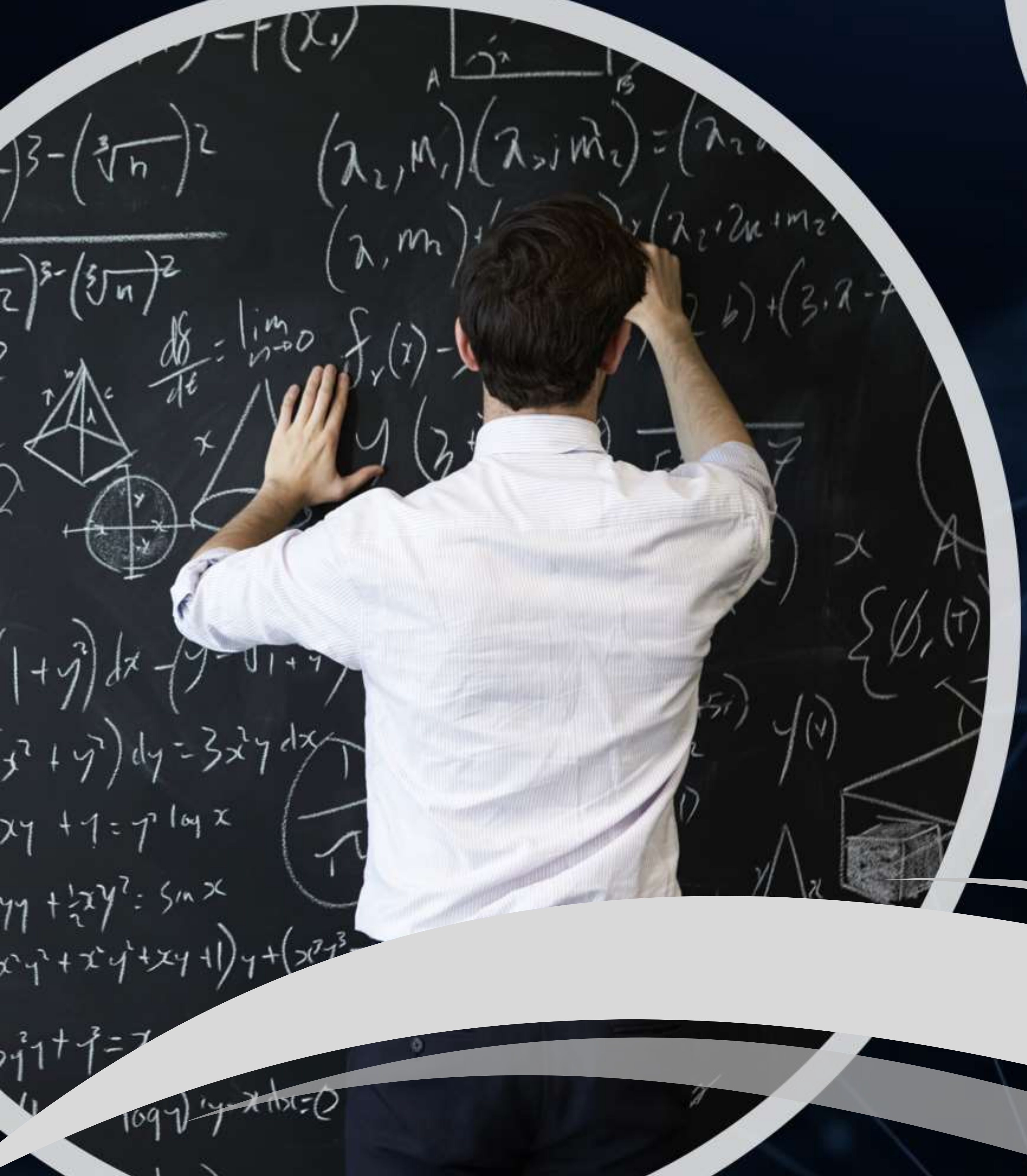


Matemática para Machine Learning

A sua base começa aqui!

