

UNIVERSIDADE FEDERAL DO NORTE DO TOCANTINS
CAMPUS DE ARAGUAÍNA
CURSO DE LICENCIATURA EM MATEMÁTICA

Luana Ferreira Plácido

O MÉTODO DOS MÍNIMOS QUADRADOS NA MODELAGEM DE FENÔMENOS

ARAGUAÍNA - TO
2026

Luana Ferreira Plácido

O MÉTODO DOS MÍNIMOS QUADRADOS NA MODELAGEM DE FENÔMENOS

Monografia apresentada ao curso de Licenciatura em Matemática da Universidade Federal do Norte do Tocantins (UFNT), como requisito parcial para a obtenção de título de Licenciada em Matemática.

Orientador: Prof. Dr. José Carlos de Oliveira Junior.

Dados Internacionais de Catalogação na Publicação (CIP)
Sistema de Geração de Ficha Catalográfica SGFC-UFNT
Gerado automaticamente mediante os dados fornecidos pelo(a) autor(a)

F383m Ferreira Plácido, Luana .

O Método dos Mínimos Quadrados na Modelagem de Fenômenos /
Luana Ferreira Plácido. - Centro de Ciências Integradas - CCI, TO,
2026.

52 f.

Monografia Graduação (Graduação - em Matemática) --
Universidade Federal do Norte do Tocantins, 2026.

Orientador: Profº Dr. José Carlos de Oliveira Junior.

1. Método dos Mínimos Quadrados. 2. Ajuste de Curvas. 3.
Modelagem Matemática.

CDD 510

TODOS OS DIREITOS RESERVADOS – A reprodução total ou parcial, de qualquer forma ou por qualquer meio deste documento é autorizado desde que citada a fonte. A violação dos direitos do autor (Lei nº 9.610/98) é crime estabelecido pelo artigo 184 do Código Penal.

Luana Ferreira Plácido

O Método dos Mínimos Quadrados na Modelagem de Fenômenos

Monografia apresentada à Universidade Federal do Norte do Tocantins (UFNT), Centro de Ciências Integradas (CCI), curso de Licenciatura em Matemática. Foi avaliada para a obtenção do título de Licenciada em Matemática e aprovada em sua forma final pelo orientador e pela banca examinadora.

Data de aprovação: 06 / 07 / 2026

Banca examinadora:

Prof. Dr. José Carlos de Oliveira Junior - Orientador (UFNT)

Prof. Dr. Alvaro Julio Yucra Hancoco - Membro interno (UFNT)

Prof. Me. André Luiz Ortiz da Silva - Membro interno (UFNT)

Dedico este trabalho aos que um dia foram meus professores de Matemática e me ensinaram a
ver a beleza que existe nessa ciência.

AGRADECIMENTOS

Nenhuma conquista se faz sozinha. Por isso, mais do que agradecer, quero registrar aqui o carinho de cada um que esteve comigo nessa caminhada.

Agradeço primeiramente a Deus, por ter me dado forças, coragem e fé para chegar até aqui. À minha família, que sempre foi meu porto seguro e minha base, e que me ajudou da forma como pôde em cada etapa desse percurso. Aos meus irmãos, que sempre me torceram e incentivaram.

Sou profundamente grata ao meu orientador, professor Dr. José Carlos de Oliveira Junior, pela paciência, pela bondade e pela disposição em me ensinar desde o nosso projeto de iniciação científica (PIVIC) até esta etapa de tanta importância na minha vida acadêmica. Estendo essa gratidão aos professores do curso, que me ensinaram do mais básico ao mais complexo com excelência.

Ao meu noivo, Rodrigo Alessandro, que desde sempre esteve ao meu lado e nunca desistiu de mim, fazendo com que eu também não desistisse de mim mesma. Obrigada por cada palavra de incentivo nos dias difíceis, por acreditar em mim quando eu mesma duvidei e por segurar minha mão em cada passo desse processo. Hoje você se tornou o meu refúgio, e sua presença fez tudo ficar mais leve e essa conquista mais doce.

Por fim, mas não menos importante, agradeço aos meus amigos Ana Laura, Francisco Uchoa, Sabrina e Tanielly, que desde o primeiro período me acolheram e fizeram desses quatro anos inesquecíveis e divertidos. Mesmo com a individualidade de cada um, formamos um laço tão forte que hoje são pessoas que levarei não apenas do curso, mas para a vida toda.

De um modo geral, sou grata a todos que, de alguma forma, fizeram parte dessa história.

A mente que se abre a uma nova ideia jamais volta ao seu tamanho original.

(Albert Einstein)

RESUMO

Este trabalho tem por objetivo estudar o Método dos Mínimos Quadrados (MMQ) e aplicá-lo à modelagem matemática de fenômenos reais, com ênfase na interpretação didática dos resultados. A pesquisa é de natureza bibliográfica e exploratória, parte-se da revisão dos fundamentos do cálculo diferencial necessários à compreensão do método, deduz-se analiticamente o critério de minimização da soma dos quadrados dos resíduos e, em seguida, aplicam-se as fórmulas obtidas a conjuntos de dados reais. No ajuste linear, modelam-se a deformação de uma mola submetida a forças crescentes e a produção agrícola em função do tempo, evidenciando a coerência dos coeficientes com as leis físicas subjacentes. No ajuste exponencial, utilizam-se os registros semanais de casos novos de COVID-19 no Brasil nas semanas epidemiológicas 13 a 29 de 2020 para calibrar o modelo e projetar o comportamento da semana 30, caso o crescimento não tivesse se alterado. Os resultados mostram que o MMQ converte dados dispersos em expressões matemáticas interpretáveis, capazes de quantificar tendências e embasar projeções de curto prazo. Discutem-se ainda os limites do modelo exponencial, válido apenas na fase inicial de crescimento acelerado, e apresentam-se, em caráter introdutório, outros modelos não lineares (logístico e de potência) cuja escolha depende da estrutura do fenômeno observado. O trabalho nasceu de um projeto de iniciação científica (PIVIC/UFNT) e destina-se a estudantes e professores de licenciatura em Matemática que desejam compreender e aplicar o MMQ de forma rigorosa e acessível.

Palavras-chave: Método dos Mínimos Quadrados. Ajuste de curvas. Modelagem matemática. Regressão linear. COVID-19. Ajuste exponencial.

ABSTRACT

This work aims to study the Least Squares Method (LSM) and apply it to the mathematical modeling of real phenomena, with emphasis on the didactic interpretation of results. The research is bibliographic and exploratory in nature, it begins with a review of the differential calculus foundations required to understand the method, analytically derives the criterion for minimizing the sum of squared residuals, and then applies the obtained formulas to real datasets. In linear fitting, the deformation of a spring under increasing forces and agricultural production as a function of time are modeled, highlighting the coherence of the coefficients with the underlying physical laws. In exponential fitting, weekly records of new COVID-19 cases in Brazil during epidemiological weeks 13 to 29 of 2020 are used to calibrate the model and project the behavior of week 30, had growth remained unchanged. The results show that the LSM converts scattered data into interpretable mathematical expressions capable of quantifying trends and supporting short-term projections. The limits of the exponential model, valid only during the initial phase of accelerated growth, are also discussed, and other non-linear models (logistic and power) are briefly introduced, emphasizing that model selection must match the mathematical structure of the observed phenomenon. This work originated from an undergraduate research project (PIVIC/UFNT) and is intended for students and professors of Mathematics teacher education programs who wish to understand and apply the LSM in a rigorous and accessible manner.

Keywords: Least Squares Method. Curve fitting. Mathematical modeling. Linear regression. COVID-19. Exponential fitting.

Sumário

1	INTRODUÇÃO	11
2	NOÇÕES PRELIMINARES	13
2.1	Máximos e Mínimos para Funções de uma Variável	13
2.2	Máximos e Mínimos para Funções de duas Variáveis.	19
2.2.1	Otimização	21
3	MOTIVAÇÃO PARA O MÉTODO DOS MÍNIMOS QUADRADOS	26
3.1	Primeira tentativa: minimizar a soma dos resíduos	27
3.2	Segunda tentativa: minimizar a soma dos módulos.	29
3.3	Terceira tentativa: minimizar a soma dos quadrados	30
3.3.1	Balanco dos três critérios	31
3.4	A formulação do método	32
4	O MÉTODO EM AÇÃO: AJUSTES LINEAR E EXPONENCIAL	34
4.1	Ajuste Linear	34
4.1.1	As equações normais	35
4.1.2	Exemplos resolvidos	37
4.2	Ajuste Exponencial	40
4.2.1	Aplicação: COVID-19 no Brasil, semanas 13 a 29 de 2020	42
4.2.2	Outros modelos não lineares: ajustes logístico e de potência	46
5	CONSIDERAÇÕES FINAIS	49
	REFERÊNCIAS	51

Capítulo 1

Introdução

A modelagem matemática traduz fenômenos do mundo real em linguagem formal, de modo a descrevê-los, compreendê-los e, quando as condições se mantêm, prever seu comportamento. Entre os instrumentos voltados a esse fim, o Método dos Mínimos Quadrados (MMQ) ocupa posição central na modelagem de dados experimentais, por conciliar fundamentação matemática precisa e resultados de interpretação direta [5, 8]. Desenvolvido no início do século XIX de forma independente por Adrien-Marie Legendre, que o publicou formalmente em 1805, e por Carl Friedrich Gauss, que reivindicou seu emprego anterior em estudos astronômicos de determinação de órbitas, o método consolidou-se como referência para o ajuste de curvas a observações sujeitas a erro.

A formulação do método parte de um conjunto de pares de dados (x_i, y_i) , com $i = 1, \dots, n$, aos quais se deseja ajustar um modelo. No caso de uma reta $y = a + bx$, define-se a função de erro quadrático

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

em que a e b são os parâmetros a determinar e cada parcela mede o quadrado do resíduo $r_i = y_i - (a + bx_i)$, isto é, a diferença entre o valor observado e o valor previsto pelo modelo. Os melhores parâmetros são aqueles que tornam S mínima; como S é uma função diferenciável de a e b , a condição necessária de mínimo impõe gradiente nulo,

$$\frac{\partial S}{\partial a} = 0 \quad \text{e} \quad \frac{\partial S}{\partial b} = 0,$$

o que origina um sistema de duas equações lineares, as chamadas equações normais, cuja solução fornece a e b . O mesmo princípio se estende a modelos não lineares por meio de transformações adequadas, como no caso do modelo exponencial $y = Ae^{\beta x}$, por exemplo, que se converte em uma relação linear ao se aplicar o logaritmo natural, $\ln y = \ln A + \beta x$, permitindo reutilizar as mesmas fórmulas sobre os dados transformados.

Este trabalho originou-se de um projeto de iniciação científica (PIVIC/UFNT) dedicado

ao uso do MMQ na modelagem de fenômenos reais. A partir dessa experiência, formulou-se a seguinte questão norteadora: “De que modo o Método dos Mínimos Quadrados pode ser utilizado para modelar matematicamente fenômenos reais, e quais são os alcances e os limites desses modelos?” Para respondê-la, define-se como objetivo geral estudar o MMQ e aplicá-lo à modelagem de fenômenos reais, conciliando rigor teórico e acessibilidade didática. Como objetivos específicos, busca-se (i) revisar os fundamentos do cálculo diferencial que embasam o método; (ii) deduzir o critério de minimização e o sistema de equações normais dele decorrente; (iii) aplicar as fórmulas obtidas a dados reais nos casos linear e exponencial; e (iv) discutir os alcances e os limites dos modelos ajustados, indicando outros modelos não lineares pertinentes.

Quanto à metodologia, a pesquisa caracteriza-se como bibliográfica e exploratória, de abordagem qualitativa voltada à compreensão e à interpretação dos modelos obtidos. O referencial teórico apoia-se em obras consolidadas de cálculo diferencial e de ajuste de curvas [3, 4, 5, 8], e as aplicações utilizam dados reais de fontes oficiais, em particular os boletins do Ministério da Saúde [10] e as projeções populacionais do Instituto Brasileiro de Geografia e Estatística [9].

Sobre essa base, ao longo do trabalho serão revisados os conceitos de máximo e mínimo, ponto crítico, Teorema de Fermat, gradiente e Teste da Segunda Derivada; será deduzido analiticamente o critério dos mínimos quadrados com as equações normais; será aplicado o ajuste linear à deformação de uma mola sob forças crescentes, em conformidade com a Lei de Hooke, e à produção agrícola; e será aplicado o ajuste exponencial aos casos novos de COVID-19 no Brasil nas semanas epidemiológicas 13 a 29 de 2020, projetando a semana 30 em um cenário contrafactual, no qual se supõe que nada teria alterado o ritmo de crescimento. Serão apresentados, ainda, em caráter introdutório, os modelos logístico e de potência.

Para expor esse percurso, o trabalho organiza-se em cinco capítulos. O Capítulo 2 reúne as noções preliminares de cálculo que sustentam o método. O Capítulo 3 trata da motivação e da fundamentação do MMQ, discutindo o conceito de resíduo, a razão de se minimizar a soma dos quadrados e a obtenção das equações normais. O Capítulo 4 apresenta o método em ação, com os ajustes linear e exponencial aplicados a dados reais e a discussão de seus alcances e limites, além do panorama dos modelos logístico e de potência. O Capítulo 5, por fim, retoma a questão norteadora e apresenta as considerações finais.

