - 2m^2x_0 - 4p\right)x + m^2x_0^2 - 2mx_0y_0 + y_0^2.$$

Essa equação do 2º grau só pode ter uma solução e, portanto, o seu discriminante precisa ser igual a zero, isto é,

$$\left(2my_0 - 2m^2x_0 - 4p\right)^2 - 4m^2\left(m^2x_0^2 - 2mx_0y_0 + y_0^2\right) = 0.$$

Fazendo as contas obtemos

$$16p(x_0m^2 - y_0m + p) = 0 \Leftrightarrow x_0m^2 - y_0m + p = 0.$$

Resolvendo para m, obtemos

$$m = \frac{y_0 + \sqrt{y_0^2 - 4x_0p}}{2x_0}.$$

Isso nos dá $m = \frac{y_0}{2x_0}$ visto que P pertence à parábola. Substituindo m na equação da reta inicial, obtemos a equação enunciada. □

R10.5. Prove o teorema Espectral 10.8: Seja $\langle\,,\,\rangle$ um produto interno em $\mathbb{R}^n$. Se $T : \mathbb{R}^n \to \mathbb{R}^n$ é um operador auto-adjunto, então existe uma base ortogonal de $\mathbb{R}^n$ composta por autovetores de T.

Solução: Vamos demonstrar por indução sobre $n = \dim \mathbb{R}^n$. No caso que $n = 1$, todo operador é autoadjunto, uma vez que toda matriz 1×1 é simétrica. Suponha o resultado verdadeiro em espaço com dimensão $n-1$, e vamos provar quando a dimensão é n. Seja $T : R^n \to \mathbb{R}^n$ um operador autoadjunto, pela proposição 10.6 b., existe um autovetor unitário $\mathbf{u}_1$ associado ao autovalor λ_1. Considere o subespaço $W = \operatorname{Span}\{\mathbf{u}_1\}$, que é um subespaço T-invariante de dimensão 1. Segue, pelo lema 10.7, que o subespaço, com dimensão $n-1$, $W^\perp$ também é T-invariante. Logo, $\bar{T} = T|_{W^\perp} : W^\perp \to W^\perp$ também é autoadjunto, pela hipótese de indução, existe uma base ortonormal $\{\mathbf{u}_2, \ldots, \mathbf{u}_n\} \subset W^\perp$ formada por autovetores de $\bar{T}$. Segue que $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ é uma base ortonormal formada por autovetores de T. □

R10.6. Seja $Q : \mathbb{R}^n \to \mathbb{R}$ uma forma quadrática associada à forma bilinear simétrica f. Mostre que vale a seguinte identidade:

$$f(\mathbf{u}, \mathbf{v}) = \frac{1}{2}\left[Q(\mathbf{u} + \mathbf{v}) - Q(\mathbf{u}) - Q(\mathbf{v})\right],$$

conhecida como fórmula de polarização.

Solução: Sabemos que $Q(\mathbf{u}) = f(\mathbf{u}, \mathbf{u})$, e seja $\mathbf{u}, \mathbf{w}$ e $\mathbf{v} \in \mathbb{R}^n$ e $\alpha \in \mathbb{R}$. Então

$$\begin{aligned} Q(\mathbf{u}+\mathbf{v}) - &Q(\mathbf{u}) - Q(\mathbf{v}) \\ &= f(\mathbf{u}+\mathbf{v}, \mathbf{u}+\mathbf{v}) - f(\mathbf{u}, \mathbf{v}) - f(\mathbf{v}, \mathbf{v}) \\ &= f(\mathbf{u}, \mathbf{u}+\mathbf{v}) + f(\mathbf{v}, \mathbf{u}+\mathbf{v}) - f(\mathbf{u}, \mathbf{v}) - f(\mathbf{v}, \mathbf{v}) \\ &= f(\mathbf{u}, \mathbf{u}) + f(\mathbf{u}, \mathbf{v}) + f(\mathbf{v}, \mathbf{u}) + f(\mathbf{v}, \mathbf{v}) - f(\mathbf{u}, \mathbf{v}) - f(\mathbf{v}, \mathbf{v}) \\ &= 2f(\mathbf{u}, \mathbf{v}). \end{aligned}$$

Dividindo por 2, obtemos a fórmula de polarização. □

Exercícios propostos

P10.1. Identifique a figura e ache sua posição quando a sua equação for:

a) $2xy + 3x - y + 1 = 0$;

b) $x^2 + y^2 + xy - x + 1 = 0$;

c) $9x^2 + 12y^2 + 9z^2 - 6xy - 6yz - 1 = 0$;

d) $11x^2 + 10y^2 + 6z^2 - 12xy - 8yz + 4xz - 12 = 0$;

e) $xy + x + y = 0$;

f) $3x^2 - 4\sqrt{3}xy - y^2 + 20y - 25 = 0$;

g) $4xy + 3y^2 + 2\sqrt{5}x + 4\sqrt{5}y = 0$;

h) $9x^2 + 16y^2 + 25z^2 + 24xy - 40x + 30y = 0$.

P10.2. Mostre que as equações paramétricas $x = a\cos t$ e $y = b\,\mathrm{sen}\, t$, $a > 0$, $b > 0$ e $0 \leq t < 2\pi$ definem uma elipse cuja equação em coordenadas cartesianas é $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$.

P10.3. Mostre que as equações paramétricas $x = t$ e $y = \frac{t^2}{4p}$, com $t \in \mathbb{R}$, definem uma parábola cuja equação em coordenadas cartesianas é $y^2 = 4px$.

P10.4. Seja $\cosh t = \frac{e^x + e^{-x}}{2}$ e $\mathrm{senh} = \frac{e^x - e^{-x}}{2}$. Mostre que as equações paramétricas $x = a\cosh t$ e $y = b\,\mathrm{senh}\, t$, $a > 0$, $b > 0$ e $t \in \mathbb{R}$ definem uma hipérbole cuja equação em coordenadas cartesianas é $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$.