Matemática para Machine Learning

Cálculo Multivariado

Data Science Academy

Cálculo Multivariado

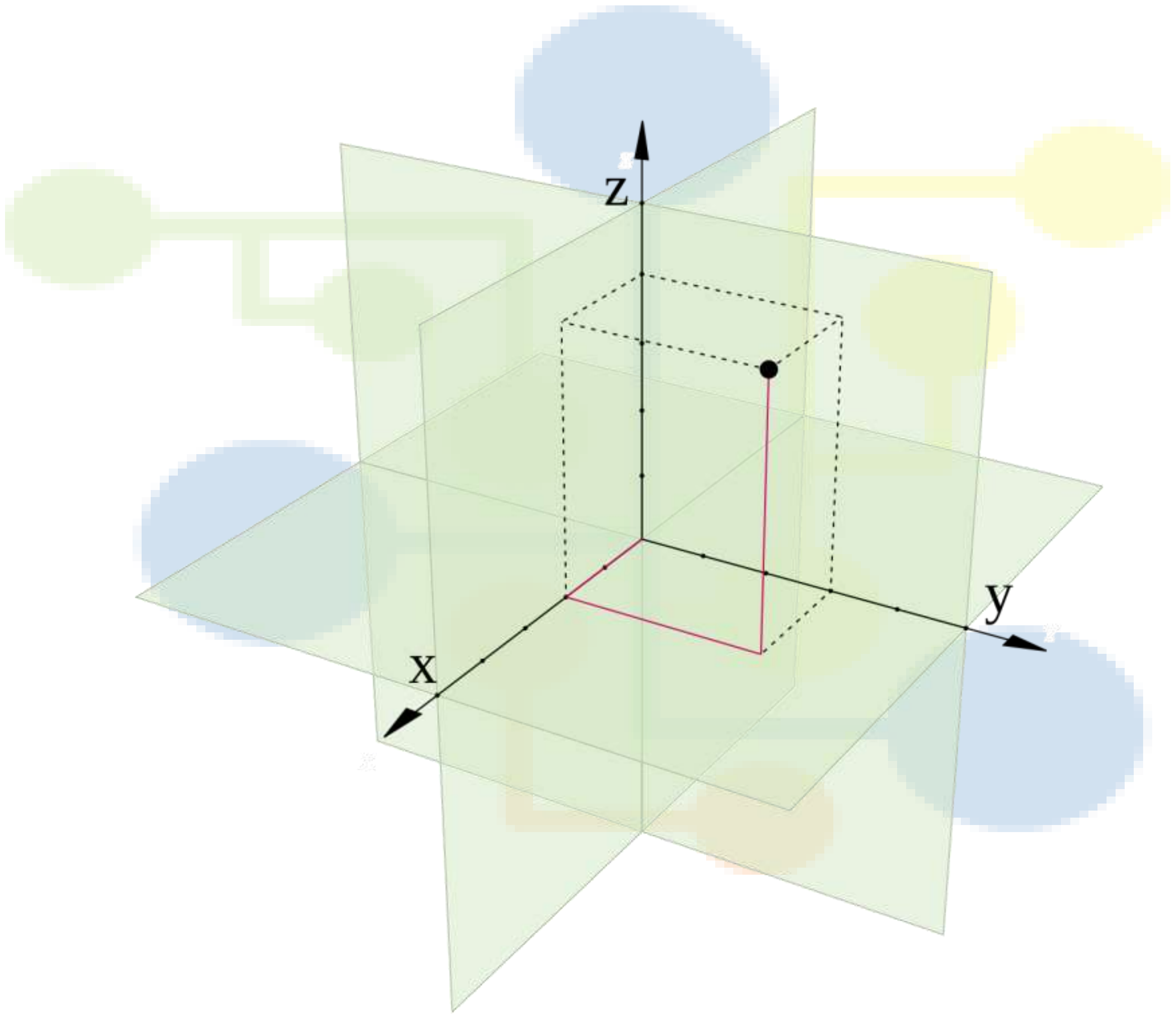
Cálculo Multivariado

Matemática para Machine Learning

O Espaço Bidimensional

Data Science Academy

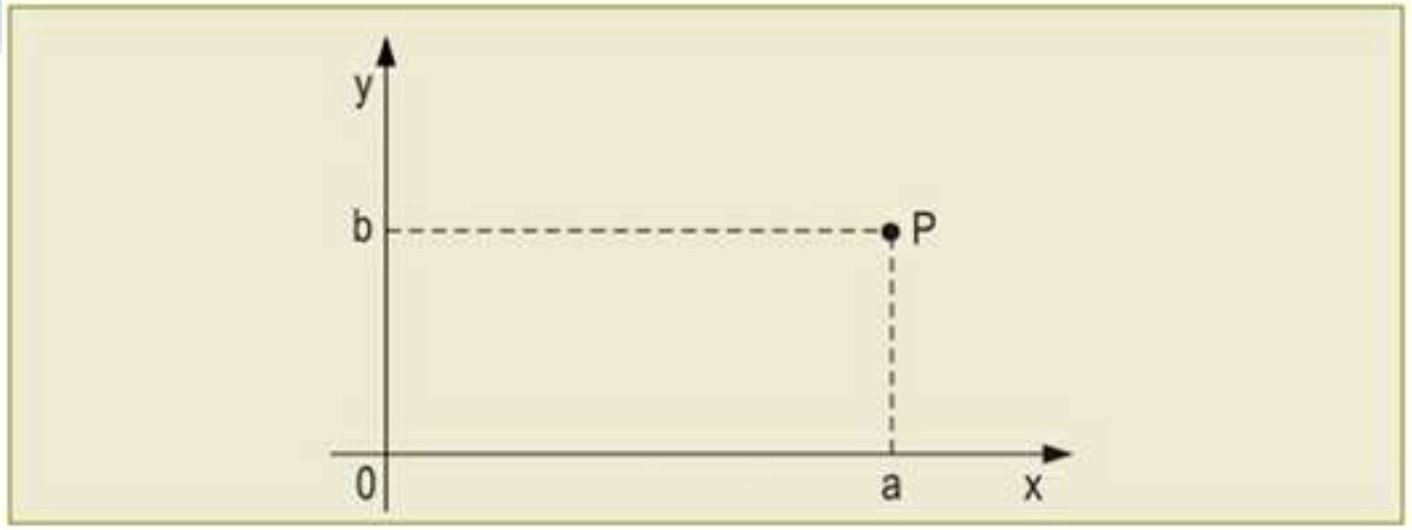
O Espaço Bidimensional



O Espaço Bidimensional

$$R^2 = \{(a, b) \mid a \in R \text{ e } b \in R\}$$

$$(3, 4); (-1, 2); (\frac{1}{2}, \frac{1}{2}); (0, \sqrt{2})$$



Matemática para Machine Learning

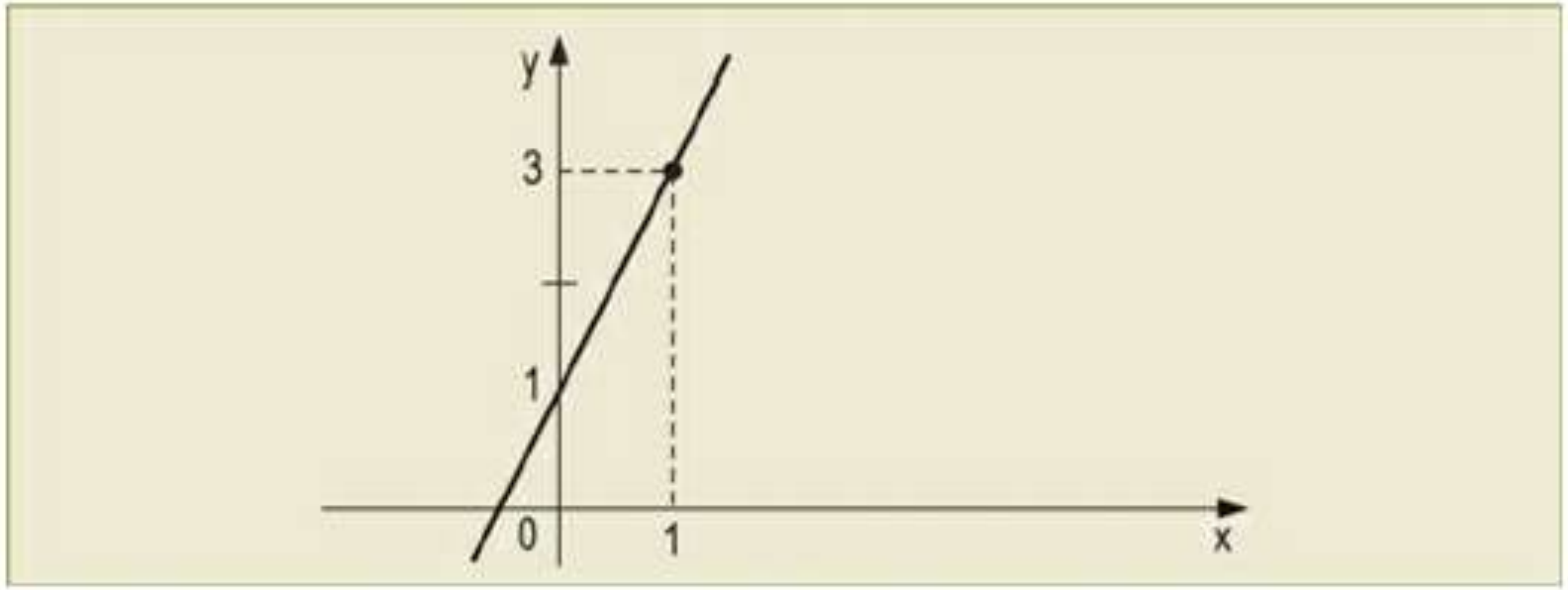
Relações em R2

Relações em R2

Chama-se relação binária, ou simplesmente relação no R2, a todo subconjunto de R2.

