RENATO NEVES DE ALMEIDA

O MÉTODO DOS MÍNIMOS QUADRADOS : ESTUDO E APLICAÇÕES PARA O ENSINO MÉDIO

UNIVERSIDADE ESTADUAL DO NORTE FLUMINENSE
DARCY RIBEIRO - UENF
CAMPOS DOS GOYTACAZES - RJ

MAIO DE 2015

RENATO NEVES DE ALMEIDA

O MÉTODO DOS MÍNIMOS QUADRADOS : ESTUDO E APLICAÇÕES PARA O ENSINO MÉDIO

"Dissertação apresentada ao Centro de Ciências e Tecnologia da Universidade Estadual do Norte Fluminense Darcy Ribeiro, como parte das exigências para obtenção do título de Mestre em Matemática."

Orientador: Prof. Paulo Sérgio Dias da Silva

UNIVERSIDADE ESTADUAL DO NORTE FLUMINENSE
DARCY RIBEIRO - UENF
CAMPOS DOS GOYTACAZES - RJ
MAIO DE 2015

RENATO NEVES DE ALMEIDA

O MÉTODO DOS MÍNIMOS QUADRADOS : ESTUDO E APLICAÇÕES PARA O ENSINO MÉDIO

"Dissertação apresentada ao Centro de Ciências e Tecnologia da Universidade Estadual do Norte Fluminense Darcy Ribeiro, como parte das exigências para obtenção do título de Mestre em Matemática."

Aprovada em 28 de maio de 2015.

Comissão Examinadora:

Prof. Oscar Alfredo Paz La Torre
D.Sc. - UENF

Prof. Mikhail Petrovich Vishnevskii
D.Sc. - UENF

Prof^a. Silvia Cristina Freitas Batista
D.Sc. - IFF

Prof. Paulo Sérgio Dias da Silva
D.Sc. - UENF
(ORIENTADOR)

À memória de meu pai, que com poucas palavras e muitos exemplos me ensinou o caminho da retidão e o valor do conhecimento.

Agradecimentos

Agradeço aos meus pais e irmãos, cujas presenças físicas nunca foram necessárias para fazê-los fundamentais em todas as minhas conquistas.

Aos meus colegas de curso, com os quais dividi nesses quase dois anos e meio a sala de aula, os contratempos e as aspirações.

À minha esposa Andréia e ao meu enteado André, dois amores, dois pilares neste momento de minha vida.

A todos os professores do curso PROFMAT, em especial ao meu orientador, Mestre Paulo Dias, pela supervisão sempre inteligente, paciente e ponderada da realização deste trabalho.

Agradeço, por fim, a Ele, sabendo que sou instrumento de sua sabedoria e de seu plano divino, por me permitir viver mais esta conquista - o nosso bom Deus.

"A Matemática é o alfabeto que Deus usou para escrever o Universo."

Galileu Galilei.

Resumo

O **Método dos Mínimos Quadrados**, bem como qualquer outro procedimento empregado no ajuste de curvas, não faz parte da grade curricular do Ensino Médio, embora os recursos matemáticos necessários para implementá-lo estejam ao alcance do aluno do final (3º ano) desse segmento. Assim sendo, dados a versatilidade do método e seu potencial como ferramenta de modelagem matemática, busca-se neste trabalho apresentar uma abordagem acessível, porém sem perda do viés científico que o norteia, ao aluno do Ensino Médio. Busca-se assim não seguir a linha adotada na maior parte dos trabalhos que se propuseram a apresentar o método ao aluno deste segmento, nos quais prioriza-se o tratamento matricial e preocupa-se quase que exclusivamente com a execução do método, ofuscando a motivação que deu origem ao mesmo. O **Método dos Mínimos Quadrados** é uma modalidade dentro do contexto do Ajuste de Curvas e, como tal, para seu bom entendimento necessita de um estudo preliminar, mesmo que breve, sobre o tema, o que é feito neste trabalho. O desenvolvimento matemático do método é efetuado até chegar-se às modalidades de ajuste que o compõem (**Ajuste Linear Simples**, **Ajuste Linear Múltiplo** e **Ajuste Polinomial**), expondo as equações normais que levarão à sua execução. Havendo para a dedução do método a necessidade de conhecimento de alguns conteúdos não contextualizados no Ensino Médio, uma estratégia de abordagem voltada para tal segmento é exposta, aproveitando conteúdos familiares ao aluno e introduzindo conceitos novos de forma diferenciada e suficiente para entendimento do desenvolvimento matemático em questão. O trabalho é concluído com algumas atividades de aplicação do método. Há sugestões para a utilização de recursos computacionais, como modo de agilizar a resolução das atividades propostas e promover o incentivo ao acesso a novas tecnologias. Tais atividades foram selecionadas de modo a não somente familiarizar o aluno com a implementação do método estudado, mas a dotá-lo de capacidade de avaliação sobre a importância e o alcance do estudo aqui realizado.