Capítulo 2

Noções Preliminares

O Método dos Mínimos Quadrados é fundamentado em conceitos essenciais do cálculo diferencial e análise de funções. Para compreender e aplicar efetivamente este método na modelagem de fenômenos reais, é necessário dominar os instrumentos matemáticos que permitem identificar extremos de funções, calcular taxas de variação e resolver problemas de otimização.

Neste capítulo, apresentamos os conceitos preliminares que formam a base teórica para o desenvolvimento do Método dos Mínimos Quadrados. Estudaremos máximos e mínimos de funções reais, derivadas ordinárias e parciais, e suas aplicações na localização de pontos críticos. Estes conceitos não apenas fundamentam o método de ajuste de curvas que será explorado nos capítulos subsequentes, mas também ilustram como a matemática pura se traduz em ferramentas práticas para análise de dados.

As discussões e exemplos apresentados neste capítulo baseiam-se principalmente em Stewart [3] e Thomas [4], referências clássicas que oferecem uma abordagem clara e rigorosa do cálculo diferencial, essencial para a compreensão da teoria que sustenta a modelagem matemática de fenômenos.

2.1 Máximos e Mínimos para Funções de uma Variável

Nesta seção, apresentamos os conceitos fundamentais para a análise de extremos de funções de uma variável. Iniciamos com as definições formais de máximos e mínimos absolutos e locais, que estabelecem com precisão quando um ponto é considerado um extremo, diferenciando entre extremos globais (em todo o domínio) e locais (em uma vizinhança do ponto). Posteriormente, introduzimos o conceito de ponto crítico, que é essencial para localizar estes extremos em funções diferenciáveis. Estes conceitos formam a base para a compreensão de problemas de otimização, em muitos contextos práticos, necessitamos identificar o maior ou menor valor que uma função assume em seu domínio, utilizando exatamente as ferramentas que serão desenvolvidas aqui.

Definição 2.1. *Uma função f tem um máximo absoluto (ou máximo global) em c se $f(c) \geq f(x)$*

para todo x em D , onde D é o domínio de f , e o número $f(c)$ é denominado valor máximo de f em D .

Definição 2.2. Uma função f tem um mínimo absoluto em c se $f(c) \leq f(x)$ para todo x em D , e o número $f(c)$ é denominado valor mínimo de f em D .

Os valores máximo e mínimo de uma função f são chamados valores extremos de f .

Definição 2.3. Uma função f tem um máximo local (ou máximo relativo) em c se $f(c) \geq f(x)$ quando x estiver nas proximidades de c . [Isso significa que $f(c) \geq f(x)$ para todo x em algum intervalo aberto contendo c].

Definição 2.4. Uma função f tem um mínimo local em c se $f(c) \leq f(x)$ quando x estiver próximo de c .

Definição 2.5. Um ponto c no domínio de uma função f é chamado de ponto crítico (ou ponto estacionário) se $f'(c) = 0$ ou se a derivada $f'(c)$ não existe.

A importância dos pontos críticos reside no fato de que todo extremo de uma função contínua em um domínio fechado deve ocorrer em um ponto crítico ou nos extremos do domínio. Portanto, para localizar extremos, nossa estratégia é encontrar todos os pontos críticos, avaliar a função nestes pontos e nos extremos do domínio, e então comparar os valores para identificar máximos e mínimos.

Para ilustrar a relevância prática destes conceitos, apresentamos alguns exemplos de otimização em contextos reais:

Observação 2.6. Na medicina, farmacologistas determinam a dose ótima de medicamentos encontrando o máximo da concentração plasmática no organismo, garantindo eficácia com segurança [1].

Observação 2.7. Na logística, empresas utilizam algoritmos de otimização para minimizar distâncias e custos de entrega, resolvendo variações do Problema do Caixeiro Viajante [2].

Os exemplos a seguir servem para ilustrar na prática os conceitos teóricos de máximos e mínimos apresentados anteriormente.

Exemplo 2.8. Considere a função $f(x) = x^2$ definida no domínio $D = [-2, 3]$. Identifique seus extremos absolutos e locais.

Resolução: Para encontrar os extremos dessa função, começamos avaliando f em alguns pontos principais, nos extremos do domínio e nos pontos onde a derivada se anula. Calculamos:

$$f(-2) = 4,$$

$$f(-1) = 1,$$

$$f(0) = 0,$$

$$f(1) = 1,$$

$$f(2) = 4,$$

$$f(3) = 9.$$

Observando esses valores, percebemos que o maior valor é $f(3) = 9$. Pela definição de máximo absoluto, como $9 \geq f(x)$ para todo $x \in [-2, 3]$, temos que $x = 3$ é o ponto de máximo absoluto com valor máximo igual a 9.

Por outro lado, o menor valor obtido é $f(0) = 0$. Como $0 \leq f(x)$ para todo x no domínio, pela definição de mínimo absoluto, o ponto $x = 0$ é o ponto de mínimo absoluto com valor mínimo igual a 0.

Além disso, o ponto $x = 0$ também é mínimo local, pois existe uma vizinhança ao redor de 0 (por exemplo, o intervalo $(-1, 1)$) onde $f(0) \leq f(x)$ para todos os pontos nessa vizinhança.

Exemplo 2.9. Analise os extremos da função $f(x) = -x^3 + 3x$ no domínio $D = [-2, 2]$.

Resolução: Para identificar os possíveis extremos dessa função, começamos calculando a derivada:

$$f'(x) = -3x^2 + 3$$

Igualando a derivada a zero para encontrar os pontos críticos:

$$-3x^2 + 3 = 0 \Rightarrow x^2 = 1 \Rightarrow x = -1 \text{ ou } x = 1$$

Agora precisamos avaliar a função nos pontos críticos encontrados e também nos extremos do intervalo. Calculamos os valores:

$$f(-2) = -(-2)^3 + 3(-2) = 8 - 6 = 2.$$

$$f(-1) = -(-1)^3 + 3(-1) = 1 - 3 = -2.$$

$$f(1) = -(1)^3 + 3(1) = -1 + 3 = 2.$$

$$f(2) = -(2)^3 + 3(2) = -8 + 6 = -2.$$

Comparando todos os valores:

$$f(-2) = 2, f(-1) = -2, f(1) = 2, f(2) = -2.$$

O maior valor é 2, que ocorre em $x = -2$ e em $x = 1$. Como $f(-2) \geq f(x)$ e $f(1) \geq f(x)$ para todo $x \in [-2, 2]$, ambos são pontos de máximo absoluto.

O menor valor é -2 , que ocorre em $x = -1$ e em $x = 2$. Como $f(-1) \leq f(x)$ e $f(2) \leq f(x)$ para todo x no domínio, ambos são pontos de mínimo absoluto.

Observe também que o ponto $x = 1$ é um máximo local (não apenas absoluto), pois existe um intervalo aberto ao redor de 1 onde $f(1)$ é o maior valor. Similarmente, $x = -1$ é um mínimo local.

Exemplo 2.10. Considere a função definida por partes:

$$f(x) = \begin{cases} \sin(\frac{1}{x}) & \text{se } x \neq 0, \\ 0 & \text{se } x = 0. \end{cases}$$

no domínio $D = [-1, 1]$. Analise a existência de extremos absolutos e locais.

Resolução: Começamos entendendo o comportamento da função. Para todo $x \neq 0$, sabemos que:

$$-1 \leq \sin\left(\frac{1}{x}\right) \leq 1$$

pois o seno sempre varia entre -1 e 1 . No ponto $x = 0$, por definição, $f(0) = 0$.

Analisamos a continuidade de f em $x = 0$. Quando $x \rightarrow 0^+$, temos $1/x \rightarrow +\infty$, de modo que o argumento $1/x$ percorre repetidamente todo o círculo trigonométrico, fazendo $\sin(1/x)$ oscilar infinitas vezes entre -1 e 1 sem se aproximar de um único valor. De fato, tomando as sequências $x_n = \frac{1}{n\pi} \rightarrow 0$ e $x'_n = \frac{2}{(4n+1)\pi} \rightarrow 0$, obtemos

$$f(x_n) = \sin(n\pi) = 0 \quad \text{e} \quad f(x'_n) = \sin\left(2n\pi + \frac{\pi}{2}\right) = 1$$

para todo n ; como as duas sequências tendem a 0 mas $f(x_n) \rightarrow 0$ e $f(x'_n) \rightarrow 1$, o limite $\lim_{x \rightarrow 0} f(x)$ não existe. Logo, f não é contínua em $x = 0$, qualquer que seja o valor atribuído a $f(0)$. Em todos os demais pontos de $[-1, 1]$, f é contínua, por ser composição de funções contínuas.

Agora analisamos os extremos. Observamos que para valores de x muito próximos de 0 (tanto à esquerda quanto à direita), a função oscila infinitas vezes entre -1 e 1 , sem se estabilizar

em um único valor.

Investigando a existência de máximo absoluto, embora a função seja limitada superiormente por 1, ela nunca realmente atinge o valor 1 no intervalo. Existem pontos arbitrariamente próximos de 0 onde $\sin(\frac{1}{x})$ fica arbitrariamente perto de 1, mas nunca iguala 1. Portanto, não existe um ponto $c \in [-1, 1]$ onde $f(c) \geq f(x)$ para todos os x no domínio. O supremo (o menor limite superior) é 1, mas não é atingido.

Para o mínimo absoluto, a situação é semelhante, o ínfimo é -1 , mas também não é atingido por nenhum ponto do intervalo.

Além disso, o ponto $x = 0$ não é sequer um extremo local: em qualquer vizinhança $(-\epsilon, \epsilon)$ de 0, por menor que seja $\epsilon > 0$, a função assume valores arbitrariamente próximos de 1 e de -1 , isto é, valores tanto maiores quanto menores que $f(0) = 0$. Assim, $f(0)$ não é nem cota superior nem cota inferior dos valores de f nessa vizinhança, e $x = 0$ não pode ser classificado como máximo nem como mínimo local.

Com este exemplo podemos ver que uma função limitada em um domínio compacto nem sempre possui máximo e mínimo absolutos. Isso não contradiz o Teorema de Weierstrass: como f não é contínua em $x = 0$, a hipótese de continuidade em todo o domínio $[-1, 1]$ não é satisfeita, de modo que o teorema simplesmente não se aplica a esta função. O exemplo ilustra, assim, por que a hipótese de continuidade é indispensável no enunciado do Teorema de Weierstrass.

Uma ferramenta fundamental para a localização de extremos em funções diferenciáveis é o Teorema de Fermat. Este teorema estabelece uma condição necessária para que um ponto seja um extremo local de uma função, em tal ponto a derivada deve ser nula. Embora essa condição não seja suficiente (pois existem pontos onde a derivada se anula mas não há extremo), ela é essencial para reduzir o conjunto de candidatos a extremantes e constitui o ponto de partida para muitos métodos de otimização.

Teorema 2.11. (Teorema de Fermat) Se f é uma função contínua no intervalo fechado $[a, b]$, diferenciável no intervalo aberto (a, b) , e f tem um extremo local (máximo ou mínimo) em um ponto $c \in (a, b)$, então $f'(c) = 0$.

Demonstração: Suponhamos, sem perda de generalidade, que f tem um máximo local em $c \in (a, b)$. Pela definição de máximo local, existe um intervalo aberto $(c - \delta, c + \delta)$ com $\delta > 0$ tal que $f(c) \geq f(x)$ para todo x nesse intervalo.

Para qualquer $h > 0$ com $h < \delta$, temos $f(c + h) \leq f(c)$, logo $f(c + h) - f(c) \leq 0$. Dividindo por $h > 0$:

$$\frac{f(c + h) - f(c)}{h} \leq 0.$$

Tomando o limite quando $h \rightarrow 0^+$:

$$f'_+(c) = \lim_{h \rightarrow 0^+} \frac{f(c+h) - f(c)}{h} \leq 0.$$

Por outro lado, para qualquer $h > 0$ com $h < \delta$, temos:

$$f(c-h) \leq f(c),$$

logo,

$$f(c-h) - f(c) \leq 0.$$

Dividindo por $-h < 0$ (inverte a desigualdade):

$$\frac{f(c-h) - f(c)}{-h} \geq 0.$$

Equivalentemente:

$$\frac{f(c+h) - f(c)}{h} \geq 0 \quad (\text{substituindo } h \text{ por } -h).$$

Tomando o limite quando $h \rightarrow 0^-$:

$$f'_-(c) = \lim_{h \rightarrow 0^-} \frac{f(c+h) - f(c)}{h} \geq 0$$

Como f é diferenciável em c , temos $f'_+(c) = f'_-(c) = f'(c)$. Das desigualdades acima, $f'(c) \leq 0$ e $f'(c) \geq 0$, concluímos que $f'(c) = 0$. O caso de mínimo local é análogo.

Observação 2.12. O Teorema de Fermat estabelecido aqui se aplica a funções de uma variável. Este resultado será posteriormente generalizado para funções de múltiplas variáveis, em que as derivadas ordinárias são substituídas por derivadas parciais. Em particular, para uma função $f(x, y)$, se existe um extremo local em (a, b) , então $f_x(a, b) = 0$ e $f_y(a, b) = 0$. Esta generalização é essencial para resolver problemas de otimização com múltiplos parâmetros, como ocorre no Método dos Mínimos Quadrados.