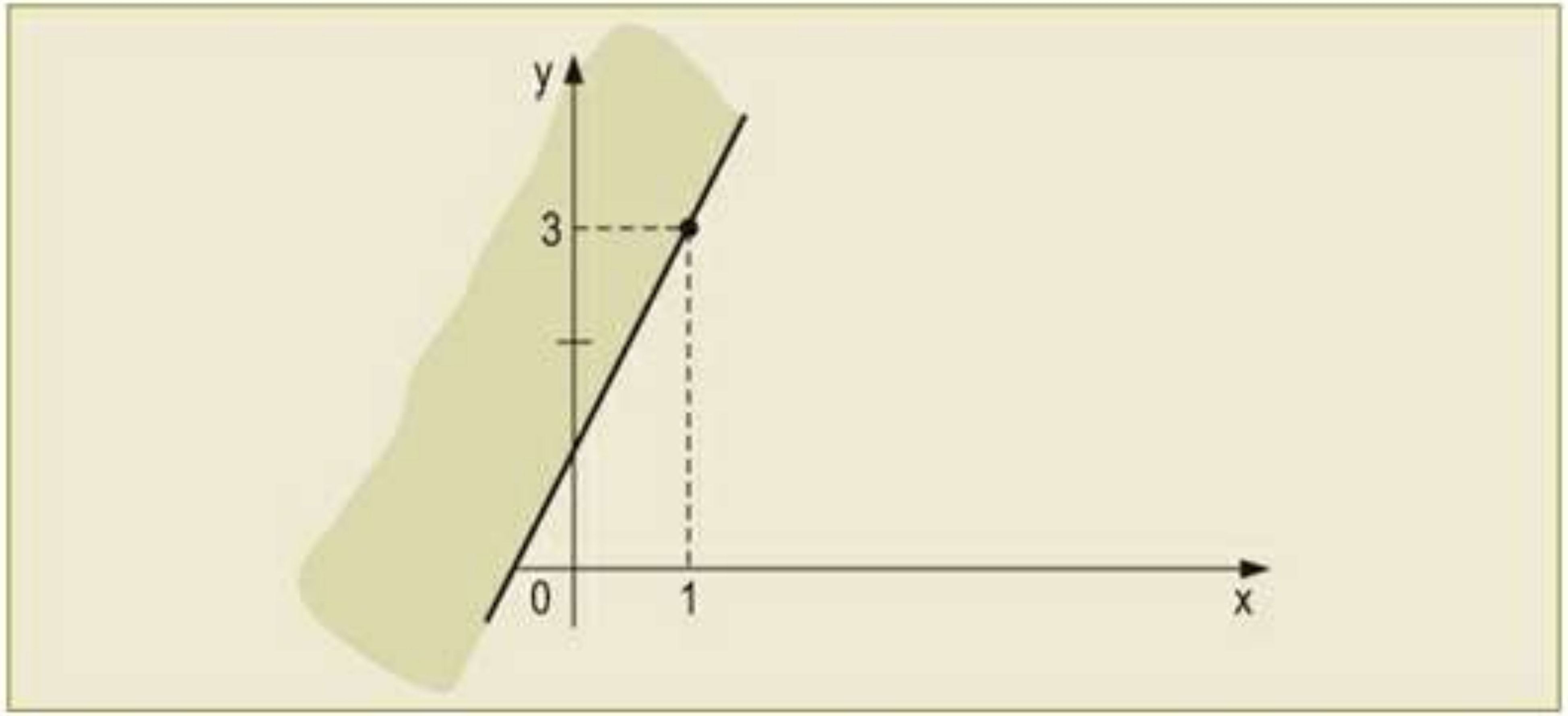
Relações em R2

$$A = \{(x, y) \in R^2 \mid y = 2x + 1\}$$



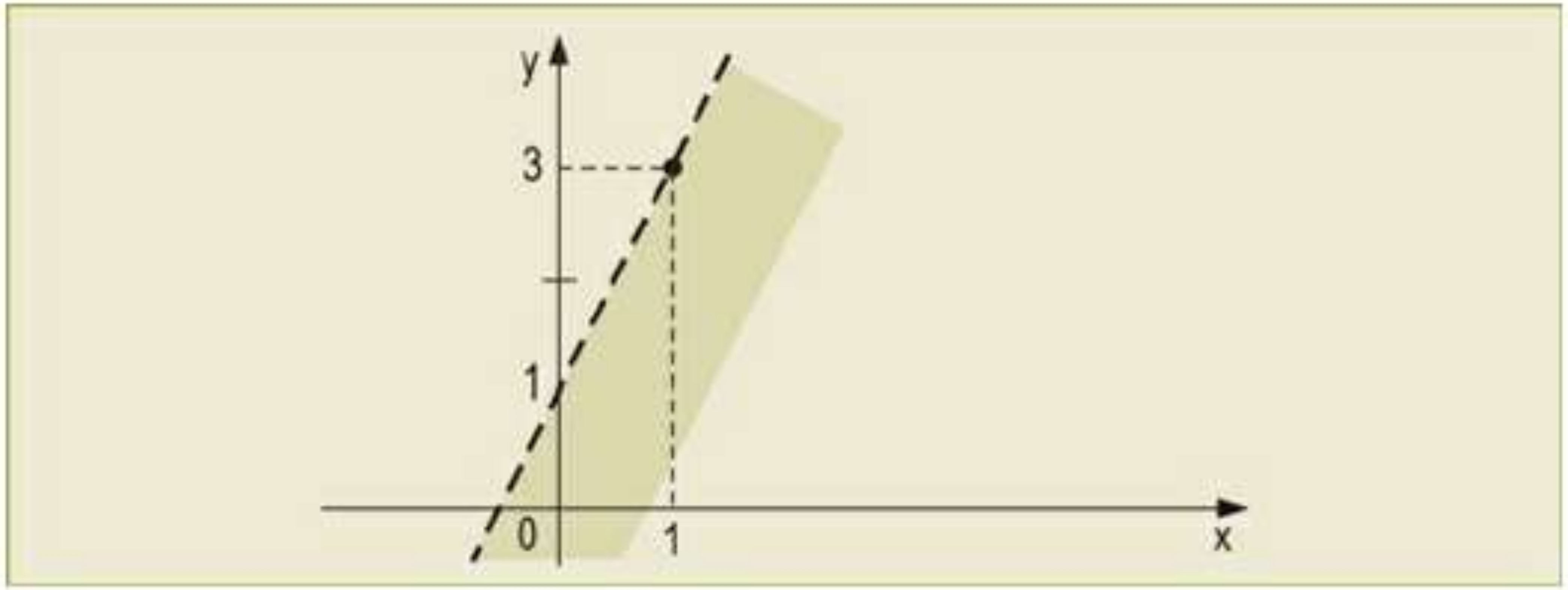
Relações em R2

$$B = \{(x, y) \in R^2 \mid y \geq 2x + 1\}$$



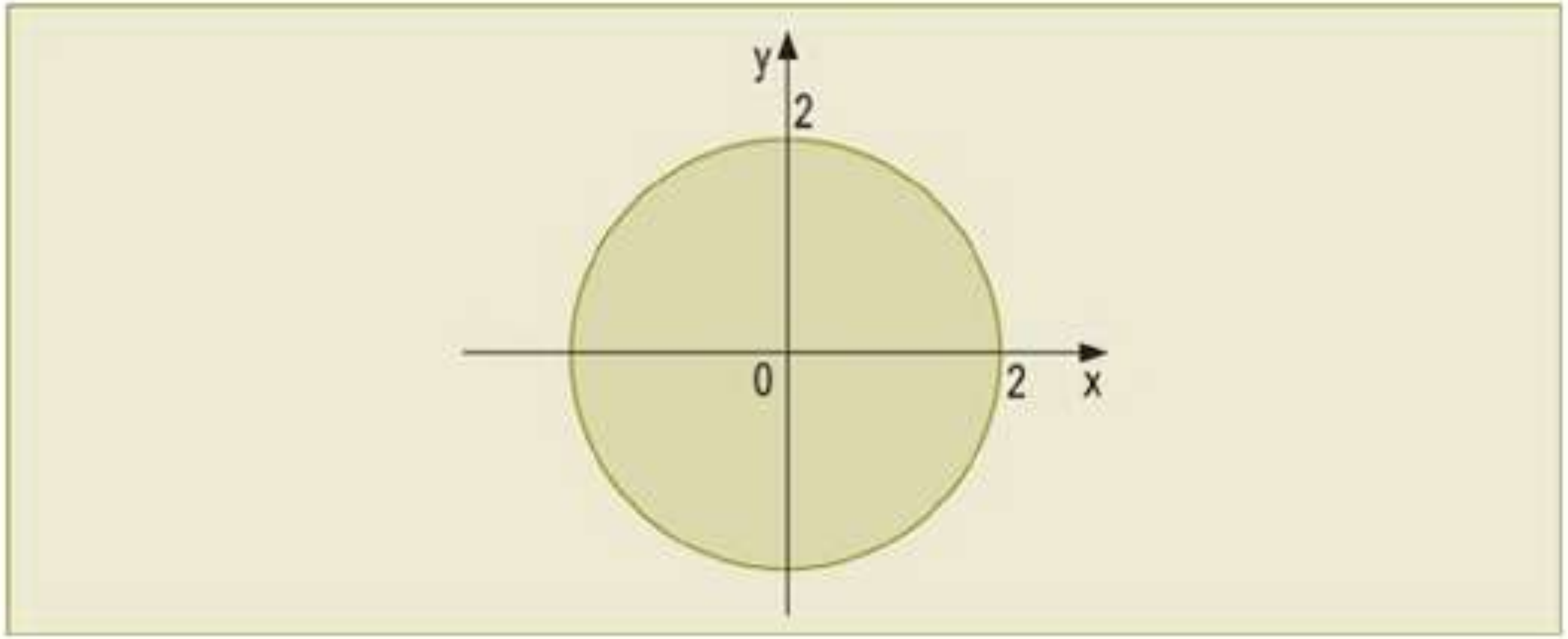
Relações em R2

$$C = \{(x, y) \in \mathbb{R}^2 \mid y < 2x + 1\}$$



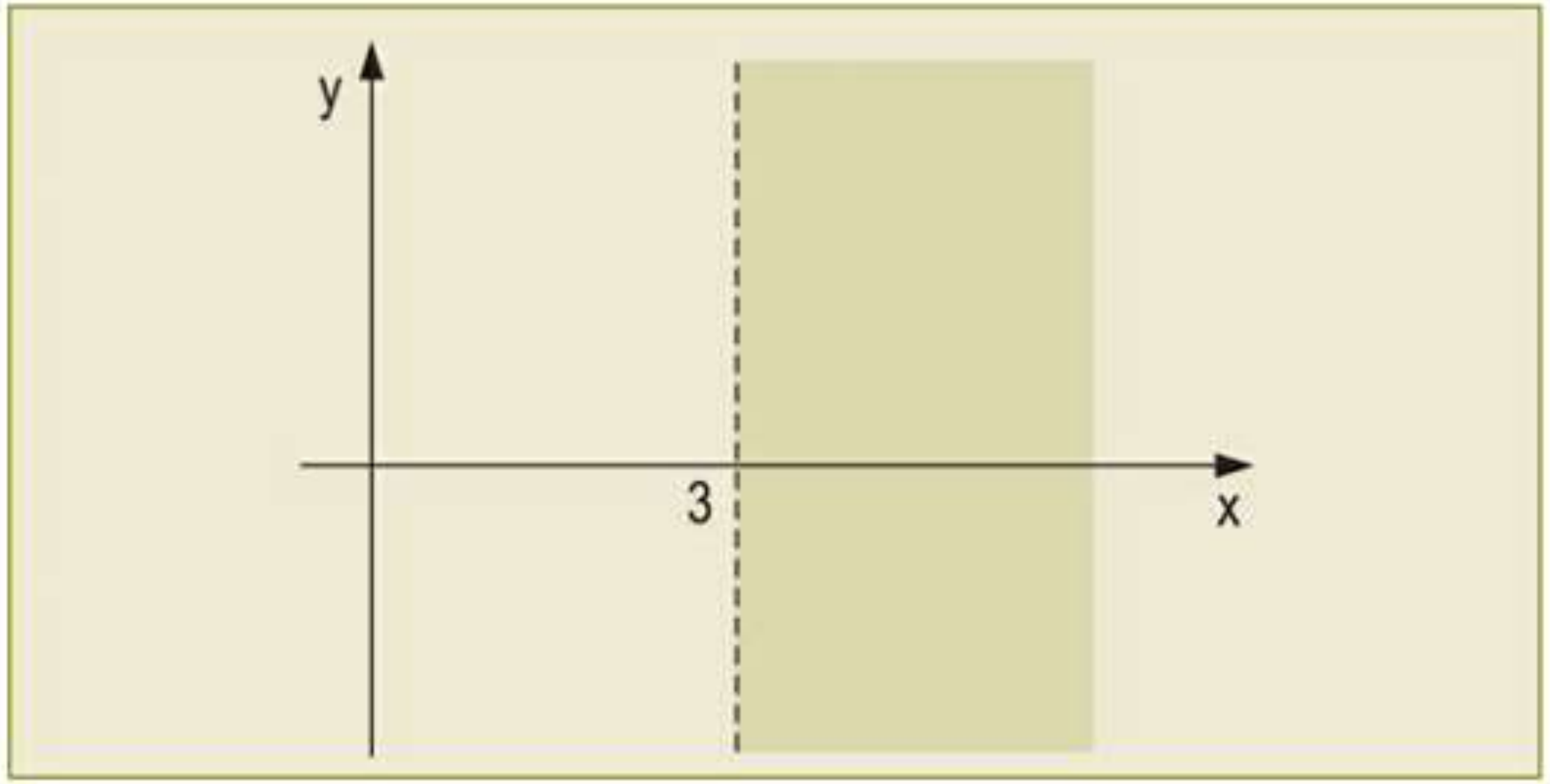
Relações em R2

$$D = \{(x, y) \in R^2 \mid x^2 + y^2 \leq 4\}$$



Relações em R2

$$E = \{(x, y) \in \mathbb{R}^2 \mid x > 3\}$$



Matemática para Machine Learning

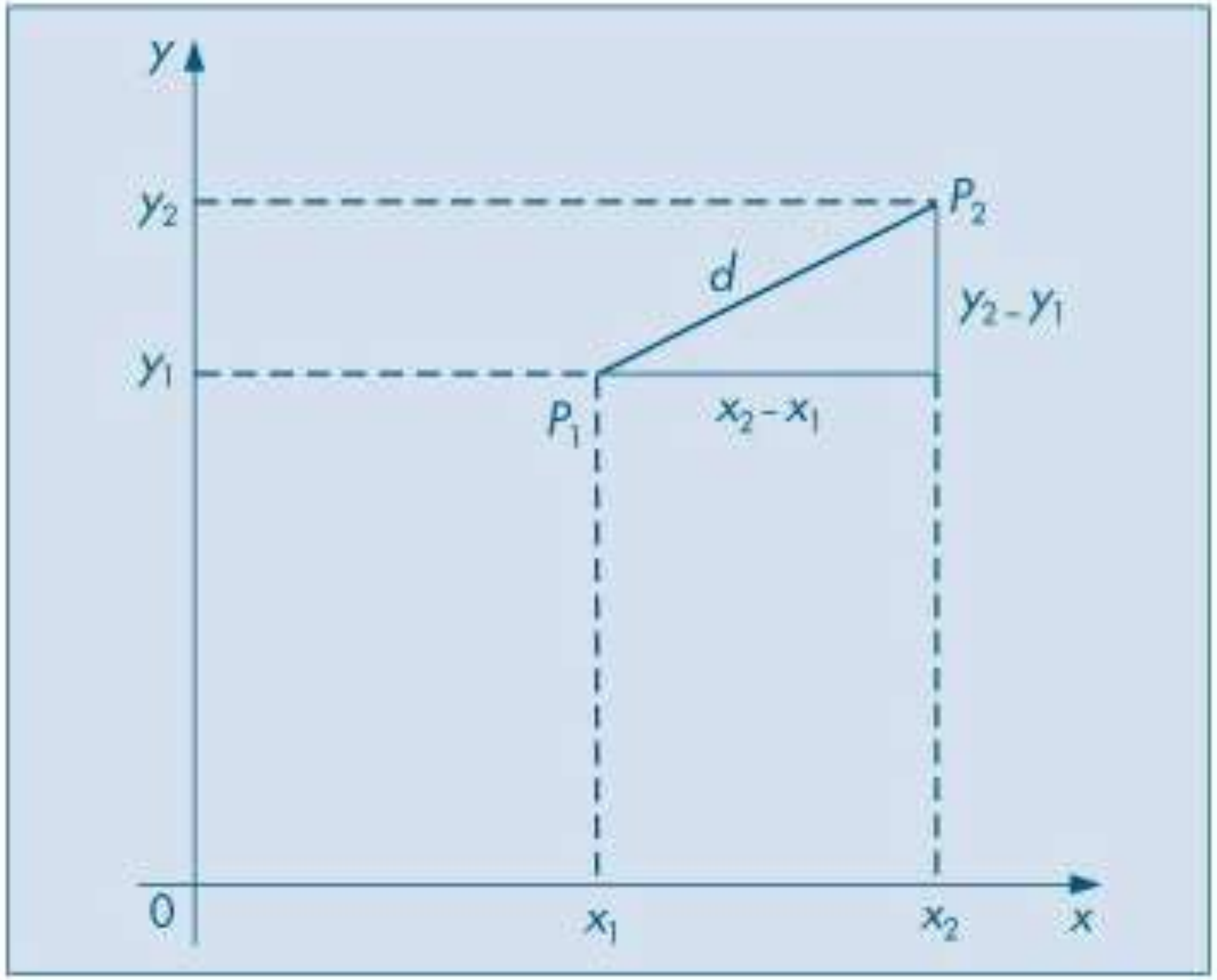
Distância Entre 2 Pontos

Distância Entre 2 Pontos

Sejam (x1, y1) e (x2, y2) dois elementos de R2, representados geometricamente pelos pontos P1 e P2.
A distância entre eles é o número:

$$d(P_1, P_2) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

Distância Entre 2 Pontos



$$d(P_1, P_2) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

Matemática para Machine Learning

O Espaço Tridimensional

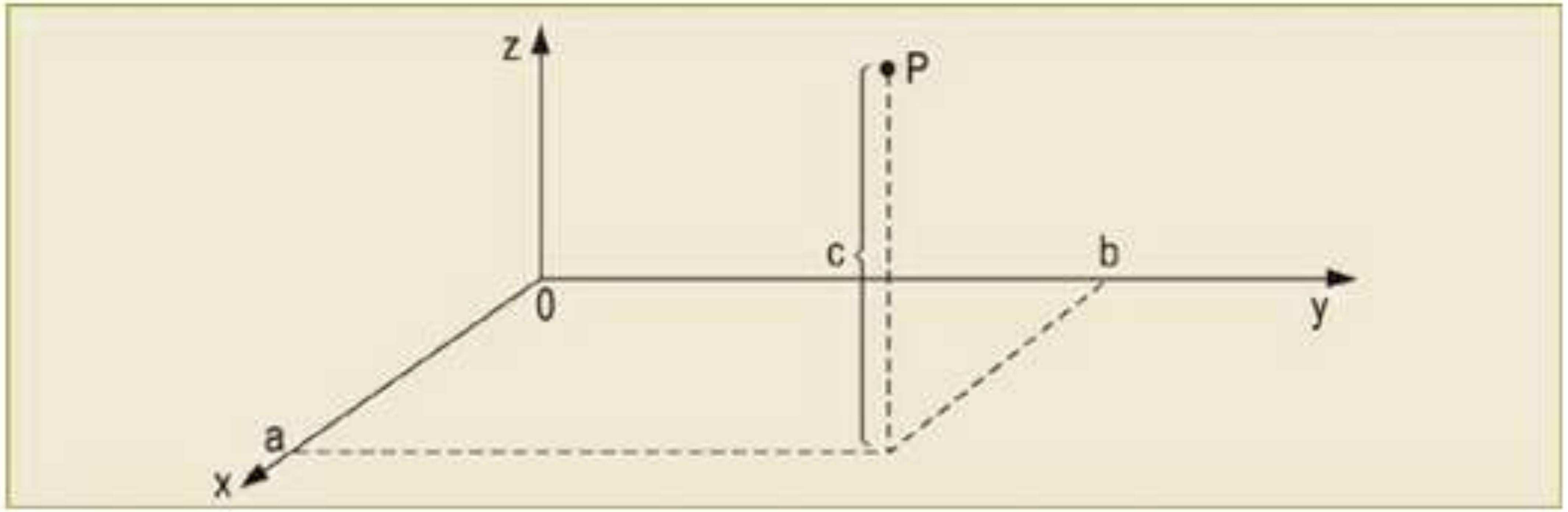
O Espaço Tridimensional

Seja R o conjunto dos números reais. O conjunto formado por todas as triplas ordenadas de números reais é chamado espaço tridimensional e é indicado por R × R × R ou simplesmente por R3. Assim:

$$R^3 = \{(a, b, c) \mid a \in R, b \in R, c \in R\}$$

$$(2, 4, 5), (3, -1, 3), (\frac{1}{2}, \frac{1}{3}, 0)$$

O Espaço Tridimensional



Geometricamente, um elemento (a, b, c) do R3 pode ser representado por um ponto P de abscissa a, ordenada b e cota c, em um sistema de eixos x, y e z perpendiculares dois a dois.

Matemática para Machine Learning

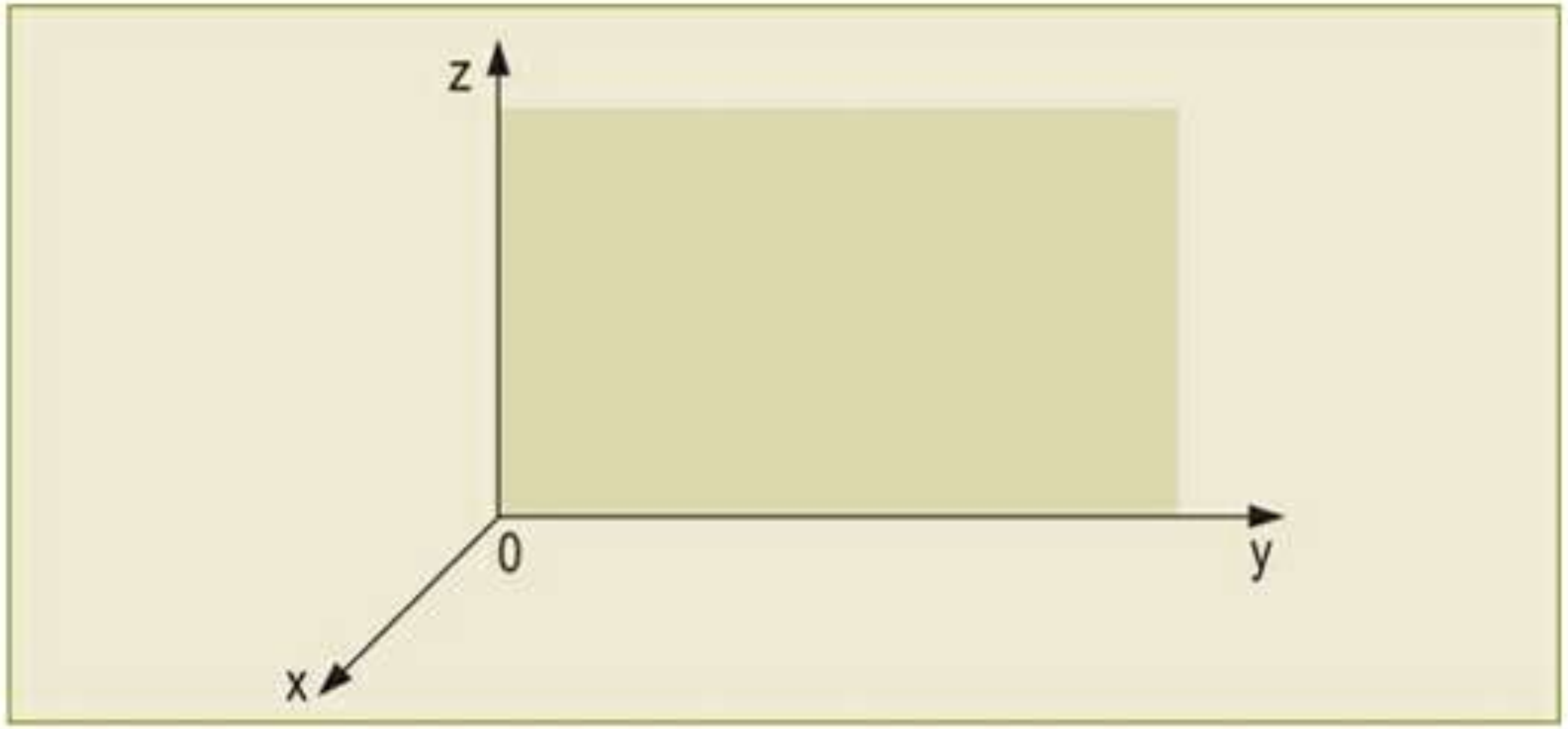
Relações em R3

Relações em R3

Chama-se relação em R3 a todo subconjunto do R3.

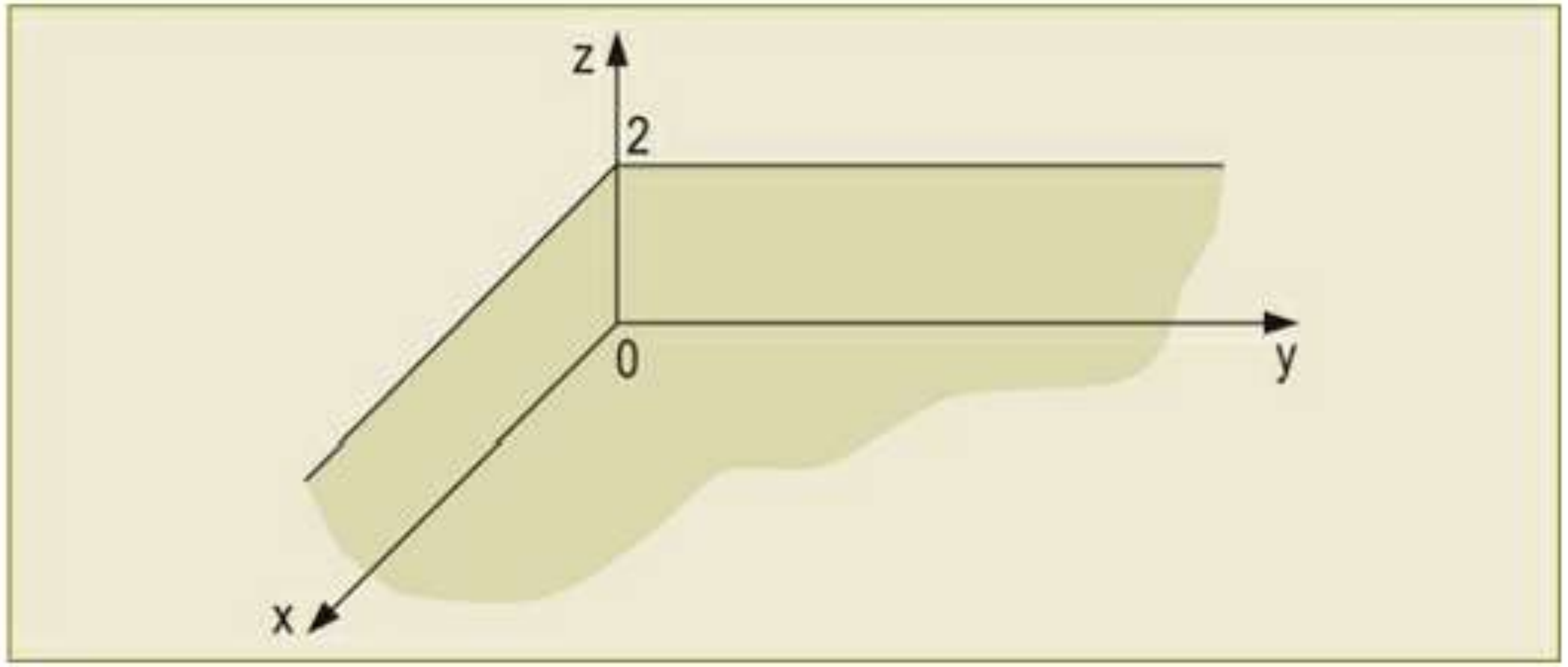
Relações em R3

$$A = \{(x, y, z) \mid x = 0\}$$



Relações em R3

$$B = \{(x, y, z) \mid z = 2\}$$



Matemática para Machine Learning

Derivadas Para Funções de 2 Variáveis

Data Science Academy

Matemática para Machine Learning

Função Derivada Parcial

Data Science Academy

Função Derivada Parcial

Se calcularmos fx e fy num ponto genérico (x, y), obteremos duas funções de x e y; a função fx(x, y) é chamada **função derivada parcial de f em relação a x** (ou, simplesmente, derivada parcial de f em relação a x). A função fy(x, y) é chamada **função derivada parcial de f em relação a y** (ou, simplesmente, derivada parcial de f em relação a y).

As derivadas parciais também podem ser indicadas por:

$$f_x \text{ ou } \frac{\partial f}{\partial x} \quad \text{e} \quad f_y \text{ ou } \frac{\partial f}{\partial y}$$

Função Derivada Parcial

Para o cálculo de fx e fy, podemos aplicar as regras de derivação estudadas em funções de uma variável no capítulo anterior, desde que:

a) No cálculo de fx consideremos y como constante.
b) No cálculo de fy consideremos x como constante.

$$f_x \text{ ou } \frac{\partial f}{\partial x} \quad \text{e} \quad f_y \text{ ou } \frac{\partial f}{\partial y}$$

Matemática para Machine Learning

Diferencial de Uma Função

Diferencial de Uma Função

Seja f uma função com duas variáveis. Se as derivadas parciais são contínuas em um conjunto aberto A, então f é diferenciável em todos os pontos de A.