Palavras-chaves: mínimos quadrados, ajuste, funções, diagrama de dispersão.

Abstract

The **Method of Least Squares**, as well as any other procedure employed in curve fitting, is not part of the curriculum of high school, although the mathematical resources needed to implement it are within the reach of the end of this year segment. Thus, given the method's versatility and its potential as a tool in mathematical modeling, we search in this work to show an accessible approach, however without loss of scientific bias that guides the high school student. Thus, we search not to follow the line adopted in most works that have proposed to show the method to the student of this segment, in hhich prioritizes the matrix treatment and is concern almost exclusively with the method's execution, overshadowing the motivation that gave rise it. The **Method of Least Squares** is a modality in the context of **Curve Fitting** and us the such, for its good understanding it needs a preliminary study, even brief, about the topic, which is made in this work. The mathematical development of the method is made until it reaches to the modalities fitting that compose it (**Simple Linear Fitting**, **Multiple Linear Fitting** and Polynomial Fitting), exposing the normal equations that will guide to the solution. Due to the need of knowledge of some topics for the deduction of the method not contextualized in high school, an strategy of approach aimed to this segment is exposed, taking advantage of familiar topics to the student and introducing new concepts in a different and suitable way for the understanding of the mathematical development in question. The work is concluded with some activities of method's applications. There are suggestions for the utilization of computational resources, in order to streamline the resolution of the proposed activities and promoting the incentive to the access of new technologies. Such activities were selected in order to make the student not only familiarize himself with the method's implementation, but as well adopt it due the capacity of avaliation about the importance and the reach the study made here.

Key-words:Least Squares, fitting, functions, scattergram.

Lista de ilustrações

Lista de tabelas

Sumário

Introdução

Na busca por um tema para elaboração de minha Dissertação de Mestrado, desde o início assumi a intenção de transitar pelo universo da Matemática Aplicada, entusiasta que sou, por acreditar no seu potencial pedagógico, da estratégia de vincular de forma significativa conteúdos matemáticos a outros domínios e áreas do conhecimento.

Procurei agir, portanto, em consonância com os Parâmetros Curriculares Nacionais (PCN), segundo os quais:

> "A Matemática no Ensino Médio tem um valor formativo, que ajuda a estruturar o pensamento e o raciocínio dedutivo, porém também desempenha um papel instrumental, pois é uma ferramenta que serve para a vida cotidiana e para muitas tarefas específicas em quase todas as atividades humanas."(BRASIL, 2000, p. 48)

Minhas pesquisas conduziram-me ao tema ***Método dos Mínimos Quadrados : Estudo e Aplicações para o Ensino Médio***. O **Método dos Mínimos Quadrados** é um assunto que, exceto em pouquíssimos cursos técnicos da área tecnológica, não é estudado nem consta da grade curricular do Ensino Médio. Porém, ao debruçar-me sobre o tema, percebi a possibilidade de apresentá-lo ao aluno desse segmento, baseado no fato de que a implementação do ajuste lança mão de métodos que requerem o domínio de conteúdos matemáticos relevantes dentro do âmbito do Ensino Médio, como as funções, os polinômios e a resolução de sistemas de equações lineares, oportunizando a manipulação prática desses conteúdos e colocando o aluno em contato com uma ferramenta matemática de grande eficácia e utilidade.