Exemplo 2.13. Considere a função $f(x) = x^3 - 3x$ no intervalo $[-2, 2]$. Use o Teorema de Fermat para encontrar os pontos críticos onde a derivada se anula.

Resolução: Calculamos a derivada da função:

$$f'(x) = 3x^2 - 3.$$

Pelo Teorema de Fermat, se f possui extremos locais no interior do intervalo, eles devem ocorrer onde $f'(x) = 0$. Igualando a derivada a zero:

$$3x^2 - 3 = 0 \Rightarrow x^2 = 1 \Rightarrow x = \pm 1.$$

Assim, os pontos críticos são

$$x = -1 \quad \text{e} \quad x = 1.$$

Verificamos que

$$f(-1) = (-1)^3 - 3(-1) = -1 + 3 = 2$$

e

$$f(1) = (1)^3 - 3(1) = 1 - 3 = -2.$$

De fato, observando o comportamento da função, $x = -1$ é um máximo local e $x = 1$ é um mínimo local, confirmando que o Teorema de Fermat forneceu os candidatos corretos aos extremos locais.

O Teorema de Fermat fornece uma condição necessária para localizar extremos: a derivada deve se anular nesses pontos. Na prática, este princípio é fundamental para resolver problemas de otimização, quando desejamos encontrar o melhor ajuste de uma curva aos dados (minimizando erros), estamos buscando valores de parâmetros que minimizam uma função de erro. O processo é idêntico ao que realizamos aqui, derivar, igualar a zero e resolver o sistema resultante. Portanto, dominar a técnica de encontrar extremos de funções é essencial para compreender o Método dos Mínimos Quadrados.

Passamos agora a desenvolver formalmente o conceito de derivada para funções com mais de uma variável, que é a ferramenta fundamental para este tipo de análise.

2.2 Máximos e Mínimos para Funções de duas Variáveis

Nesta seção, apresentamos os conceitos fundamentais para a análise de extremos de funções reais de duas variáveis. Pressupomos que o leitor já domina os conceitos de derivadas parciais de primeira e segunda ordem; para uma revisão detalhada, recomendamos Stewart [3] e Thomas [4].

Definição 2.14. *Uma função real de duas variáveis é uma regra que associa a cada par ordenado (x, y) de um domínio $D \subseteq \mathbb{R}^2$ um único número real $f(x, y) \in \mathbb{R}$. Denotamos essa função por $f : D \subseteq \mathbb{R}^2 \rightarrow \mathbb{R}$.*

Para localizar extremos de funções de duas variáveis, uma ferramenta central é o *vetor gradiente*, que generaliza o conceito de derivada; o gradiente aponta na direção de maior crescimento da função.

Definição 2.15. *Seja $f : D \subseteq \mathbb{R}^2 \rightarrow \mathbb{R}$ uma função diferenciável. O vetor gradiente (ou gradiente) de f no ponto (x, y) é o vetor:*

$$\nabla f(x, y) = \left(\frac{\partial f}{\partial x}(x, y), \frac{\partial f}{\partial y}(x, y) \right) = (f_x(x, y), f_y(x, y)).$$

Observação 2.16. *O vetor gradiente $\nabla f(x, y)$ é um vetor bidimensional situado no plano do domínio de f , isto é, no plano (x, y) . Ele não pertence ao gráfico tridimensional da função, mas aponta a direção e o sentido de máxima taxa de crescimento de f no domínio. Esta propriedade o torna essencial em problemas de otimização e em métodos numéricos de busca por extremos [3].*

Como exemplo concreto, calculamos o gradiente de uma função simples:

Exemplo 2.17. *Encontre o vetor gradiente $\nabla f(x, y)$ da função $f(x, y) = xy + x^2$ e avalie-o no ponto $(2, 1)$.*

Resolução: Primeiramente, calculamos as derivadas parciais de primeira ordem:

$$f_x(x, y) = y + 2x$$

$$f_y(x, y) = x$$

Portanto, o vetor gradiente é:

$$\nabla f(x, y) = (y + 2x, x).$$

No ponto $(2, 1)$:

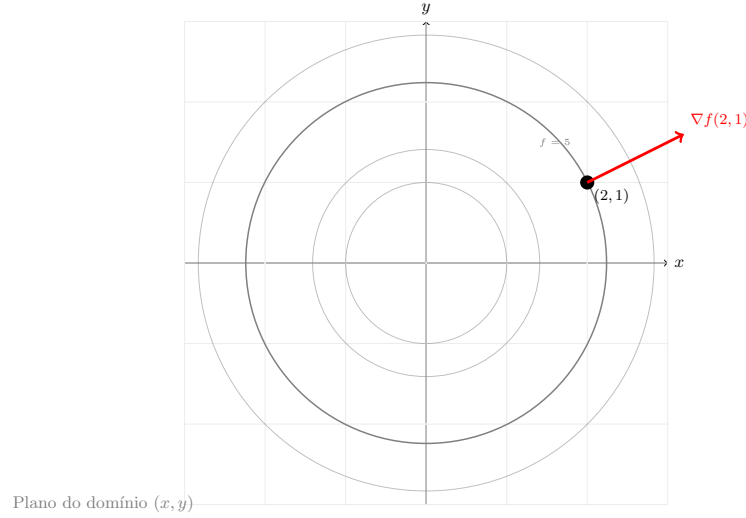
$$\nabla f(2, 1) = (1 + 2 \cdot 2, 2) = (5, 2).$$

Este vetor indica a direção e o sentido de máxima taxa de crescimento de f naquele ponto do domínio. A Figura 2.1 ilustra geometricamente o vetor gradiente de uma função no plano do domínio:

Exemplo 2.18. *Seja $f(x, y) = x^2 + 2y^2$. Calcule o vetor gradiente $\nabla f(x, y)$, avalie-o no ponto $(1, 2)$ e identifique as curvas de nível da função.*

Resolução: Calculamos as derivadas parciais de primeira ordem:

$$f_x(x, y) = 2x, \quad f_y(x, y) = 4y.$$

Figura 2.1. Vetor gradiente de $f(x, y) = xy + x^2$ no ponto $(2, 1)$.

Fonte: Elaboração própria (2026).

Portanto, o vetor gradiente é:

$$\nabla f(x, y) = (2x, 4y).$$

Avaliando no ponto $(1, 2)$:

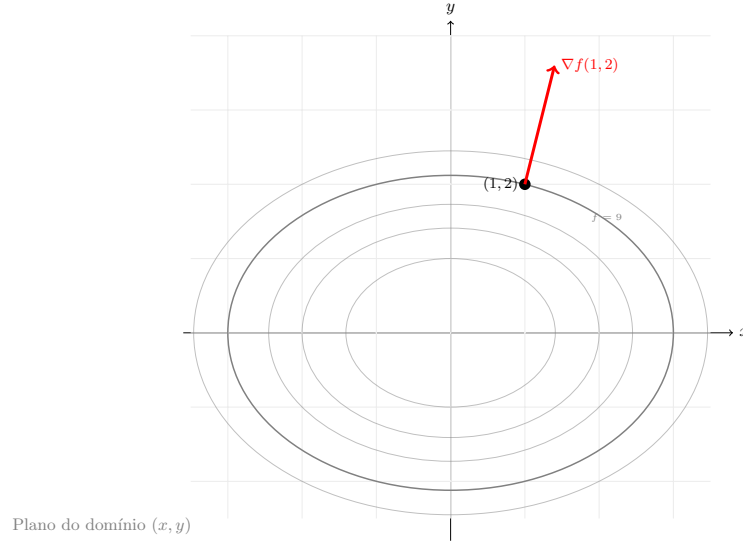
$$\nabla f(1, 2) = (2 \cdot 1, 4 \cdot 2) = (2, 8).$$

As curvas de nível de f são as elipses $x^2 + 2y^2 = c$, para $c > 0$. Em particular, o ponto $(1, 2)$ pertence à elipse $x^2 + 2y^2 = 9$, pois $f(1, 2) = 1 + 8 = 9$. O vetor $\nabla f(1, 2) = (2, 8)$ é perpendicular a essa elipse no ponto $(1, 2)$ e aponta no sentido de crescimento de f , em direção às elipses de nível mais alto. Note que a componente em y é dominante ($8 \gg 2$), o que reflete o maior impacto de y na função neste ponto. A Figura 2.2 ilustra essa relação geometricamente:

Nas subseções seguintes, utilizaremos o conceito de gradiente para caracterizar máximos e mínimos de funções de duas variáveis.

2.2.1 Otimização

Como vimos, o Teorema de Fermat estabelece que, em um extremo local de uma função diferenciável de uma variável, necessariamente $f'(c) = 0$. Para funções de duas variáveis, ocorre um fenômeno semelhante, nos pontos de extremo local o vetor gradiente é nulo. Os conceitos de máximo e mínimo local estendem-se naturalmente da Seção 2.1, substituindo o ponto $c \in \mathbb{R}$ pelo par $(a, b) \in \mathbb{R}^2$ e o intervalo aberto por um disco aberto centrado em (a, b) .

Figura 2.2. Vetor gradiente de $f(x, y) = x^2 + 2y^2$ no ponto $(1, 2)$.

Fonte: Elaboração própria (2026).

Definição 2.19. Um ponto (a, b) no domínio de f é chamado de ponto crítico de f se $\nabla f(a, b) = (0, 0)$, ou seja, se

$$f_x(a, b) = 0 \quad e \quad f_y(a, b) = 0,$$

ou se alguma das derivadas parciais não existe em (a, b) . Todo extremo local de f ocorre em um ponto crítico; portanto, os pontos críticos são os candidatos a extremo.

A condição $\nabla f = (0, 0)$ é necessária, mas não suficiente; um ponto crítico pode ser máximo local, mínimo local ou nenhum dos dois. O teste a seguir permite distinguir esses casos.

Teorema 2.20 (Teste da Segunda Derivada). Seja f uma função com derivadas parciais de segunda ordem contínuas em uma vizinhança de um ponto crítico (a, b) . Define-se o *discriminante* de f em (a, b) por

$$D(a, b) = f_{xx}(a, b) f_{yy}(a, b) - [f_{xy}(a, b)]^2.$$

- (i) Se $D > 0$ e $f_{xx}(a, b) > 0$, então $f(a, b)$ é um *mínimo local*.
- (ii) Se $D > 0$ e $f_{xx}(a, b) < 0$, então $f(a, b)$ é um *máximo local*.
- (iii) Se $D < 0$, então (a, b) é um *ponto de sela* (não é extremo local).
- (iv) Se $D = 0$, o teste é *inconclusivo*.

Demonstração. A prova deste resultado, baseada na fórmula de Taylor de segunda ordem e na análise da forma quadrática associada à matriz Hessiana, pode ser encontrada em [7], p. 569.

Observação 2.21. *A discriminante D depende apenas das derivadas parciais de segunda ordem de f avaliadas no ponto crítico, sendo calculado sem informações globais sobre a função.*

Para localizar e classificar os extremos locais de uma função de duas variáveis, o procedimento consiste em resolver o sistema $\nabla f = (0, 0)$ para obter os pontos críticos e, em seguida, aplicar o Teste da Segunda Derivada a cada um deles.

Exemplo 2.22. Encontre e classifique os pontos críticos de $f(x, y) = x^2 + y^2 - 2x - 4y + 5$.

Resolução: Calculamos as derivadas parciais de primeira ordem e igualamos a zero:

$$f_x(x, y) = 2x - 2 = 0 \implies x = 1,$$

$$f_y(x, y) = 2y - 4 = 0 \implies y = 2.$$

O único ponto crítico é $(1, 2)$. As derivadas de segunda ordem são:

$$f_{xx} = 2, \quad f_{yy} = 2, \quad f_{xy} = 0.$$

O discriminante é:

$$D(1, 2) = 2 \cdot 2 - 0^2 = 4 > 0.$$

Como $D > 0$ e $f_{xx} = 2 > 0$, pelo Teste da Segunda Derivada, $f(1, 2) = 0$ é um **mínimo local**.

Exemplo 2.23. Encontre e classifique todos os pontos críticos de $f(x, y) = x^3 - 3x + y^2$.

Resolução: Igualamos as derivadas parciais a zero:

$$f_x(x, y) = 3x^2 - 3 = 0 \implies x^2 = 1 \implies x = 1 \text{ ou } x = -1,$$

$$f_y(x, y) = 2y = 0 \implies y = 0.$$

Os pontos críticos são $(1, 0)$ e $(-1, 0)$. As derivadas de segunda ordem são:

$$f_{xx}(x, y) = 6x, \quad f_{yy}(x, y) = 2, \quad f_{xy}(x, y) = 0.$$

O discriminante é:

$$D(x, y) = 6x \cdot 2 - 0^2 = 12x.$$

Em $(1, 0)$: $D(1, 0) = 12 > 0$ e $f_{xx}(1, 0) = 6 > 0 \implies$ mínimo local, com $f(1, 0) = -2$.

Em $(-1, 0)$: $D(-1, 0) = -12 < 0 \implies$ ponto de sela, com $f(-1, 0) = 2$.

Portanto, f possui um mínimo local em $(1, 0)$ e um ponto de sela em $(-1, 0)$.

Exemplo 2.24. Uma empresa fabrica dois produtos, A e B , em quantidades x e y (em centenas de unidades por mês). O lucro mensal, em milhares de reais, é modelado pela função

$$L(x, y) = 8x + 10y - x^2 - y^2 - xy.$$

Determine as quantidades que maximizam o lucro e o valor do lucro máximo.