Diferencial de Uma Função

Consideremos a função dada por f(x, y) = 2x^2 + 3y^2 e calculemos a variação Δf sofrida pela função quando x e y apresentam variações Δx e Δy a partir do ponto (x0, y0).

$$\Delta f = f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0, y_0)$$

$$= 2(x_0 + \Delta x)^2 + 3(y_0 + \Delta y)^2 - (2x_0^2 + 3y_0^2)$$

$$= 2(x_0^2 + 2x_0\Delta x + \Delta x^2) + 3(y_0^2 + 2y_0\Delta y + \Delta y^2) - 2x_0^2 - 3y_0^2$$

$$\boxed{= 4x_0\Delta x + 6y_0\Delta y + 2\Delta x^2 + 3\Delta y^2}$$

Diferencial de Uma Função

$$\Delta f = f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0, y_0)$$

$$= 2(x_0 + \Delta x)^2 + 3(y_0 + \Delta y)^2 - (2x_0^2 + 3y_0^2)$$

$$= 2(x_0^2 + 2x_0\Delta x + \Delta x^2) + 3(y_0^2 + 2y_0\Delta y + \Delta y^2) - 2x_0^2 - 3y_0^2$$

$$\boxed{= 4x_0\Delta x + 6y_0\Delta y + 2\Delta x^2 + 3\Delta y^2}$$

Por exemplo, se x0 = 5, y0 = 6 e Δx = Δy = 0,01, teremos:

$$\Delta f = 4 . (5) . 0{,}01 + 6 . (6) . 0{,}01 + 2(0{,}01)^2 + 3(0{,}01)^2$$

$$= 0{,}2 + 0{,}36 + 0{,}0002 + 0{,}0003$$

Diferencial de Uma Função

Por exemplo, se x0 = 5, y0 = 6 e Δx = Δy = 0,01, teremos:

$$\Delta f = 4 \,.\, (5) \,.\, 0{,}01 + 6 \,.\, (6) \,.\, 0{,}01 + 2(0{,}01)^2 + 3(0{,}01)^2$$

$$= 0{,}2 + 0{,}36 + 0{,}0002 + 0{,}0003$$

Como as parcelas 0,0002 e 0,0003 são desprezíveis comparadas com 0,2 e 0,36, podemos dizer que:

$$\mathbf{\Delta f \cong 0{,}2 + 0{,}36 = 0{,}56}$$

Diferencial de Uma Função

Voltando à expressão de Δf notamos que:

$$4x_0 = \frac{\partial f}{\partial x}(x_0, y_0) \text{ e } 6y_0 = \frac{\partial f}{\partial y}(x_0, y_0)$$

os termos 2Δx^2 + 3Δy^2 são desprezíveis quando comparados com 4x0Δx + 6y0Δy desde que Δx e Δy sejam próximos de zero;

Diferencial de Uma Função

O resultado que acabamos de ver não é um caso isolado, mas vale para a maioria das funções; isto é, a variação sofrida por f(x, y) quando variamos simultaneamente x e y de valores pequenos Δx e Δy é aproximadamente igual a:

$$\Delta f \cong \frac{\partial f}{\partial x}(x_0, y_0)\,\Delta x + \frac{\partial f}{\partial y}(x_0, y_0)\,\Delta y$$

Diferencial de Uma Função

Esse exemplo preliminar nos leva à seguinte definição: Seja f uma função com duas variáveis e seja (x0, y0) um ponto de seu domínio. Seja Δf a variação sofrida por f(x, y) ao passarmos do ponto (x0, y0) para o ponto (x0 + Δx, y0 + Δy). Isto é, Δf = f(x0 + Δx, y0 + Δy) − f (x0, y0).

Dizemos que f é diferenciável no ponto (x0, y0) se Δf puder ser escrita sob a forma:

$$\Delta f = \frac{\partial f}{\partial x}(x_0, y_0) \,.\, \Delta x + \frac{\partial f}{\partial y}(x_0, y_0) \,.\, \Delta y + \Delta x \,.\, h_1(\Delta x, \Delta y) + \Delta y \,.\, h_2(\Delta x, \Delta y)$$

onde as funções h1 e h2 têm limites iguais a zero quando (Δx, Δy) tende a (0,0).

Diferencial de Uma Função

A parcela:

$$\Delta f \cong \frac{\partial f}{\partial x}(x_0, y_0)\,\Delta x + \frac{\partial f}{\partial y}(x_0, y_0)\,\Delta y$$

é chamada **diferencial de f** e é indicada por df, no caso de f ser diferenciável.

Diferencial de Uma Função

Seria bastante trabalhoso termos de verificar pela definição se uma função é ou não diferenciável, para calcularmos a diferencial como resultado aproximado de Δf. Felizmente, a maioria das funções é diferenciável.

Existe um teorema que nos fornece condições facilmente verificáveis para vermos se uma função é diferenciável. Seu enunciado é o seguinte:

Seja f uma função com duas variáveis. Se as derivadas parciais são contínuas em um conjunto aberto A, então f é diferenciável em todos os pontos de A.

Matemática para Machine Learning

Função Composta - Regra da Cadeia

Função Composta - Regra da Cadeia

Consideremos uma função usada em uma fábrica:

$$P(x, y) = 6x^{0,5}y^{0,5}$$

em que x e y são as quantidades de dois insumos, capital e trabalho, e P, a quantidade produzida de um produto.

Suponhamos que o capital x cresça com o tempo t, de acordo com a relação x = 0,16t, e que o trabalho cresça de acordo com a relação y = 0,09t.

Função Composta - Regra da Cadeia

Se quisermos expressar a produção em função do tempo, temos que substituir x = 0,16t e y = 0,09t na relação:

$$P(x, y) = 6x^{0,5}y^{0,5}$$

$$P(t) = 6(0,16t)^{0,5}(0,09t)^{0,5} = 0,72t$$

A função de t, dada por P(t) = 0,72t, chamamos de **função composta** de P com x e y.

A derivada da função composta dada por P(t) = 0,72t em relação a t é imediata (função de uma variável):

$$P'(t) = \frac{dP}{dt} = 0,72$$

Função Composta - Regra da Cadeia

Seja f uma função de duas variáveis x e y, diferenciável num ponto (x_0, y_0) do domínio, e sejam as funções dadas por $x(t)$ e $y(t)$ diferenciáveis em t_0, de modo que $x(t_0) = x_0$ e $y(t_0) = y_0$. Então a função F composta de f com x e y é tal que:

$$\frac{dF}{dt}(t_0) = \frac{\partial f}{\partial x}(x_0, y_0) \cdot \frac{dx}{dt}(t_0) + \frac{\partial f}{\partial y}(x_0, y_0) \cdot \frac{dy}{dt}(t_0)$$

Ou abreviadamente:

$$\frac{dF}{dt} = \frac{\partial f}{\partial x} \cdot \frac{dx}{dt} + \frac{\partial f}{\partial y} \cdot \frac{dy}{dt}$$

Função Composta - Regra da Cadeia

De forma prática a regra da cadeia se faz derivando a função que esta de fora (f'(g(x)) multiplicada pela função de dentro derivada (g'(x)).