P10.5. Para as matrizes abaixo, aplique o algoritmo para encontrar uma matriz invertível P, tal que $D = P^t AP$ seja diagonal, onde

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

P10.6. a) Dê a transformação linear que descreve o movimento rígido que leva o segmento de extremos $\left[\begin{smallmatrix} -6 \\ 2 \end{smallmatrix}\right]$ e $\left[\begin{smallmatrix} -1 \\ 2 \end{smallmatrix}\right]$ ao segmento de extremos $\left[\begin{smallmatrix} -2 \\ 6 \end{smallmatrix}\right]$ e $\left[\begin{smallmatrix} 1 \\ 2 \end{smallmatrix}\right]$, respectivamente.

b) Mostre que essa transformação é uma rotação, e encontre o seu ângulo.

P10.7. Mostre que a equação da reta tangente à elipse $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ em $P = \left[\begin{smallmatrix} x_0 \\ y_0 \end{smallmatrix}\right]$ é

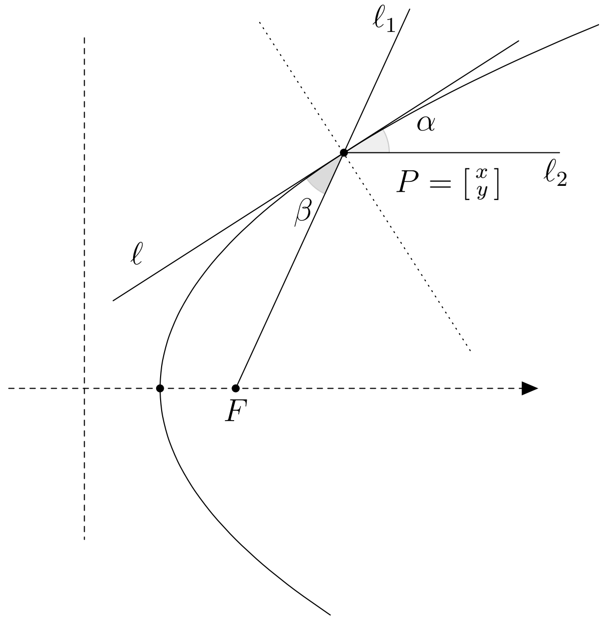
$$b^2 x_0 x + a^2 y_0 y = a^2 b^2.$$

P10.8. Suponha que sejam dadas as retas $y = m_1 x + k_1$ e $y = m_2 x + k_2$, as quais se interceptam em um ponto P. Seja θ o ângulo entre elas. Então,

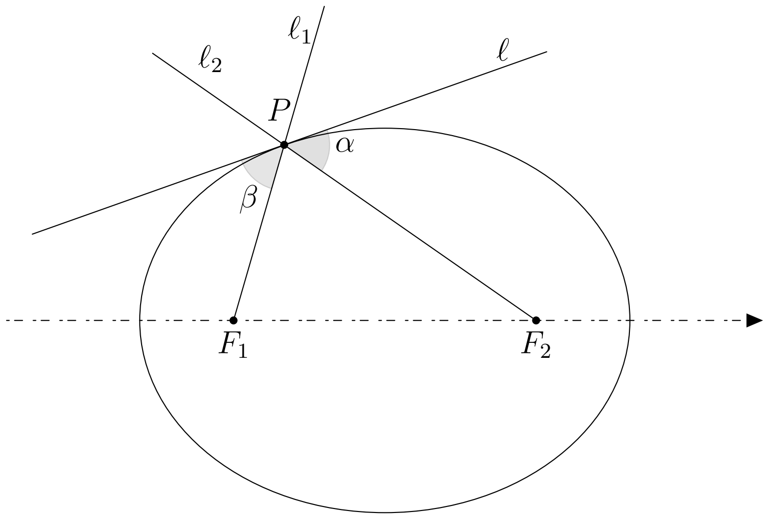
$$\mathrm{tg}\,\theta = \frac{m_2 - m_1}{1 + m_1 m_2},\ m_1 m_2 \neq -1,$$

onde m_1 é o coeficiente da reta onde se inicia o ângulo.

P10.9. Seja P um ponto de uma parábola e ℓ a reta tangente à parábola em P. Sejam ℓ_1 a reta passando pelo foco F e o ponto P e ℓ_2 a reta que passa por P, e que é paralela ao eixo x, então os ângulos, no ponto P, α entre ℓ e ℓ_1 e β entre ℓ e ℓ_2, são iguais.



P10.10. Sejam P um ponto sobre uma elipse e ℓ a reta tangente a esta em P. Considere ℓ_1 a reta ligando o foco em F_1 e ℓ_2, a reta ligando o foco F_2 a P. Então, os ângulos: α, entre as retas ℓ e ℓ_1, e β, entre ℓ e ℓ_2, são iguais.



Bibliografia

[1] Paul R. Halmos, *Finite-Dimensional Vector Spaces*. Springer-Verlag, 1974.

[2] Hamilton S. Bueno, *Álgebra Linear: um segundo curso*

[3] S. Lipschutz e M.L. Lipson, *Álgebra Linear*, Coleção Schaum, Bookman, 4ª edição, 2011.

[4] K. Hoffman e R. Kunze, *Álgebra Linear*, Prentice-Hall, INC, Englewood Cliff, New Jersey, 1971.

[5] H, Anton e Robert C. Busby, *Contemporary Linear Algebra*, John Wiley & Sons, Inc.

[6] Elon L. Lima, *Álgebra Linear*, Coleção Matemática Universitária, 2ª edição, 1996.