Resolução: Calculamos as derivadas parciais de primeira ordem e igualamos a zero para obter os pontos críticos:

$$L_x(x, y) = 8 - 2x - y = 0,$$

$$L_y(x, y) = 10 - 2y - x = 0.$$

Da primeira equação, $y = 8 - 2x$. Substituindo na segunda:

$$10 - 2(8 - 2x) - x = 0 \implies 10 - 16 + 4x - x = 0 \implies 3x - 6 = 0 \implies x = 2.$$

Logo, $y = 8 - 2 \cdot 2 = 4$. O único ponto crítico é $(2, 4)$.

As derivadas de segunda ordem são:

$$L_{xx} = -2, \quad L_{yy} = -2, \quad L_{xy} = -1.$$

O discriminante é:

$$D(2, 4) = L_{xx} L_{yy} - [L_{xy}]^2 = (-2)(-2) - (-1)^2 = 4 - 1 = 3 > 0.$$

Como $D > 0$ e $L_{xx} = -2 < 0$, pelo Teste da Segunda Derivada, $(2, 4)$ é um **máximo local**. O lucro máximo é

$$L(2, 4) = 8 \cdot 2 + 10 \cdot 4 - 2^2 - 4^2 - 2 \cdot 4 = 16 + 40 - 4 - 16 - 8 = 28.$$

Portanto, a empresa maximiza o lucro produzindo 200 unidades de A e 400 unidades de B por mês, obtendo um lucro mensal de 28 mil reais. Vale notar que esse mesmo procedimento, isto é, anular o gradiente e classificar o ponto crítico pelo discriminante, é o que sustentará a obtenção dos parâmetros do Método dos Mínimos Quadrados, no qual a função a minimizar é a soma dos quadrados dos resíduos, também uma função de duas variáveis.

Com os conceitos de ponto crítico, gradiente e teste da segunda derivada estabelecidos, estamos preparados para avançar. No próximo capítulo, apresentamos a motivação para

o Método dos Mínimos Quadrados, discutimos por que a soma dos quadrados dos resíduos é o critério de ajuste mais adequado e como o problema de encontrar a melhor curva para um conjunto de dados se reduz, essencialmente, a um problema de otimização das ferramentas desenvolvidas aqui.

Capítulo 3

Motivação para o Método dos Mínimos Quadrados

Em muitas situações práticas, dispomos de um conjunto de dados coletados experimentalmente (pares (x_i, y_i) , $i = 1, \dots, n$) e desejamos encontrar uma função f que os represente de forma satisfatória. Essa necessidade surge, por exemplo, quando um médico acompanha a concentração de um medicamento no sangue ao longo das horas, quando um economista observa a inflação mês a mês, ou quando um físico mede a posição de um corpo em queda em instantes sucessivos. Em todos esses casos, os dados são pontos isolados, mas o fenômeno por trás deles é contínuo e aqui gostaríamos de uma *função* que o descrevesse.

Há uma inversão fundamental no problema de ajuste de curvas, os dados são conhecidos e fixos, e a incógnita é a função; isso inverte o hábito que trazemos do Cálculo. Estamos acostumados a receber uma função f e perguntar coisas sobre ela (sua derivada, seus máximos, suas raízes), mas agora queremos saber: dados os pontos, qual é a melhor função? É exatamente o mesmo tipo de inversão que motiva os Métodos Variacionais, em vez de avaliar funções, procuramos, dentro de uma família de funções, aquela que otimiza um certo critério.

Raramente, porém, existe uma função simples que passe exatamente por todos os pontos. Medições carregam erros de instrumento, de leitura, de arredondamento; fenômenos reais sofrem perturbações aleatórias. Insistir em uma curva que passe por todos os pontos significaria modelar também o erro, e não apenas o fenômeno. O objetivo é encontrar a função que mais se aproxima dos dados, segundo algum critério de qualidade de ajuste; este capítulo é dedicado a uma única pergunta: que critério é esse, e por quê?

Antes de medir o erro total de um ajuste, precisamos medir o erro em cada ponto. Suponha que escolhemos uma função candidata f . No ponto x_i , o valor *observado* foi y_i , mas o modelo *prevê* $f(x_i)$. A diferença entre o que vimos e o que o modelo diz é o que chamamos de *resíduo*.

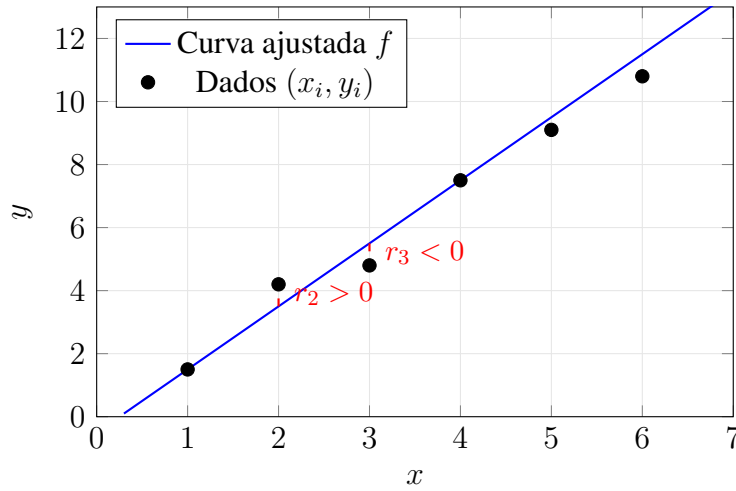
Definição 3.1. Dado um conjunto de pontos (x_i, y_i) , $i = 1, \dots, n$, e uma função f , o *resíduo*

(ou erro) no i -ésimo ponto é definido por

$$r_i = y_i - f(x_i).$$

Geometricamente, $|r_i|$ é o comprimento do segmento vertical que liga o ponto observado (x_i, y_i) à curva. O sinal de r_i informa a posição $r_i > 0$ quando o ponto está *acima* da curva (o modelo subestimou) e $r_i < 0$ quando está *abaixo* (o modelo superestimou). A Figura 3.1 ilustra a situação.

Figura 3.1. Diagrama de dispersão com curva ajustada e resíduos



Fonte: Elaboração própria (2026).

Cada resíduo mede um erro local, e como podemos combinar os n resíduos $r_1, \dots, r_n$ em um único número que meça o erro *total* do ajuste? Esse número será a quantidade a minimizar; a escolha de como construí-lo determina todo o método. Analisaremos três candidatos naturais e veremos, por meio de exemplos numéricos, por que os dois primeiros falham e o terceiro funciona.

3.1 Primeira tentativa: minimizar a soma dos resíduos

A ideia mais ingênua é somar os erros:

$$E_1(f) = \sum_{i=1}^n r_i = \sum_{i=1}^n (y_i - f(x_i)) \longrightarrow \min.$$

Parece fácil, já que cada r_i mede um erro, então a soma mede o erro total. O problema é que os resíduos têm sinal, e sinais opostos se cancelam. Um modelo pode errar muito para cima em alguns pontos e muito para baixo em outros, e a soma resultar zero, mascarando um ajuste péssimo. O exemplo a seguir torna isso concreto.

Exemplo 3.2 (O empate escandaloso). *Considere os três pontos*

$$(1, 1), \quad (2, 2), \quad (3, 3),$$

que estão perfeitamente alinhados sobre a reta $y = x$. Vamos comparar duas retas candidatas usando o critério E_1 .

Candidata A: $f(x) = x$ (o ajuste perfeito). *Os resíduos são*

$$r_1 = 1 - 1 = 0, \quad r_2 = 2 - 2 = 0, \quad r_3 = 3 - 3 = 0,$$

logo $E_1 = 0 + 0 + 0 = 0$.

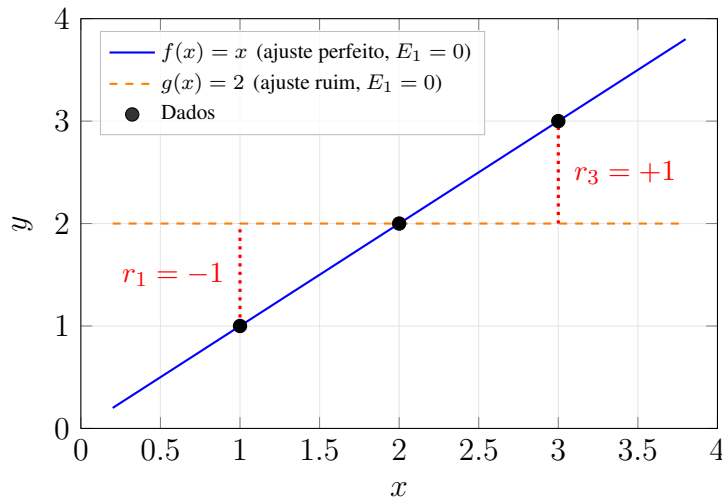
Candidata B: $g(x) = 2$ (uma reta horizontal, que ignora completamente a tendência crescente dos dados). *Os resíduos são*

$$r_1 = 1 - 2 = -1, \quad r_2 = 2 - 2 = 0, \quad r_3 = 3 - 2 = 1,$$

logo $E_1 = -1 + 0 + 1 = 0$.

Conclusão: *pelo critério da soma dos resíduos, a reta horizontal $g(x) = 2$ é tão boa quanto o ajuste perfeito $f(x) = x$, pois ambas atingem o valor mínimo possível em módulo. O critério é incapaz de distinguir um ajuste exato de um ajuste visivelmente ruim. A Figura 3.2 mostra a situação: o erro negativo em $x = 1$ cancela exatamente o erro positivo em $x = 3$.*

Figura 3.2. Duas retas com soma de resíduos nula: o critério não as distingue



Fonte: Elaboração própria (2026).

Na verdade, a situação é ainda pior do que um empate. Como os resíduos podem ser arbitrariamente negativos, é possível tornar E_1 tão negativo quanto se queira, basta escolher uma função que passe muito acima de todos os pontos. Por exemplo, com os mesmos dados, a reta

$h(x) = 100$ produz $E_1 = (1 - 100) + (2 - 100) + (3 - 100) = -294$. Pelo critério “minimizar E_1 ”, a reta h seria melhor que o ajuste perfeito! Ou seja, o problema de minimização sequer está bem posto, pois E_1 não possui mínimo, por poder ser empurrado para $-\infty$. Descartamos a primeira tentativa.

3.2 Segunda tentativa: minimizar a soma dos módulos

O problema ficou claro, o cancelamento de sinais. O remédio direto é tomar o valor absoluto de cada resíduo:

$$E_2(f) = \sum_{i=1}^n |r_i| = \sum_{i=1}^n |y_i - f(x_i)| \longrightarrow \min.$$

Agora todos os erros contribuem positivamente, e o problema do Exemplo 3.2 desaparece, para a reta horizontal $g(x) = 2$, teríamos $E_2 = |-1| + |0| + |1| = 2 > 0$, enquanto o ajuste perfeito mantém $E_2 = 0$. O critério agora distingue corretamente os dois ajustes.

Então, por que não paramos aqui? Porque, ao consertar o cancelamento, criamos um problema novo, de natureza diferente, a função módulo não é diferenciável em zero. E isso não é um detalhe técnico, é fatal para a nossa estratégia. Como vimos no Capítulo 2, para minimizar, derivamos, igualamos a zero e resolvemos ($\nabla S = 0$, Teorema de Fermat, Teste da Segunda Derivada). Se a função de erro tem “bicos”, esse procedimento simplesmente não se aplica nos pontos mais importantes, que são, justamente, os candidatos a mínimo. O exemplo a seguir mostra as duas patologias de uma só vez.

Exemplo 3.3 (O fundo achatado). *Tomemos o caso mais simples possível de ajuste: aproximar dois dados por uma constante. Observamos os valores*

$$y_1 = 0 \quad e \quad y_2 = 10$$

(digamos, duas medições de uma mesma grandeza) e queremos a constante c que melhor os representa. Aqui a família de funções candidatas é $f_c(x) = c$, e o parâmetro a determinar é o número c . Pelo critério dos módulos, devemos minimizar

$$E_2(c) = |0 - c| + |10 - c| = |c| + |10 - c|.$$

Vamos calcular E_2 em alguns valores de c :

c	-2	0	2	5	8	10	12
$E_2(c)$	$2+12=14$	$0+10=10$	$2+8=10$	$5+5=10$	$8+2=10$	$10+0=10$	$12+2=14$

Note o que aconteceu: todos os valores de c entre 0 e 10 produzem exatamente o mesmo erro total $E_2 = 10$. De fato, se $0 \leq c \leq 10$, então $|c| + |10 - c| = c + (10 - c) = 10$, constante.

O gráfico de E_2 (Figura 3.3, à esquerda) tem um “fundo achatado”, um intervalo inteiro de mínimos.

Disso decorrem dois problemas:

(i) **Não unicidade.** O critério não escolhe uma resposta: $c = 0,5$, $c = 5$ e $c = 9,9$ são todos “igualmente ótimos”. Um método de ajuste que devolve infinitas respostas, algumas delas claramente desequilibradas (como $c = 0,5$, coladíssima em um dado e longíssima do outro), não cumpre seu papel de selecionar o melhor modelo.