$$y' = f'(g(x))\, g'(x)$$

ou

$$\frac{dy}{dx} = \frac{dy}{du}\frac{du}{dx}$$

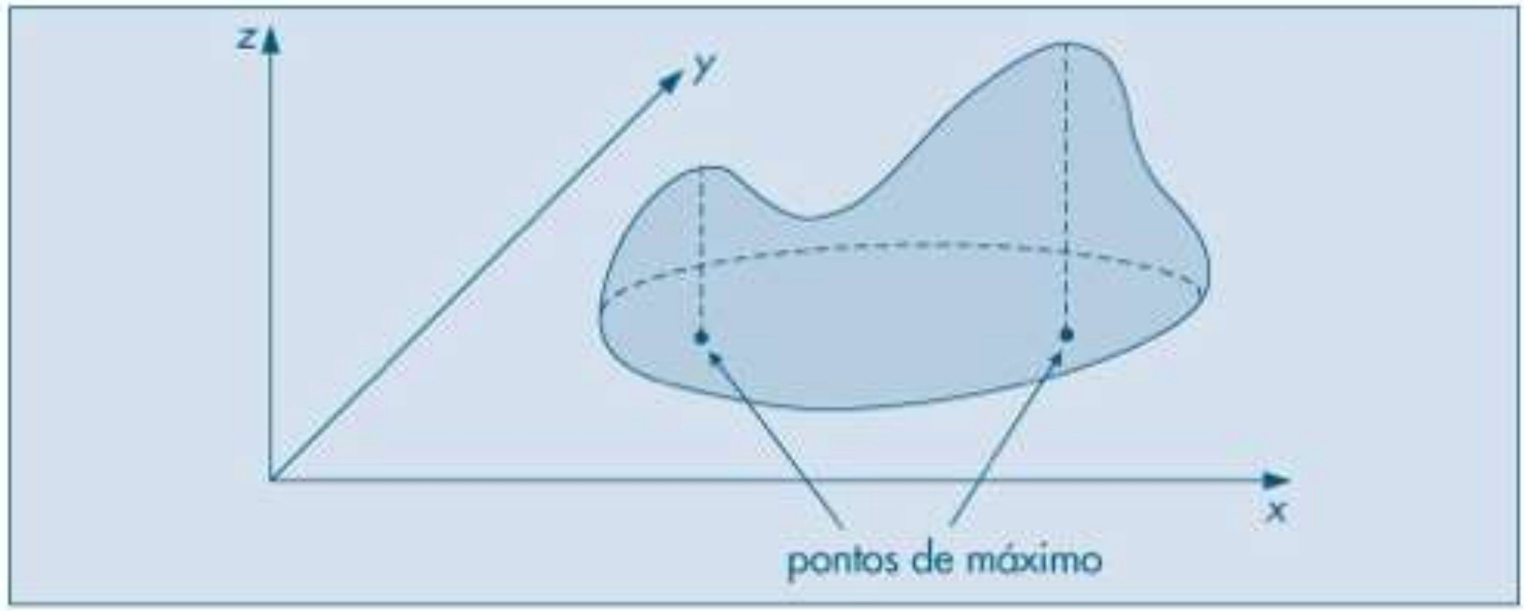
Matemática para Machine Learning

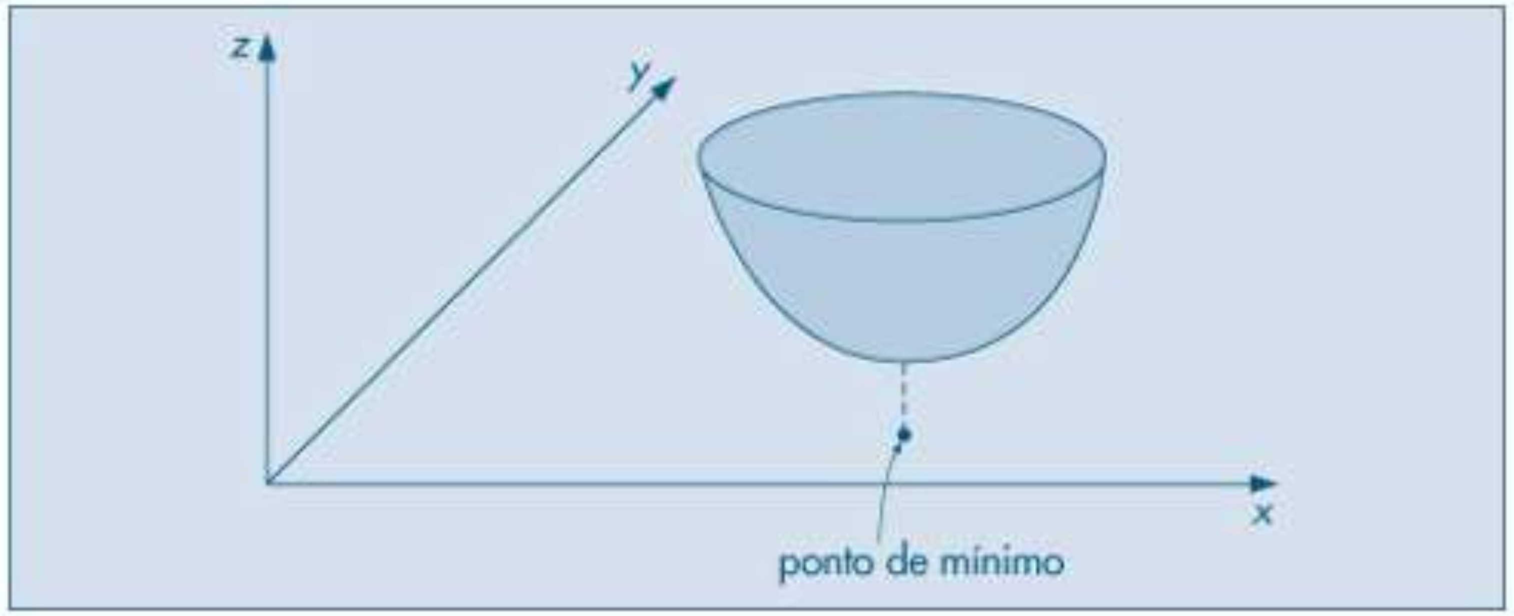
Máximos e Mínimos Para Funções de Duas Variáveis

Data Science Academy

Máximos e Mínimos Para Funções de Duas Variáveis

Uma importante aplicação do estudo das derivadas parciais é a da otimização de funções. Otimizar uma função significa encontrar seu ponto de máximo ou de mínimo.





Máximos e Mínimos Para Funções de Duas Variáveis

Como determinamos os pontos máximo e mínimo?

Seja f uma função de duas variáveis x e y, contínua, com derivadas parciais até segunda ordem contínuas. Seja (x0, y0) um ponto crítico de f (os pontos que anulam simultaneamente as derivadas parciais fx e fy). Chamemos o determinante:

$$H(x_0, y_0) = \begin{vmatrix} f_{xx}(x_0, y_0) & f_{xy}(x_0, y_0) \\ f_{yx}(x_0, y_0) & f_{yy}(x_0, y_0) \end{vmatrix}$$

Máximos e Mínimos Para Funções de Duas Variáveis

$$H(x_0, y_0) = \begin{vmatrix} f_{xx}(x_0, y_0) & f_{xy}(x_0, y_0) \\ f_{yx}(x_0, y_0) & f_{yy}(x_0, y_0) \end{vmatrix}$$

H(x0, y0) > 0 e fxx(x0, y0) < 0, então (x0, y0) será ponto de máximo de f.

H(x0, y0) > 0 e fxx(x0, y0) > 0, então (x0, y0) será ponto de mínimo de f.

Máximos e Mínimos Para Funções de Duas Variáveis

Consideremos a função f(x, y) = $x^2 + y^2 - 2x$

Os pontos críticos de f são soluções do sistema: $\begin{cases} f_x = 2x - 2 = 0 \\ f_y = 2y = 0 \end{cases}$

ou seja, x = 1 e y = 0. Portanto, (1, 0) é o único ponto crítico.

Por outro lado: $\begin{cases} f_{xx} = 2 \\ f_{xy} = 0 \\ f_{yx} = 0 \\ f_{yy} = 2 \end{cases} \Rightarrow \begin{cases} f_{xx}(1,0) = 2 \\ f_{xy}(1,0) = 0 \\ f_{yx}(1,0) = 0 \\ f_{yy}(1,0) = 2 \end{cases}$ $H(1,0) = \begin{vmatrix} 2 & 0 \\ 0 & 2 \end{vmatrix} = 4$ $\begin{cases} H(1,0) > 0 \\ \text{e} \\ f_{xx}(1,0) = 2 > 0 \end{cases} \Rightarrow (1,0)$ é ponto de mínimo de f.