(ii) **A ferramenta do Capítulo 2 quebra.** Tentemos aplicar o procedimento padrão: derivar E_2 e igualar a zero. Para $c < 0$, temos $E_2(c) = -c + (10 - c) = 10 - 2c$, com derivada -2 ; para $c > 10$, temos $E_2(c) = c + (c - 10) = 2c - 10$, com derivada $+2$; e no platô $0 < c < 10$ a derivada é 0 em todos os pontos. Nos pontos $c = 0$ e $c = 10$, onde o gráfico forma bicos, a derivada nem existe. A condição $E'_2(c) = 0$ não isola candidato algum: ela é satisfeita por um intervalo inteiro. O Teorema de Fermat, o Teste da Segunda Derivada, o sistema $\nabla S = 0$, nada disso fornece uma resposta única aqui.

O critério dos módulos é conceitualmente honesto, mede mesmo o erro total, mas operacionalmente intratável com as ferramentas do cálculo diferencial, além de poder admitir infinitas soluções.¹ Precisamos de um critério que mantenha a virtude (erros não se cancelam) e elimine o defeito (falta de diferenciabilidade).

3.3 Terceira tentativa: minimizar a soma dos quadrados

Há outra maneira, além do módulo, de tornar um número não negativo, elevá-lo ao quadrado. Isso sugere o critério

$$E_3(f) = \sum_{i=1}^n r_i^2 = \sum_{i=1}^n (y_i - f(x_i))^2 \longrightarrow \min.$$

A diferença crucial é que a função $r \mapsto r^2$ é polinomial, suave, com derivadas de todas as ordens em todos os pontos. Voltemos ao exemplo anterior e vejamos o critério em ação.

Exemplo 3.4 (O mesmo problema, agora com final feliz). Com os mesmos dados $y_1 = 0$ e $y_2 = 10$ do Exemplo 3.3, o critério dos quadrados pede para minimizar

$$E_3(c) = (0 - c)^2 + (10 - c)^2 = c^2 + (10 - c)^2 = 2c^2 - 20c + 100.$$

¹Registre-se, por honestidade, que minimizar $\sum |r_i|$ não é uma ideia absurda: ela define a chamada *regressão L_1* , útil quando há dados muito discrepantes (*outliers*), e é tratada com técnicas de otimização não diferenciável, como programação linear. O ponto aqui é que ela não se resolve com “derivar e igualar a zero”, que é o instrumental deste trabalho.

Calculando em alguns valores:

c	0	2	4	5	6	10
$E_3(c)$	100	68	52	50	52	100

O fundo achatado sumiu, os valores decrescem até $c = 5$ e voltam a crescer. E agora conseguimos aplicar o teorema perfeitamente. Derivando:

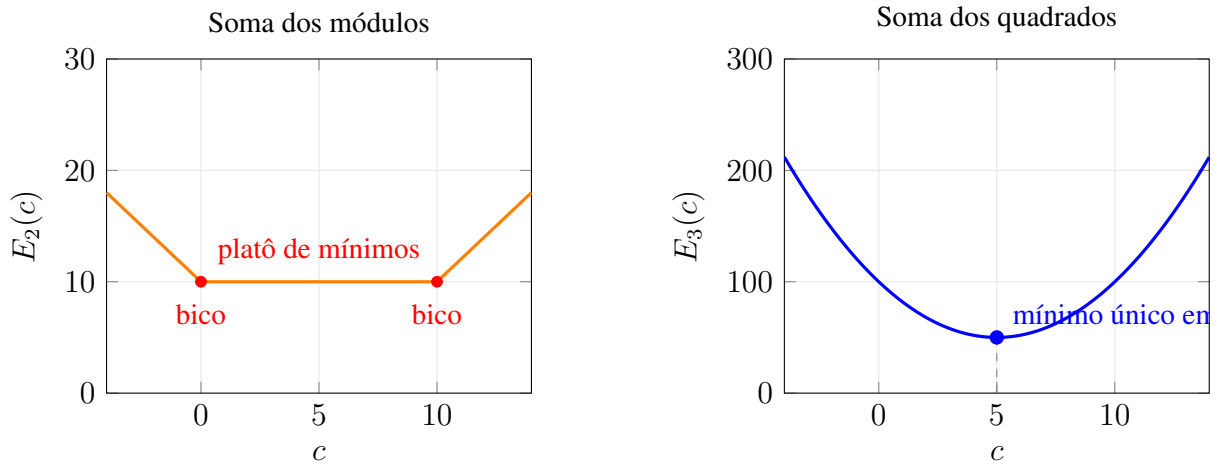
$$E'_3(c) = 2c - 2(10 - c) = 4c - 20.$$

Igualando a zero (Teorema de Fermat): $4c - 20 = 0 \Rightarrow c = 5$. Como $E''_3(c) = 4 > 0$, trata-se de um mínimo, e por ser E_3 uma parábola com concavidade para cima, o mínimo é global e único.

Vale observar quem é a resposta: $c = 5 = \frac{0+10}{2}$, a média aritmética dos dados. Isso não é coincidência: ajustar uma constante por mínimos quadrados sempre devolve a média, o método formaliza e generaliza a noção mais intuitiva que temos de valor representativo.

A Figura 3.3 compara os dois critérios lado a lado no mesmo problema. À esquerda, E_2 com seu platô de mínimos e seus bicos; à direita, E_3 , uma parábola lisa com um único ponto de mínimo, localizável pela condição $E'_3 = 0$.

Figura 3.3. Função de erro dos módulos (E_2) versus dos quadrados (E_3) para os dados $\{0, 10\}$



Fonte: Elaboração própria (2026).

3.3.1 Balanço dos três critérios

Os exemplos evidenciaram por que o critério dos quadrados supera os dois anteriores. A primeira vantagem é a ausência de cancelamento, cada parcela r_i^2 é não negativa, de modo que $E_3(f) = 0$ ocorre se, e somente se, todos os resíduos são nulos, a curva passa por todos

os pontos. O ajuste ruim do Exemplo 3.2, que enganava E_1 com $E_1 = 0$, agora recebe $E_3 = (-1)^2 + 0^2 + 1^2 = 2 > 0$, e o ajuste perfeito vence, como deveria.

A segunda vantagem é a diferenciabilidade. A função $r \mapsto r^2$ é suave, polinomial, com derivadas de todas as ordens em todos os pontos. Quando f depende de parâmetros, E_3 é uma função diferenciável (polinomial, nos modelos lineares) nesses parâmetros. Todo o arsenal do Capítulo 2, Teorema de Fermat, gradiente nulo, Teste da Segunda Derivada aplica-se diretamente, como vimos no Exemplo 3.4, não há bicos, não há platôs.

Há ainda um terceiro benefício, mais sutil. Elevar ao quadrado penaliza desproporcionalmente os erros grandes, um resíduo de tamanho 2 contribui com 4, mas um de tamanho 10 contribui com 100. O método prefere, portanto, distribuir erros pequenos por todos os pontos a tolerar um erro grande em apenas um deles o que favorece ajustes equilibrados e, em condições bastante gerais, garante que o mínimo seja único.

3.4 A formulação do método

Para formular o problema com precisão, falta um último ingrediente conceitual. Dissemos no início que “a incógnita é a função”, mas procurar entre todas as funções não faz sentido (sempre existiria alguma curva esdrúxula passando por todos os pontos, com erro zero e nenhum poder explicativo). Na prática, escolhemos primeiro uma família de funções com formato definido, retas $f(x) = a + bx$, parábolas $f(x) = a + bx + cx^2$, exponenciais, etc. guiados pelo aspecto dos dados ou pela teoria do fenômeno. A função fica então determinada por um número finito de parâmetros, e é sobre eles que a otimização atua.

Definição 3.5. *Dados os pontos (x_i, y_i) , $i = 1, \dots, n$, e uma família de funções f_θ parametrizada por $\theta = (\theta_1, \dots, \theta_k)$, define-se a função de erro quadrático*

$$S(\theta) = \sum_{i=1}^n (y_i - f_\theta(x_i))^2. \quad (3.1)$$

O Método dos Mínimos Quadrados consiste em encontrar o vetor de parâmetros θ^ que minimiza $S(\theta)$.*

Note a mudança de papéis em relação ao Cálculo usual, os dados (x_i, y_i) , que pareciam as variáveis, são constantes fixadas; as variáveis de verdade são os parâmetros θ . O problema (3.1) é, assim, exatamente um problema de otimização do tipo estudado no Capítulo 2, em geral com várias variáveis, uma para cada parâmetro, e a condição necessária de mínimo é

$$\nabla_\theta S(\theta^*) = 0,$$

isto é, um sistema de k equações (uma derivada parcial nula para cada parâmetro). O próximo

capítulo especializa essa condição para dois modelos concretos, o linear e o exponencial e leva o método até dados reais.

Capítulo 4

O Método em Ação: Ajustes Linear e Exponencial

Nos capítulos anteriores, o Método dos Mínimos Quadrados foi construído em dois tempos: o Capítulo 2 desenvolveu as ferramentas de otimização, ponto crítico, gradiente nulo, Teste da Segunda Derivada; e o Capítulo 3 motivou a escolha da soma dos quadrados dos resíduos como critério de ajuste e formalizou o problema como a minimização de $S(\theta) = \sum_{i=1}^n (y_i - f_\theta(x_i))^2$, cuja condição necessária de mínimo é $\nabla_\theta S(\theta^*) = 0$. Aqui esse programa é executado.

Na Seção 4.1, parte-se do modelo mais simples que é o ajuste linear, $f(x) = a + bx$. Para ele, $\theta = (a, b)$ e a condição $\nabla_{(a,b)} S = 0$ gera um sistema de duas equações com duas incógnitas, resolvido de forma fechada e ilustrado com exemplos didáticos. Na Seção 4.2, a mesma lógica é levada ao ajuste exponencial e aplicada a dados reais, os registros de casos confirmados de COVID-19 no Brasil em 2020 [10], evidenciando que o método é capaz de descrever fenômenos concretos e subsidiar análises de interesse prático. As discussões e exemplos apresentados neste capítulo baseiam-se em [5] e [8], com suporte nas ferramentas de cálculo apresentadas em [3], [4], [6] e [7].

4.1 Ajuste Linear

O modelo mais simples dentro da estrutura desenvolvida no Capítulo 3 é a família das *retas*, $f(x) = a + bx$, parametrizada por $\theta = (a, b) \in \mathbb{R}^2$. Ajustar linearmente n pontos $(x_1, y_1), \dots, (x_n, y_n)$ significa encontrar o par (a^*, b^*) que minimiza a função de erro quadrático

$$S(a, b) = \sum_{i=1}^n (y_i - a - bx_i)^2, \quad (4.1)$$

caso particular de $S(\theta)$ definido na Seção 3.4, com $f_\theta(x) = a + bx$. Ajustar a nuvem de pontos por uma reta é, portanto, resolver o problema de otimização $\min_{(a,b) \in \mathbb{R}^2} S(a, b)$.

4.1.1 As equações normais

Pela condição necessária de mínimo estabelecida no Capítulo 2 (gradiente nulo), o par minimizador (a^*, b^*) satisfaz

$$\nabla_{(a,b)} S(a^*, b^*) = \mathbf{0}, \quad \text{isto é,} \quad \frac{\partial S}{\partial a}(a^*, b^*) = 0 \quad \text{e} \quad \frac{\partial S}{\partial b}(a^*, b^*) = 0.$$

Calculamos cada derivada parcial explicitamente, passo a passo.

Derivada em relação a a :

$$\frac{\partial S}{\partial a} = \sum_{i=1}^n 2(y_i - a - bx_i)(-1) = -2 \sum_{i=1}^n (y_i - a - bx_i).$$

Igualando a zero e dividindo por -2 :

$$\sum_{i=1}^n y_i - na - b \sum_{i=1}^n x_i = 0. \quad (4.2)$$

Derivada em relação a b :

$$\frac{\partial S}{\partial b} = \sum_{i=1}^n 2(y_i - a - bx_i)(-x_i) = -2 \sum_{i=1}^n x_i(y_i - a - bx_i).$$

Igualando a zero e dividindo por -2 :

$$\sum_{i=1}^n x_i y_i - a \sum_{i=1}^n x_i - b \sum_{i=1}^n x_i^2 = 0. \quad (4.3)$$

Reorganizando (4.2) e (4.3) com as incógnitas a e b à esquerda, obtém-se o sistema de equações:

$$\begin{cases} na + \left(\sum_{i=1}^n x_i\right)b = \sum_{i=1}^n y_i, \\ \left(\sum_{i=1}^n x_i\right)a + \left(\sum_{i=1}^n x_i^2\right)b = \sum_{i=1}^n x_i y_i. \end{cases} \quad (4.4)$$

Em notação matricial, (4.4) se escreve como $A \begin{pmatrix} a \\ b \end{pmatrix} = \mathbf{v}$, com

$$A = \begin{pmatrix} n & \sum_{i=1}^n x_i \\ \sum_{i=1}^n x_i & \sum_{i=1}^n x_i^2 \end{pmatrix}, \quad \mathbf{v} = \begin{pmatrix} \sum_{i=1}^n y_i \\ \sum_{i=1}^n x_i y_i \end{pmatrix}. \quad (4.5)$$

Unicidade da solução. O determinante da matriz A em (4.5) é

$$\Delta = n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2. \quad (4.6)$$

Denotando $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$, vale a identidade

$$\Delta = n \sum_{i=1}^n (x_i - \bar{x})^2. \quad (4.7)$$

De fato, expandindo o quadrado em (4.7) e usando $\sum_{i=1}^n x_i = n\bar{x}$:

$$\begin{aligned} n \sum_{i=1}^n (x_i - \bar{x})^2 &= n \sum_{i=1}^n (x_i^2 - 2\bar{x} x_i + \bar{x}^2) \\ &= n \sum_{i=1}^n x_i^2 - 2n\bar{x} \sum_{i=1}^n x_i + n^2 \bar{x}^2 \\ &= n \sum_{i=1}^n x_i^2 - 2n\bar{x}(n\bar{x}) + n^2 \bar{x}^2 \\ &= n \sum_{i=1}^n x_i^2 - n^2 \bar{x}^2 = n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2 = \Delta. \end{aligned}$$

Como $\Delta = n \sum_{i=1}^n (x_i - \bar{x})^2$ é uma soma de quadrados multiplicada por $n > 0$, conclui-se que $\Delta \geq 0$, com $\Delta = 0$ somente quando $x_i = \bar{x}$ para todo i , ou seja, quando todos os x_i são iguais. Portanto $\Delta > 0$ sempre que os pontos $x_1, \dots, x_n$ não sejam todos iguais, condição satisfeita em qualquer conjunto de dados com variação horizontal. Nesse caso, as colunas de A são linearmente independentes, A é inversível, e o sistema (4.4) possui solução única [12, p. 76].

Aplicando a Regra de Cramer [12, p. 77, 78] ao sistema (4.4):

$$b = \frac{n \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \cdot \sum_{i=1}^n y_i}{\Delta}, \quad (4.8)$$

$$a = \frac{\sum_{i=1}^n y_i - b \sum_{i=1}^n x_i}{n} = \bar{y} - b\bar{x}, \quad (4.9)$$

onde $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$. A equação (4.9) revela uma propriedade geométrica importante, a reta ajustada passa necessariamente pelo centróide $(\bar{x}, \bar{y})$ da nuvem de pontos [5].

4.1.2 Exemplos resolvidos

Exemplo 4.1. Comprimento de mola em função da força aplicada.

Um experimento registrou o comprimento total y (em cm) de uma mola submetida a diferentes forças x (em N). Os dados coletados estão na Tabela 4.1.

Tabela 4.1. Comprimento da mola em função da força aplicada.

x_i (N)	1	2	3	4	5
y_i (cm)	12,5	14,3	16,8	18,5	20,9

Fonte: Adaptado de [5].

Para aplicar as fórmulas (4.8) e (4.9), organizam-se os somatórios auxiliares na Tabela 4.2.

Tabela 4.2. Somatórios auxiliares — experimento da mola.

i	x_i	y_i	x_i^2	$x_i y_i$
1	1	12,5	1	12,5
2	2	14,3	4	28,6
3	3	16,8	9	50,4
4	4	18,5	16	74,0
5	5	20,9	25	104,5
Σ	15	83,0	55	270,0

Fonte: Elaboração própria (2026).

Com $n = 5$, o determinante (4.6) vale:

$$\Delta = 5 \cdot 55 - 15^2 = 275 - 225 = 50.$$

Substituindo em (4.8):

$$b = \frac{5 \cdot 270,0 - 15 \cdot 83,0}{50} = \frac{1350,0 - 1245,0}{50} = \frac{105,0}{50} = 2,10.$$

E em (4.9):

$$a = \frac{83,0 - 2,10 \cdot 15}{5} = \frac{83,0 - 31,5}{5} = \frac{51,5}{5} = 10,30.$$

A reta de mínimos quadrados é:

$$f(x) = 10,30 + 2,10x. \quad (4.10)$$

A Lei de Hooke estabelece que a deformação de uma mola é diretamente proporcional à força aplicada [3]. Em termos matemáticos, se Δl denota o alongamento da mola e x a força aplicada, então:

$$\Delta l = \frac{1}{k}x, \quad (4.11)$$

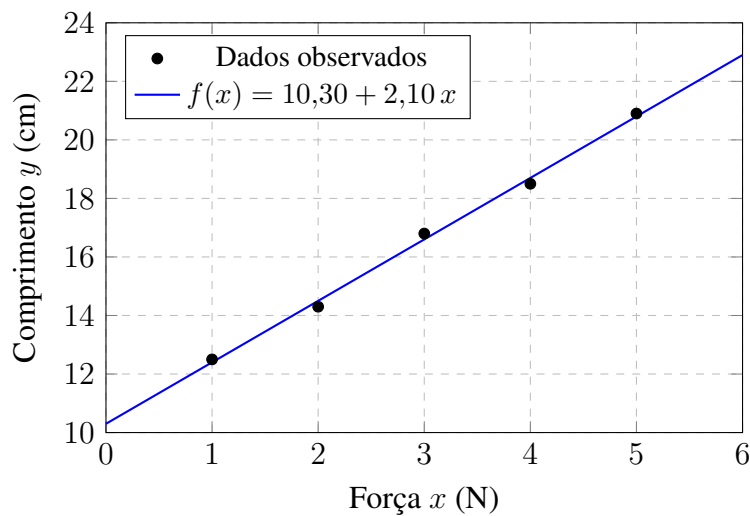
onde k (em N/cm) é a constante elástica da mola, que mede sua rigidez. Como o comprimento total é a soma do comprimento natural y_0 com o alongamento, tem-se:

$$y = y_0 + \frac{1}{k}x,$$

que é exatamente um modelo linear $y = a + bx$, com $a = y_0$ e $b = 1/k$. Os valores obtidos pelo MMQ confirmam essa estrutura: $a = 10,30$ cm estima o comprimento natural da mola (sem carga), e $b = 2,10$ cm/N estima o inverso da constante elástica, ou seja, $k \approx 0,48$ N/cm.

A Figura 4.1 exibe o diagrama de dispersão e a reta ajustada.

Figura 4.1. Diagrama de dispersão e reta ajustada — experimento da mola.



Fonte: Elaboração própria (2026).

Exemplo 4.2. Produção agrícola em função da precipitação mensal.

Uma estação meteorológica registrou a precipitação mensal x (em mm) e a produção agrícola y (em kg) de uma pequena lavoura ao longo de cinco meses, conforme a Tabela 4.3.

Tabela 4.3. Produção agrícola em função da precipitação mensal.

x_i (mm)	10	20	30	40	50
y_i (kg)	10	20	25	30	40

Fonte: Adaptado de [5].

Os somatórios auxiliares estão na Tabela 4.4.

Tabela 4.4. Somatórios auxiliares — produção agrícola.

i	x_i	y_i	x_i^2	$x_i y_i$
1	10	10	100	100
2	20	20	400	400
3	30	25	900	750
4	40	30	1600	1200
5	50	40	2500	2000
Σ	150	125	5500	4450

Fonte: Elaboração própria (2026).

Com $n = 5$:

$$\Delta = 5 \cdot 5500 - 150^2 = 27500 - 22500 = 5000.$$

Aplicando (4.8):

$$b = \frac{5 \cdot 4450 - 150 \cdot 125}{5000} = \frac{22250 - 18750}{5000} = \frac{3500}{5000} = 0,70.$$

E (4.9):

$$a = \frac{125 - 0,70 \cdot 150}{5} = \frac{125 - 105}{5} = \frac{20}{5} = 4,00.$$

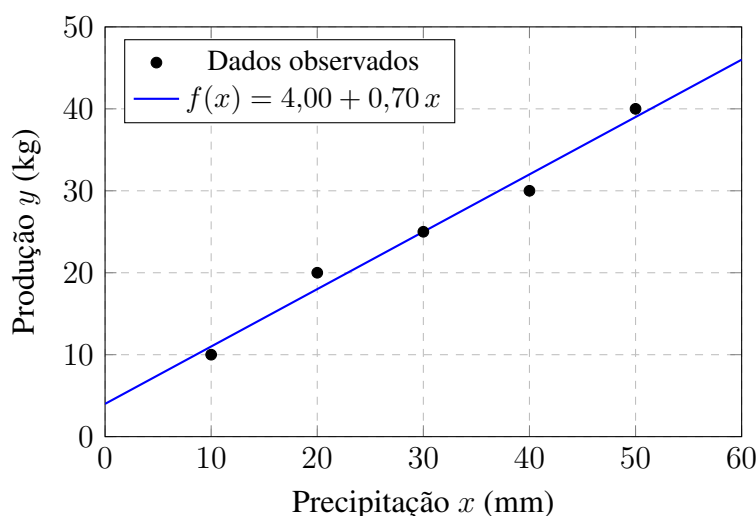
A reta ajustada é:

$$f(x) = 4,00 + 0,70x. \quad (4.12)$$

Interpretação. O valor $a = 4,00$ kg pode ser interpretado como a produção mínima estimada em condições de precipitação quase nula (irrigação residual ou umidade armazenada no solo); $b = 0,70$ kg/mm indica que cada milímetro adicional de chuva está associado, em média, a um acréscimo de 0,70 kg na colheita. O modelo é adequado na faixa observada, mas pode subestimar a produção para precipitações muito elevadas, onde fatores como solo encharcado introduzem não linearidades. A Figura 4.2 exibe os dados e a reta ajustada.

Os dois exemplos mostram que o mesmo procedimento algébrico, resolver o sistema

Figura 4.2. Diagrama de dispersão e reta ajustada — produção agrícola.



Fonte: Elaboração própria (2026).

(4.4), descreve fenômenos de domínios inteiramente distintos. A estrutura linear $f(x) = a + bx$ é suficientemente flexível para aproximar relações proporcionais em contextos variados; ao mesmo tempo, é o modelo mais parcimonioso, com apenas dois parâmetros a determinar.

Há, contudo, situações em que a proporcionalidade linear é claramente inadequada. Fenômenos de crescimento acelerado, como a proliferação de um vírus em fase inicial de epidemia, exibem comportamento de duplicação periódica, incompatível com uma reta. Para esses casos, a Seção 4.2 mostra que uma transformação logarítmica converte o ajuste exponencial num ajuste linear, tornando as fórmulas (4.8) e (4.9) diretamente reutilizáveis sobre os dados transformados, com aplicação aos registros de COVID-19 no Brasil em 2020 [10, 13].

4.2 Ajuste Exponencial

No início de 2020, o mundo foi surpreendido pela rápida disseminação de um novo coronavírus, o SARS-CoV-2, responsável pela doença denominada COVID-19. Em 11 de março de 2020, a Organização Mundial da Saúde declarou estado de pandemia [11], reconhecendo a propagação global e sustentada do vírus. No Brasil, os primeiros casos confirmados foram registrados ao final de fevereiro de 2020 e, a partir da semana epidemiológica 13 (22 a 28 de março), os boletins do Ministério da Saúde passaram a divulgar semanalmente o número de casos novos [10]. Semana após semana, esse número crescia de forma acelerada, levantando uma questão urgente, se nada mudasse, quantos casos novos seriam registrados nas semanas seguintes.

Responder a essa pergunta por meios matemáticos exige construir um modelo que descreva o padrão de crescimento observado e, em seguida, estendê-lo para além dos dados dispo-

níveis. É precisamente esse o papel do ajuste de curvas pelo Método dos Mínimos Quadrados. Nesta seção, utilizamos os registros semanais de casos novos de COVID-19 no Brasil correspondentes às semanas epidemiológicas 13 a 29 de 2020, dezessete semanas consecutivas, para ajustar um modelo exponencial e projetar o que teria ocorrido na semana 30, caso o ritmo de crescimento não tivesse se alterado [13].

No início de uma epidemia, cada pessoa infectada transmite o vírus a um número médio de $R_0 > 1$ outras pessoas suscetíveis. Após a primeira geração de transmissões há R_0 casos novos; após a segunda, R_0^2 ; após k gerações, R_0^k . Esse crescimento multiplicativo, em que a quantidade se multiplica por um mesmo fator a cada período, é chamado de crescimento *exponencial*. Em tempo contínuo, escreve-se

$$y(t) = A e^{\beta t}, \quad A > 0, \beta > 0,$$

onde A representa uma quantidade de referência no instante inicial e β é a taxa de crescimento por unidade de tempo. Esse padrão se mantém enquanto houver muitas pessoas suscetíveis e o comportamento da população não mudar; ao longo do tempo, fatores como imunidade adquirida, medidas de distanciamento e mudanças de comportamento tendem a reduzir β e romper a trajetória exponencial.

Reconhecida essa dinâmica, o desafio seguinte é determinar os parâmetros A e β a partir dos dados reais. O modelo exponencial $y = A e^{\beta x}$, contudo, não é linear nesses parâmetros, de modo que as fórmulas deduzidas na Seção 4.1 não se aplicam diretamente. A solução é *transformar* os dados de modo a obter uma relação linear. Aplicando o logaritmo natural a ambos os membros:

$$\ln y = \ln A + \beta x.$$

Definindo as variáveis auxiliares $Y = \ln y$ e $a = \ln A$, obtemos

$$Y = a + \beta x, \tag{4.13}$$

que é exatamente uma reta com intercepto a e inclinação β . Isso significa que, se calcularmos $Y_i = \ln y_i$ para cada observação e ajustarmos uma reta aos pares (x_i, Y_i) usando as fórmulas (4.8) e (4.9) da Seção 4.1, obtemos os valores de β e a . O parâmetro original A é então recuperado por $A = e^a$.