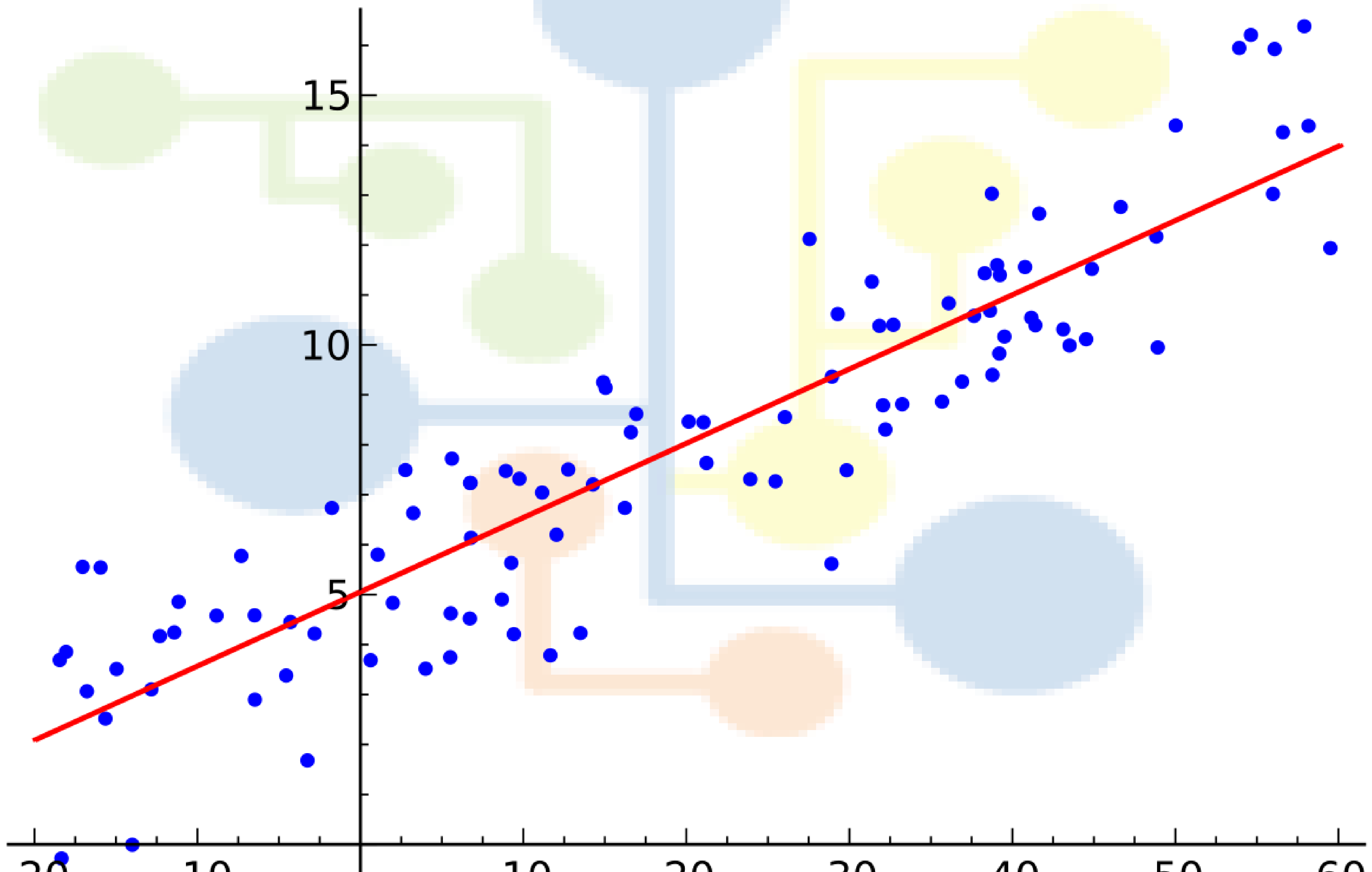
Matemática para Machine Learning

Aplicação - Ajuste de Retas Pelo Método dos Mínimos Quadrados

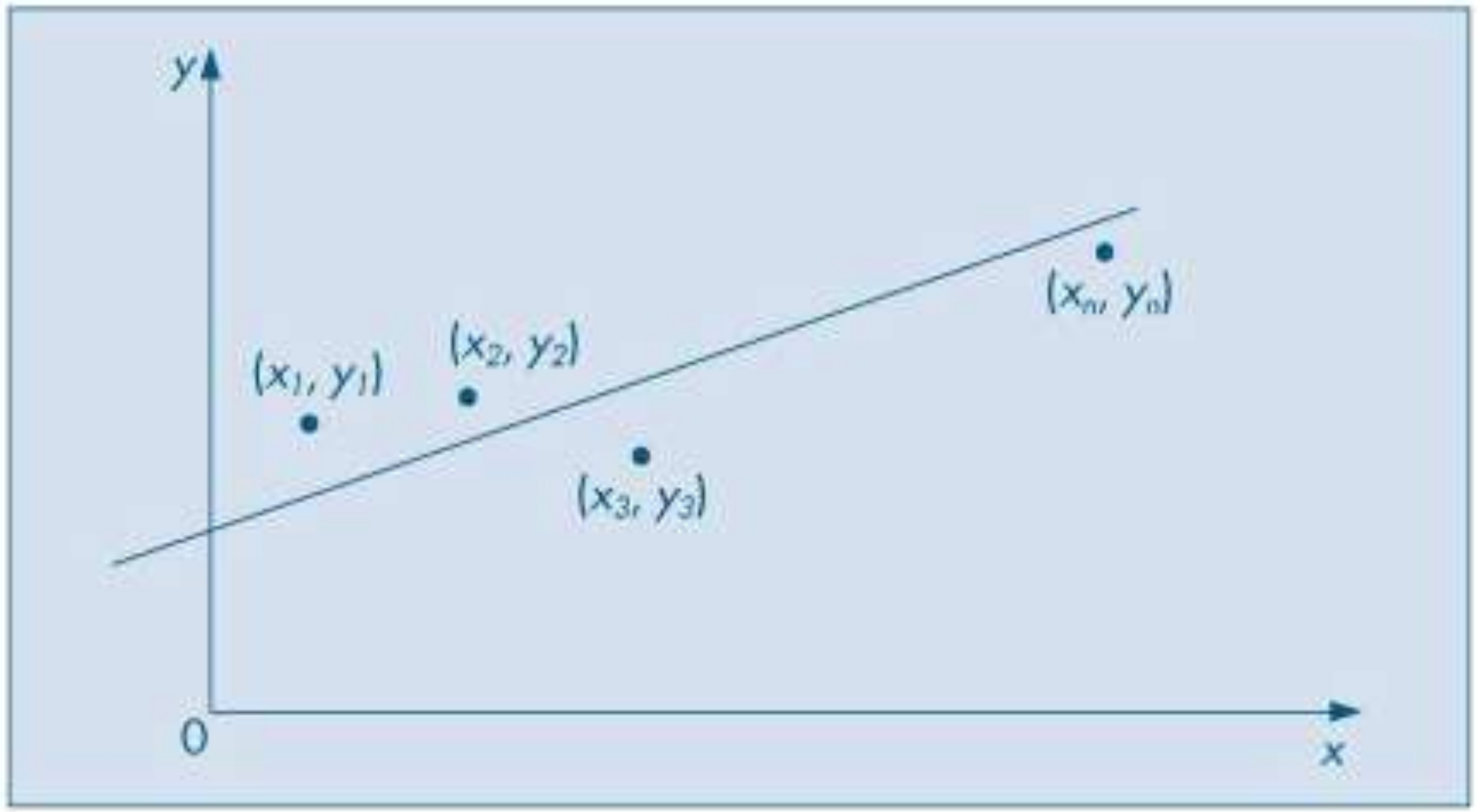
Data Science Academy

Aplicação - Ajuste de Retas Pelo Método dos Mínimos Quadrados

Aplicação - Ajuste de Retas Pelo Método dos Mínimos Quadrados



Aplicação - Ajuste de Retas Pelo Método dos Mínimos Quadrados

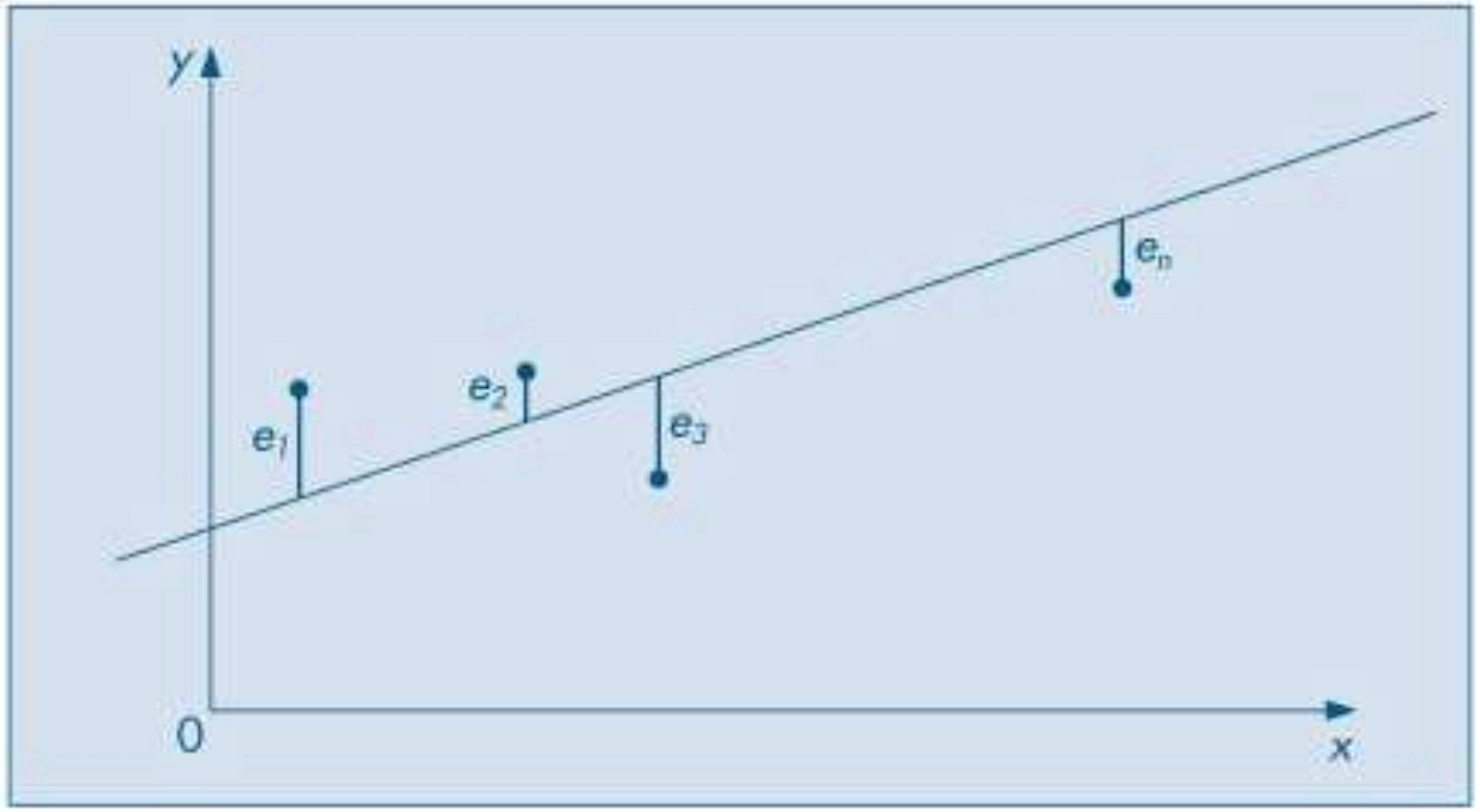


Para obter a reta, podemos usar o Método dos Mínimos Quadrados.

Definição:

Entre as infinitas retas que existem, uma delas, de equação y = ax + b, será aquela que tornará mínima a soma dos quadrados dos desvios

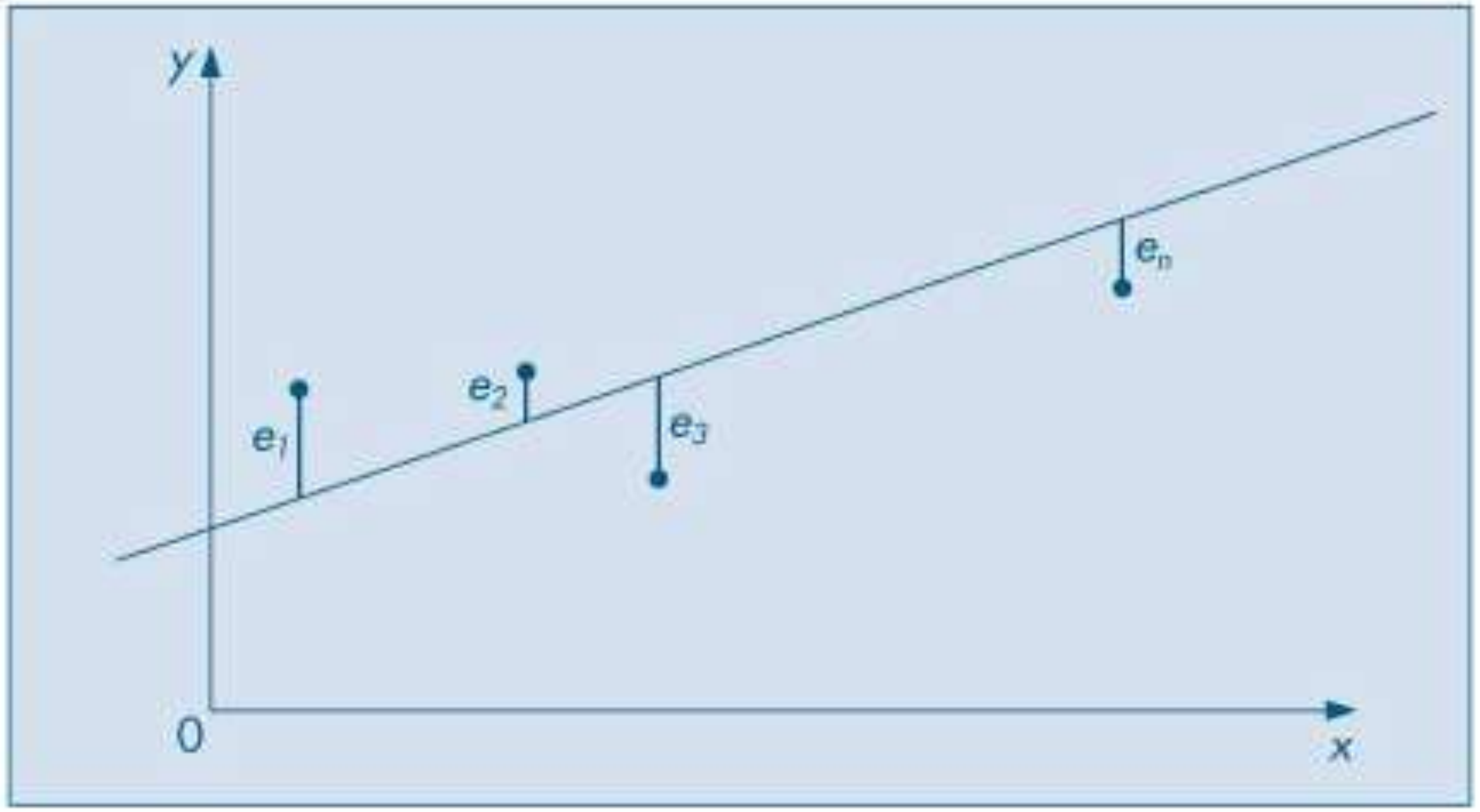
Aplicação - Ajuste de Retas Pelo Método dos Mínimos Quadrados



Entre as infinitas retas que existem, uma delas, de equação y = ax + b, será aquela que tornará mínima a soma dos quadrados dos desvios