O intercepto a da reta linearizada é obtido exatamente pelas mesmas fórmulas da Seção 4.1, por isso mantém o mesmo símbolo. Já a inclinação é denotada por β , e não por b , para lembrar que ela representa a taxa de crescimento do modelo exponencial; e o parâmetro original, não linearizado, é denotado por A (maiúsculo), para distingui-lo do intercepto a , ao qual se relaciona por $A = e^a$.

4.2.1 Aplicação: COVID-19 no Brasil, semanas 13 a 29 de 2020

Exemplo 4.3. Queremos ajustar um modelo exponencial $\hat{y}(x) = A e^{\beta x}$ ao número semanal de casos novos de COVID-19 no Brasil, usando as semanas 13 a 29 de 2020 como dados de ajuste, e em seguida projetar quantos casos novos seriam esperados na semana 30, caso o crescimento observado nas semanas anteriores continuasse sem alteração.

A Tabela 4.5 apresenta, para cada semana (indexada por $x = 1, 2, \dots, 17$), o número de casos novos registrado (y_i), o logaritmo natural $Y_i = \ln y_i$ e o produto $x_i Y_i$. Os valores de Y_i e $x_i Y_i$ foram arredondados a duas casas decimais.

Tabela 4.5. Dados de COVID-19 no Brasil (semanas 13 a 29 de 2020) e valores transformados para o ajuste exponencial

x_i	Semana	y_i	$Y_i = \ln y_i$	$x_i Y_i$
1	13/2020	3 912	8,27	8,27
2	14/2020	6 386	8,76	17,52
3	15/2020	10 520	9,26	27,78
4	16/2020	15 921	9,68	38,70
5	17/2020	22 234	10,01	50,05
6	18/2020	37 586	10,53	63,21
7	19/2020	59 380	10,99	76,94
8	20/2020	77 203	11,25	90,03
9	21/2020	114 107	11,64	104,80
10	22/2020	151 191	11,93	119,26
11	23/2020	174 406	12,07	132,76
12	24/2020	177 668	12,09	145,05
13	25/2020	217 065	12,29	159,74
14	26/2020	246 088	12,41	173,79
15	27/2020	263 337	12,48	187,22
16	28/2020	262 846	12,48	199,67
17	29/2020	235 010	12,37	210,25
$\sum$			188,52	1 805,05

Fonte: Elaboração própria (2026), com dados de [10].

Calculando as somas necessárias. Com $n = 17$ semanas e $x_i = 1, 2, \dots, 17$, as somas $\sum x_i$ e $\sum x_i^2$ resultam das fórmulas de progressões aritméticas:

$$\sum_{i=1}^{17} x_i = \frac{17 \cdot 18}{2} = 153, \quad \sum_{i=1}^{17} x_i^2 = \frac{17 \cdot 18 \cdot 35}{6} = 1\,785.$$

As demais somas são lidas diretamente da Tabela 4.5: $\sum Y_i = 188,52$ e $\sum x_i Y_i = 1\,805,05$.

Determinando β e a . O determinante da matriz normal, calculado pela fórmula (4.6), é:

$$\Delta = n \sum x_i^2 - \left(\sum x_i \right)^2 = 17 \cdot 1\,785 - 153^2 = 30\,345 - 23\,409 = 6\,936.$$

Aplicando a fórmula (4.8) aos pares transformados (x_i, Y_i) , encontramos a inclinação da reta

linearizada:

$$\begin{aligned}
 \beta &= \frac{n \sum x_i Y_i - \sum x_i \cdot \sum Y_i}{\Delta} \\
 &= \frac{17 \cdot 1\,805,05 - 153 \cdot 188,52}{6\,936} \\
 &= \frac{30\,685,85 - 28\,843,56}{6\,936} \\
 &= \frac{1\,842,29}{6\,936} \\
 &\approx 0,266.
 \end{aligned}$$

E pela fórmula (4.9), o intercepto:

$$a = \frac{\sum Y_i - \beta \sum x_i}{n} = \frac{188,52 - 0,266 \cdot 153}{17} = \frac{188,52 - 40,70}{17} = \frac{147,82}{17} \approx 8,698.$$

O modelo ajustado. Lembrando que $a = \ln A$, temos $A = e^a = e^{8,698}$. O modelo exponencial de melhor ajuste aos dados é, portanto:

$$\hat{y}(x) = e^{8,698} \cdot e^{0,266x}. \quad (4.14)$$

Resolução: O coeficiente $\beta \approx 0,266$ representa a taxa de crescimento logarítmica por semana. Como $e^{0,266} \approx 1,30$, o número de casos novos crescia, em média, cerca de 30% a cada semana durante o período analisado, o equivalente a dobrar a cada $\ln 2 / 0,266 \approx 2,6$ semanas. $\square$

A Figura 4.3 mostra os pares (x_i, y_i) junto com a curva (4.14). Nota-se que o modelo acompanha bem a tendência de crescimento acelerado observada nas primeiras semanas, confirmando visualmente a adequação do ajuste exponencial para esse período.

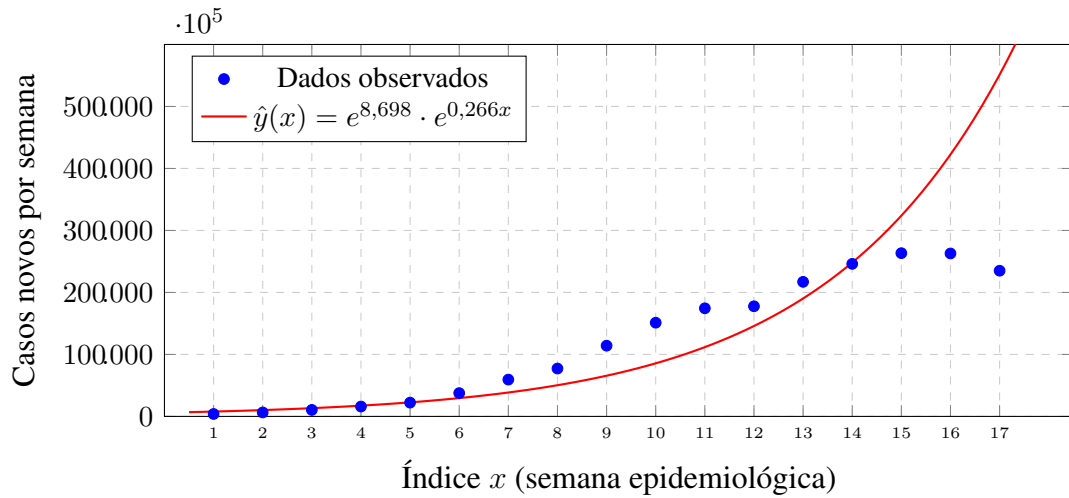
Projeção para a semana 30 (se nada tivesse mudado). A semana 30 de 2020 corresponde a $x = 18$. Substituindo em (4.14):

$$\hat{y}(18) = e^{8,698} \cdot e^{0,266 \times 18} = e^{8,698} \cdot e^{4,788} \approx 690\,278.$$

O modelo prevê, portanto, que se a dinâmica de crescimento das semanas 13 a 29 continuasse inalterada, seriam registrados aproximadamente **690 278** casos novos na semana 30 de 2020.

O Ministério da Saúde registrou, contudo, 319 653 casos novos naquela semana [10], menos da metade do valor projetado pelo modelo. A diferença entre previsão e realidade ($690\,278 - 319\,653 \approx 370\,625$ casos a menos que o esperado) traduz, numericamente, o efeito

Figura 4.3. Casos novos de COVID-19 no Brasil (semanas 13 a 29 de 2020) e curva exponencial ajustada pelo MMQ.



Fonte: Elaboração própria (2026), com dados de [10].

combinado das medidas de distanciamento social, das restrições de circulação e da mudança de comportamento da população naquele período. O modelo exponencial cumpriu seu papel: não de descrever o futuro com exatidão, mas de mostrar o que teria acontecido caso nada tivesse mudado, e por contraste, evidenciar o quanto as intervenções efetivamente modificaram a trajetória da pandemia.

Para avaliar o desempenho preditivo do modelo em um horizonte mais próximo, ajusta-se agora o modelo exponencial às nove semanas disponíveis ($x = 1, \dots, 9$, correspondentes às semanas 13 a 21 de 2020) e usa-se o resultado para estimar a semana 22 ($x = 10$), cujo valor real registrado pelo Ministério da Saúde é 151 191 casos [10]. Para distinguir dos parâmetros do ajuste completo (dezessete semanas), os novos parâmetros são denotados $\tilde{\beta}$ e $\tilde{\alpha}$.

Os valores $\tilde{Y}_i = \ln y_i$ e os produtos $x_i \tilde{Y}_i$ para essa janela são os mesmos da Tabela 4.5; a Tabela 4.6 os reproduz de forma compacta.

Tabela 4.6. Janela de ajuste restrita: semanas 13 a 21 de 2020 ($x = 1, \dots, 9$).

x_i	Semana	y_i	$\tilde{Y}_i = \ln y_i$	$x_i \tilde{Y}_i$
1	13/2020	3 912	8,27	8,27
2	14/2020	6 386	8,76	17,52
3	15/2020	10 520	9,26	27,78
4	16/2020	15 921	9,68	38,70
5	17/2020	22 234	10,01	50,05
6	18/2020	37 586	10,53	63,21
7	19/2020	59 380	10,99	76,94
8	20/2020	77 203	11,25	90,03
9	21/2020	114 107	11,64	104,80
Σ			90,39	477,30

Fonte: Elaboração própria (2026), com dados de [10].

Com $n = 9$, as somas $\sum x_i$ e $\sum x_i^2$ resultam das fórmulas de progressões aritméticas:

$$\sum_{i=1}^9 x_i = \frac{9 \cdot 10}{2} = 45, \quad \sum_{i=1}^9 x_i^2 = \frac{9 \cdot 10 \cdot 19}{6} = 285.$$

O determinante é:

$$\tilde{\Delta} = 9 \cdot 285 - 45^2 = 2\,565 - 2\,025 = 540.$$

Aplicando as fórmulas (4.8) e (4.9) aos pares $(x_i, \tilde{Y}_i)$:

$$\begin{aligned} \tilde{\beta} &= \frac{9 \cdot 477,30 - 45 \cdot 90,39}{540} = \frac{4\,295,70 - 4\,067,55}{540} = \frac{228,15}{540} \approx 0,423; \\ \tilde{\alpha} &= \frac{90,39 - 0,423 \cdot 45}{9} = \frac{90,39 - 19,04}{9} = \frac{71,35}{9} \approx 7,928. \end{aligned}$$

A previsão para $x = 10$ (semana 22) é, portanto:

$$\hat{y}(10) = e^{7,928} \cdot e^{0,423 \times 10} = e^{7,928+4,230} = e^{12,158} \approx 190\,612.$$

O valor real registrado na semana 22 foi 151 191 casos, enquanto o modelo previu 190 612, uma diferença absoluta de aproximadamente 39 421 casos. Para contextualizar essa diferença, em 2020 o Brasil contava com cerca de 212 000 000 de habitantes [9]: a discrepância entre previsão e realidade representa menos de 0,02% da população total do país, um valor numericamente ínfimo na escala demográfica brasileira. Isso demonstra que o modelo exponencial ajustado pelo Método dos Mínimos Quadrados é capaz de estimar a ordem de grandeza dos casos novos com precisão relevante para fins de planejamento e tomada de decisão, cumprindo seu objetivo de descrever matematicamente a dinâmica da epidemia e projetar cenários a partir de dados passados.

A diferença expressiva entre a previsão de 690 278 casos para a semana 30 e o valor real de 319 653 casos revela uma limitação estrutural do modelo exponencial, a suposição de que a taxa de crescimento β permanece constante ao longo do tempo. Em epidemias reais, porém, essa taxa é continuamente influenciada por fatores que o modelo não capta. Medidas de distanciamento social e restrições de circulação reduzem o contato entre pessoas e diminuem a velocidade de transmissão; mudanças nos critérios de testagem e de notificação alteram o volume de casos confirmados; o surgimento de variantes do vírus pode modificar a transmissibilidade; e, com o avanço das campanhas de vacinação, a proporção de indivíduos suscetíveis na população decresce progressivamente. Ao se afastar temporalmente dos dados de ajuste, a extrapolação exponencial acumula todos esses efeitos não modelados e tende a superestimar a realidade.

O ajuste exponencial é adequado apenas enquanto a epidemia se encontra em sua fase

inicial de crescimento acelerado, quando a proporção de indivíduos suscetíveis ainda é elevada e as condições de transmissão permanecem aproximadamente constantes. Superada essa fase, a taxa de novos casos tende a desacelerar e a curva começa a se estabilizar, comportamento incompatível com o crescimento ilimitado previsto pelo modelo exponencial.

4.2.2 Outros modelos não lineares: ajustes logístico e de potência

O modelo exponencial explorado na seção anterior é apenas um entre vários modelos não lineares aos quais o Método dos Mínimos Quadrados pode ser adaptado. A escolha do modelo adequado depende fundamentalmente da natureza do fenômeno e do padrão exibido pelos dados, não existe um modelo único válido para todas as situações. A seguir, apresentam-se resumidamente dois outros modelos não lineares relevantes, o logístico e o de potência, e discute-se por que, no contexto deste trabalho, o exponencial é a escolha natural.