$$f(a,b) = \sum_{i=1}^{n} (y_i - ax_i - b)^2$$

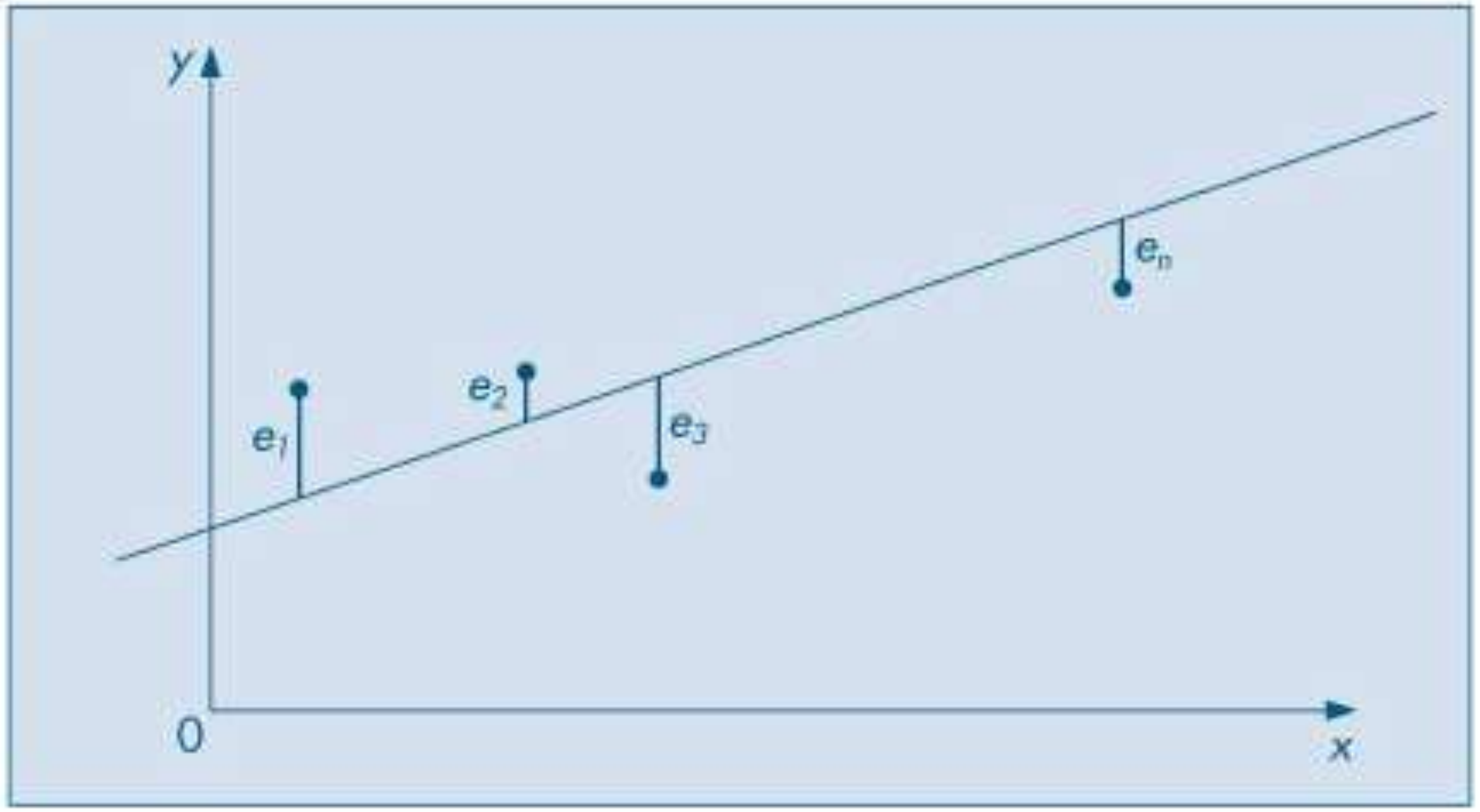
Aplicação - Ajuste de Retas Pelo Método dos Mínimos Quadrados



A reta é chamada de Reta dos Mínimos Quadrados.

$$f(a, b) = \sum_{i=1}^{n} (y_i - ax_i - b)^2$$

Aplicação - Ajuste de Retas Pelo Método dos Mínimos Quadrados



Nosso problema consiste em encontrar o par (a, b) que minimiza f(a, b).

$$f(a,b) = \sum_{i=1}^{n} (y_i - ax_i - b)^2$$

Aplicação - Ajuste de Retas Pelo Método dos Mínimos Quadrados

$$a = \frac{\sum x_i y_i - \frac{(\sum x_i)(\sum y_i)}{n}}{\sum x_i^2 - \frac{(\sum x_i)^2}{n}}$$

e

$$b = \frac{\sum y_i}{n} - a\frac{\sum x_i}{n}$$

Nosso problema consiste em encontrar o par (a, b) que minimiza f(a, b).

$$f(a,b) = \sum_{i=1}^{n} (y_i - ax_i - b)^2$$

Aplicação - Ajuste de Retas Pelo Método dos Mínimos Quadrados

Um empresário deseja obter empiricamente uma equação de demanda para seu produto. Ele admite que a quantidade média demandada (y) relaciona-se com seu preço unitário (x) por meio de uma função do 1o grau. Para estimar essa reta, ele fixou os preços em vários níveis e observou a quantidade demandada, obtendo os dados a seguir:

Preço unitário (x)	Quantidade demandada (y)
1	45
2	43
3	35
4	33
5	30
6	21
7	12

Aplicação - Ajuste de Retas Pelo Método dos Mínimos Quadrados

$$a = \frac{\sum x_i y_i - \frac{(\sum x_i)(\sum y_i)}{n}}{\sum x_i^2 - \frac{(\sum x_i)^2}{n}}$$

e

$$b = \frac{\sum y_i}{n} - a\frac{\sum x_i}{n}$$

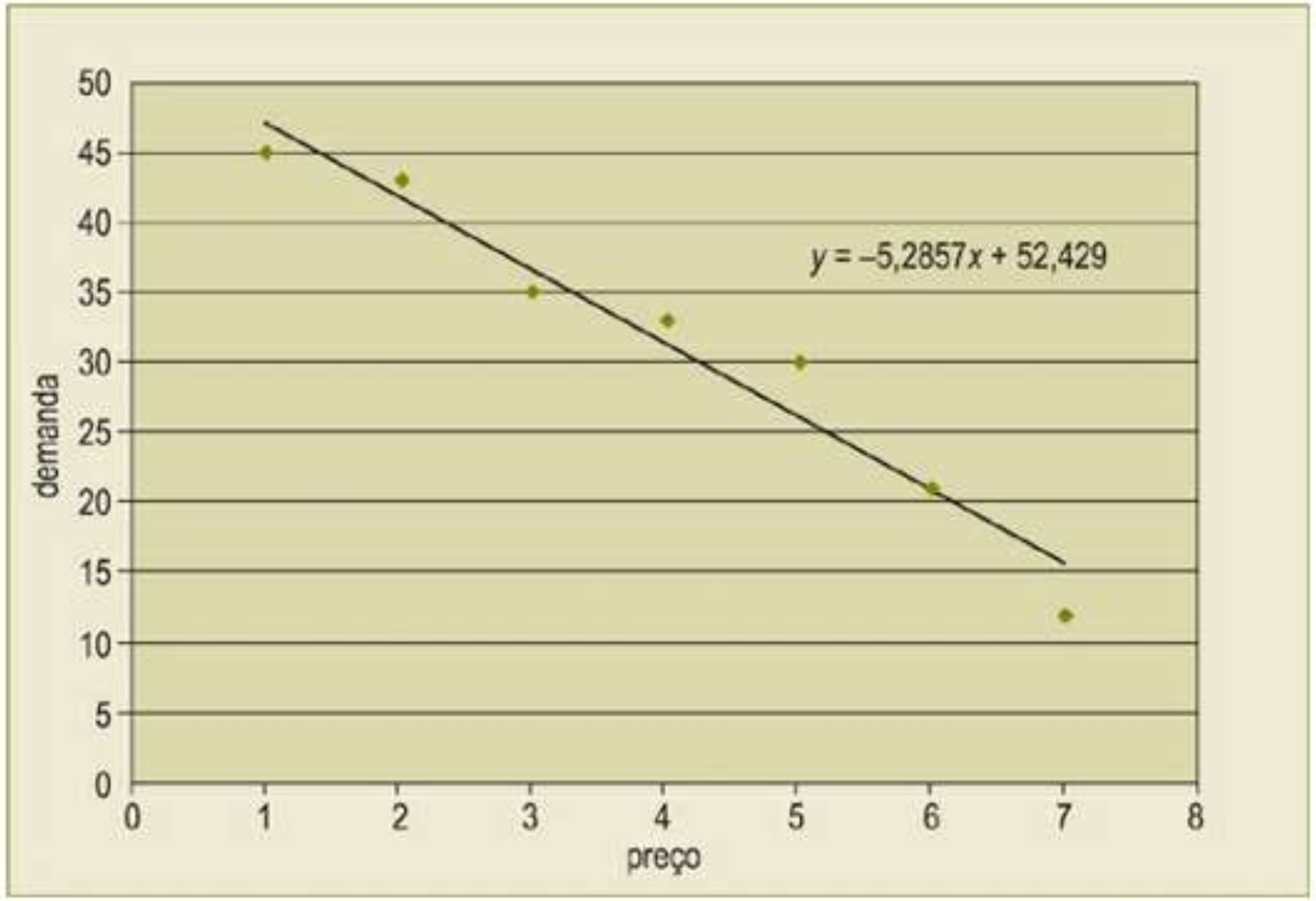
	x_i	y_i	$x_i y_i$	x_i^2
	1	45	45	1
	2	43	86	4
	3	35	105	9
	4	33	132	16
	5	30	150	25
	6	21	126	36
	7	12	84	49
$\sum$	**28**	**219**	**728**	**140**

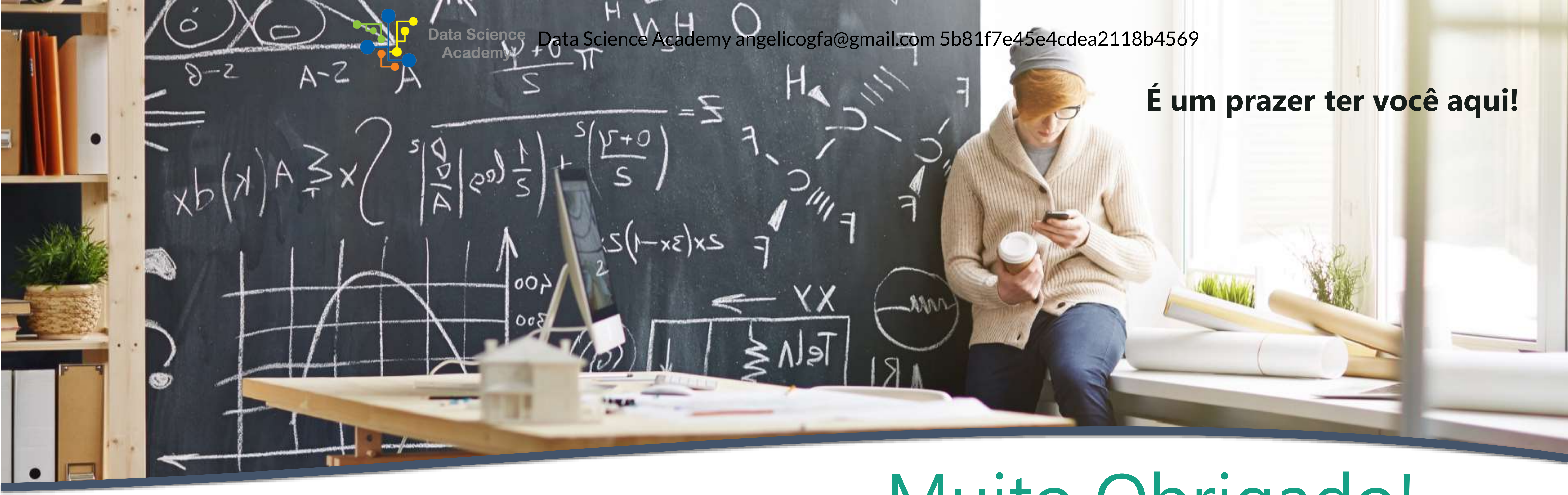
Aplicação - Ajuste de Retas Pelo Método dos Mínimos Quadrados

$$a = \frac{728 - \frac{(28)(219)}{7}}{140 - \frac{(28)^2}{7}} = \frac{-148}{28} = -5,2857$$

$$b = \frac{219}{7} - (-5,2857).\frac{28}{7} = 52,4285$$

Portanto, a equação da reta procurada é y = −5,2857x + 52,4285.





É um prazer ter você aqui!

Muito Obrigado!

Pela Confiança em Nosso Trabalho.

Continue Trilhando Uma Excelente Jornada de Aprendizagem!

Data Science Academy