Modelo logístico

O modelo logístico descreve crescimentos que apresentam uma limitação natural: crescem rapidamente no início, desaceleram progressivamente e tendem a um valor máximo. Sua forma mais comum é

$$C(t) = \frac{L}{1 + A e^{-kt}}, \quad (4.15)$$

onde $C(t)$ é a quantidade acumulada no instante t ; $L > 0$ é o *valor limite* ou capacidade de saturação; $A > 0$ está relacionado à condição inicial; e $k > 0$ mede a velocidade de crescimento. A curva tem formato de “S”: cresce de modo quase exponencial no início, desacelera na fase intermediária e se aproxima assintoticamente de L [14, 15].

Esse comportamento é adequado quando o crescimento é limitado — por exemplo, os *casos acumulados* de uma única onda epidêmica, em que L estima o total de casos ao final da onda. Note que a variável aqui seria diferente da utilizada neste trabalho: os casos novos semanais formam uma curva com picos (sobem e depois caem), não uma curva em “S”.

O modelo (4.15) é não linear nos parâmetros L , A e k , exigindo em geral métodos numéricos para sua minimização. Contudo, se o valor de L for previamente estimado, é possível linearizá-lo: manipulando a equação e aplicando logaritmo natural, obtém-se

$$Y = \ln\left(\frac{L}{C(t)} - 1\right) = \ln A - kt,$$

que é uma reta $Y = a + bt$ com $a = \ln A$ e $b = -k$, permitindo aplicar diretamente as fórmulas da Seção 4.1 aos dados transformados.

Modelo de potência

O modelo de potência tem a forma

$$y = a x^b, \quad (4.16)$$

onde $a > 0$ e b são parâmetros a determinar [13]. Ele é adequado quando a relação entre as grandezas segue uma lei de escala, onde a variável resposta varia como uma potência da variável preditora, e não de forma aditiva ou multiplicativa constante. Aplicações incluem relações geométricas (área em função do comprimento), leis físicas, relações biológicas entre massa e metabolismo, entre outras.

Uma vantagem importante é que o modelo (4.16) pode ser linearizado por logaritmos: tomando $\ln$ em ambos os membros,

$$\ln y = \ln a + b \ln x.$$

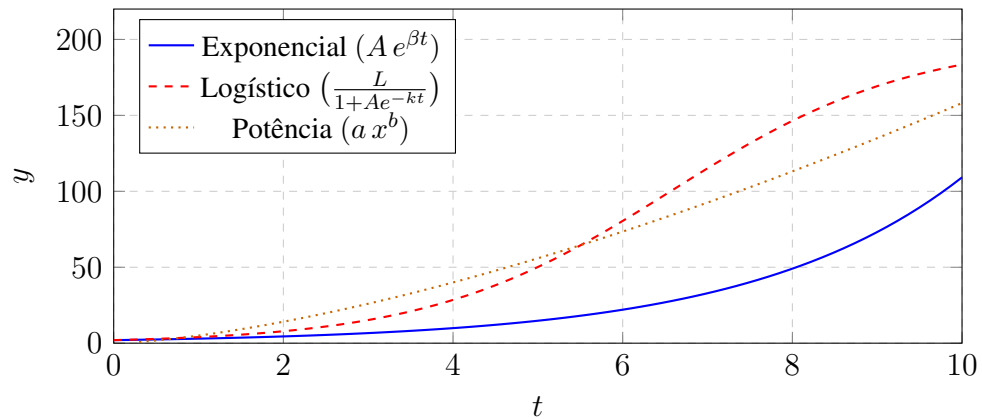
Definindo $X = \ln x$, $Y = \ln y$ e $\alpha = \ln a$, chega-se ao modelo linear $Y = \alpha + bX$, ao qual se aplicam diretamente as fórmulas da Seção 4.1. Após obter α e b pelo MMQ, recupera-se $a = e^\alpha$. Para que a transformação seja válida, é necessário que $x_i > 0$ e $y_i > 0$ para todo i [13].

Ao linearizar por logaritmos, tanto no modelo de potência quanto no logístico (quando L é conhecido), o MMQ minimiza os erros na escala logarítmica, e não na escala original. Em contextos onde os resíduos relativos importam mais do que os absolutos, isso é matematicamente coerente; em outros, pode introduzir uma pequena diferença em relação a um ajuste direto na escala original. Para trabalhos introdutórios, a técnica é amplamente aceita [13].

A Figura 4.4 ilustra curvas típicas dos três modelos. Na fase inicial de uma epidemia, os casos novos crescem de forma multiplicativa, cada pessoa infectada transmite o vírus a outras, gerando crescimento proporcional ao número atual de casos, sem que haja ainda limitação por imunidade de rebanho ou saturação de suscetíveis. Esse padrão, com taxa de crescimento aproximadamente constante na escala logarítmica e sem saturação visível, corresponde precisamente ao modelo exponencial. O modelo logístico seria mais apropriado para os casos *acumulados* de uma onda completa, e o de potência, para relações de escala entre grandezas distintas. A escolha do modelo, portanto, não é arbitrária: ela deve casar com a estrutura matemática do fenômeno observado.

Ao longo deste capítulo, o Método dos Mínimos Quadrados mostrou-se capaz de capturar tanto tendências lineares quanto exponenciais em dados reais, sempre a partir do mesmo procedimento analítico, minimizar a soma dos quadrados dos resíduos para obter os parâmetros de melhor ajuste. No ajuste linear, modelaram-se situações de proporcionalidade direta, a deformação de uma mola e a produção agrícola; no ajuste exponencial, tratou-se de um fenômeno

Figura 4.4. Curvas típicas dos modelos exponencial, logístico e de potência (valores ilustrativos).



Fonte: Elaboração própria (2026).

de multiplicação contínua, o crescimento de uma epidemia. Em ambos os casos, o método converteu dados dispersos em uma expressão matemática interpretável, capaz de quantificar tendências e embasar projeções.

Capítulo 5

Considerações Finais

Este trabalho buscou responder à questão norteadora proposta na Introdução, isto é, de que modo o Método dos Mínimos Quadrados pode ser utilizado para modelar matematicamente fenômenos reais e quais são os alcances e os limites desses modelos. Os capítulos desenvolvidos construíram uma resposta progressiva a essa pergunta, partindo dos fundamentos do cálculo, passando pela dedução do critério de ajuste e chegando à aplicação a dados reais. Com isso, conclui-se que os objetivos específicos anunciados foram atendidos.

O primeiro objetivo específico, revisar os fundamentos do cálculo que embasam o método, foi cumprido no Capítulo 2, que reuniu as noções de máximo e mínimo, ponto crítico, Teorema de Fermat, gradiente e Teste da Segunda Derivada, exatamente os instrumentos necessários para formular e resolver o problema de minimização que está no centro do método. O segundo objetivo, deduzir o critério de minimização e o sistema de equações normais, concretizou-se no Capítulo 3, no qual se definiu o resíduo de cada observação, justificou-se a escolha de minimizar a soma dos quadrados e, a partir da condição de gradiente nulo, obtiveram-se as fórmulas dos coeficientes de melhor ajuste.

O terceiro objetivo, aplicar as fórmulas a dados reais nos casos linear e exponencial, materializou-se no Capítulo 4. No ajuste linear, os coeficientes obtidos revelaram sentido concreto. No experimento da mola, a reta $f(x) = 10,30 + 2,10x$ estima em cerca de 10,30 cm o comprimento natural da mola e associa cada Newton de força a um alongamento médio de 2,10 cm, o que corresponde a uma constante elástica de aproximadamente 0,48 N/cm, em coerência com a Lei de Hooke. No exemplo da produção agrícola, a reta $f(x) = 4,00 + 0,70x$ indica um acréscimo médio de 0,70 kg de colheita por milímetro de chuva. No ajuste exponencial, o modelo da forma $y = A e^{\beta x}$, calibrado com os casos novos de COVID-19 no Brasil nas semanas 13 a 29 de 2020, descreveu um crescimento médio próximo de 30% por semana, com duplicação a cada 2,6 semanas aproximadamente.

O quarto objetivo, discutir os alcances e os limites dos modelos, foi atendido pela própria aplicação à COVID-19, que expôs com clareza o que o método consegue fazer e onde ele encontra fronteiras. A projeção contrafactual para a semana 30, da ordem de 690 mil casos pelo

modelo, confrontada com o valor real registrado, próximo de 320 mil casos, evidenciou o efeito das medidas de distanciamento e da mudança de comportamento da população, fatores que o modelo não incorpora. Já a validação preditiva de curto prazo, feita com a janela das semanas 13 a 21 para estimar a semana 22, aproximou-se do valor real, com diferença que representa menos de 0,02% da população brasileira de 2020, o que confirma a pertinência do ajuste dentro da fase de crescimento acelerado. Reconhece-se, contudo, um limite estrutural. O modelo exponencial pressupõe taxa de crescimento constante e descreve bem apenas a fase inicial da epidemia, perdendo aderência quando o crescimento desacelera por efeito da imunidade adquirida, do distanciamento, das mudanças nos critérios de testagem e da vacinação. Essa limitação não decorre de falha do método, mas da escolha do modelo, uma vez que o exponencial representa crescimentos sem saturação.

Esse reconhecimento abre desdobramentos naturais para trabalhos futuros, entre eles o ajuste logístico aplicado a séries de casos acumulados de uma onda completa, no qual o valor limite estima o total ao final da onda, o ajuste de potência para relações de escala entre grandezas, e a comparação sistemática entre diferentes modelos para um mesmo conjunto de dados, com o intuito de avaliar qual deles melhor se adéqua à natureza do fenômeno. Tais extensões preservam o mesmo arcabouço analítico empregado neste trabalho, a minimização da soma dos quadrados dos resíduos, variando apenas o modelo e, quando necessário, a transformação aplicada aos dados para linearizá-lo. Conclui-se, assim, que o Método dos Mínimos Quadrados articula fundamentação teórica rigorosa e aplicação a dados de relevância social, oferecendo um caminho ao mesmo tempo preciso e acessível para descrever fenômenos, quantificar tendências e, com consciência de seus próprios limites, embasar projeções.

Referências

- [1] GIBALDI, M.; PERRIER, D. **Pharmacokinetics**. 2. ed. New York: Marcel Dekker, 1982.
- [2] APPLGATE, D. L.; BIXBY, R. E.; CHVÁTAL, V.; COOK, W. J. **The Traveling Salesman Problem: A Computational Study**. Princeton: Princeton University Press, 2006.
- [3] STEWART, James. **Cálculo**. 8. ed. São Paulo: Cengage Learning, 2017.
- [4] THOMAS, George B.; WEIR, Maurice D.; HASS, Joel R. **Cálculo**. 12. ed. São Paulo: Pearson, 2012.
- [5] ALMEIDA, Renato Neves de. **O método dos mínimos quadrados: estudo e aplicações para o ensino médio**. 2015. Dissertação (Mestrado Profissional – PROFMAT) – Universidade Estadual do Norte Fluminense Darcy Ribeiro, Campos dos Goytacazes, 2015.
- [6] ANTON, Howard; BIVENS, Irl; DAVIS, Stephen. **Cálculo**. 10. ed. Porto Alegre: Bookman, 2014.
- [7] GUIDORIZZI, Hamilton Luiz. **Um curso de cálculo: volume 2**. 6. ed. Rio de Janeiro: LTC, 2019.
- [8] PEDROSA, Diogo Pinheiro Fernandes. **Ajuste de curvas**. Caicó: Universidade Federal do Rio Grande do Norte, 2014.
- [9] INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE). **Projeção da população do Brasil e das Unidades da Federação**. Rio de Janeiro: IBGE, 2020. Disponível em: <<https://www.ibge.gov.br/estatisticas/sociais/populacao/9109-projecao-da-populacao.html>>. Acesso em: 19 maio 2026.

- [10] BRASIL. Ministério da Saúde. **Painel Coronavírus**. Brasília, 2020. Disponível em: <https://infoms.saude.gov.br/extensions/covid-19_html/covid-19_html.html>. Acesso em: 15 set. 2025.
- [11] ORGANIZAÇÃO MUNDIAL DA SAÚDE (OMS). **WHO Director-General's opening remarks at the media briefing on COVID-19 – 11 March 2020**. Geneva: OMS, 2020. Disponível em: <<https://www.who.int/news-room/speeches/item/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19—11-march-2020>>. Acesso em: 19 maio 2026.
- [12] BOLDRINI, José Luiz et al. **Álgebra Linear**. 3. ed. São Paulo: Harbra, 1986.
- [13] VELOSO, Eriky Abreu. **Um estudo da disseminação da COVID-19 no Brasil a partir do Método dos Mínimos Quadrados**. Orientadora: Fabiane Regina da Cunha Dantas Araújo. 2021. Trabalho de Conclusão de Curso (Bacharelado em Ciência e Tecnologia) — Universidade Federal Rural do Semi-Árido (UFERSA), Mossoró, 2021.
- [14] BASSANEZI, Rodney Carlos. **Ensino-aprendizagem com modelagem matemática**. São Paulo: Contexto, 2002.
- [15] BOYCE, William E.; DIPRIMA, Richard C. **Elementary differential equations**. 10. ed. Hoboken: John Wiley & Sons, 2012.