Foi, portanto, essa característica, essa possibilidade de integrar num mesmo tema conteúdos do Ensino Médio a um recurso matemático de ampla aplicação, capaz de inserir-se em diferentes áreas de atividades, o principal elemento motivador na realização deste trabalho.

Estratégias desse tipo, acredito eu, convidam o aluno a repensar o saber matemático, enxergando o seu perfil catalisador, sua versatilidade e, inevitavelmente, sua importância como suporte a diversas atividades do imenso organismo social ao qual pertencemos. Isso resulta não apenas em consolidação do conhecimento, mas também em incentivo e entusiasmo na busca por novos saberes.

Percebi ainda que o tema possui grande afinidade com recursos computacionais, notadamente softwares de geração de curvas e programas cujos algoritmos automatizam os procedimentos envolvidos nos métodos de ajustes. Por isso, as atividades propostas, em muitas das ocasiões, sugerem a utilização dessas ferramentas, proporcionando a possibilidade de interação com recursos tecnológicos que agilizam e enriquecem o aprendizado.

Portanto, o objetivo geral deste trabalho é construir algo significativo para o aluno do Ensino Médio, apresentando-lhe uma ferramenta nova, o **Método dos Mínimos Quadrados**. Entende-se que a utilização deste recurso, no qual o aluno irá defrontar-se com conteúdos e competências já por ele adquiridos, poderá levá-lo à construções de novos saberes e a uma visão mais prática do universo matemático.

O assunto poderá ser tratado com a profundidade que se achar conveniente, pois a aplicação do método contempla atividades simples, como ajustes lineares a partir de conjuntos de tês pontos, e outras consideravelmente mais complexas. Estima-se que uma boa resposta possa ser obtida na apresentação do método a alunos de cursos técnicos da área tecnológica

A organização do trabalho obedece à seguinte estrutura:

O capítulo 1 contém uma breve exposição dos caminhos que conduziram ao desenvolvimento do ***Método dos Mínimos Quadrados***. Uma pequena exposição sobre o surgimento do conceito de funções á apresentada, visto que tal tópico está no cerne do tema estudado, culminando, de forma resumida, com os detalhes principais da criação do método. A abordagem busca deixar claro para quem lê este trabalho a importância e a eficácia do método como ferramenta de modelagem matemática em diversos ramos de atividades e conhecimentos, com seu uso ainda se fazendo muito presente nos dias atuais.

O capítulo 2 propõe-se a expor o método de uma forma tão ampla quanto necessária para que se atinjam as metas deste trabalho. Nele há a abordagem dos quesitos que motivam a utilização do método, situando-o como um importante desdobramento dentro do tema ***Ajuste de Curvas***. Conteúdos considerados de relação estreita com o método, ou por estarem vinculados à sua fundamentação teórica (como as derivadas parciais e o conceito de otimização), ou por constituírem um suporte à sua aplicação (como os sistemas de equações lineares), são abordados de forma sucinta ou apenas citados, de modo a conferirem o devido suporte ao tema principal. Em seguida, são apresentados os fundamentos matemáticos do método para, por fim, passar-se a uma abordagem individualizada de cada um dos três tipos de ajustes de curvas por ele contemplados : o ***Ajuste Linear Simples***, o ***Ajuste Linear Múltiplo*** e o ***Ajuste Polinomial***. Aqui busca-se apresentar o procedimento matemático que implementa o método para cada tipo de ajuste e sedimentá-lo com a exposição de exemplos simples, porém eficazes, capazes de conferir razoável desenvoltura na utilização do mesmo.

O capítulo 3 apresenta estratégias para apresentação do **Método dos Mínimos Quadrados** a alunos do Ensino Médio, bem como atividades que se constituam em exemplos considerados significativos da utilização do método. Buscou-se, na seleção e elaboração dessas atividades, demonstrar a abrangência do método, diversificando-se as áreas para sua aplicação.

O capítulo 4 contém as considerações finais sobre o trabalho realizado e aqui exposto.

Capítulo 1

Contextualização Histórica

1.1 O Surgimento das Funções

Embora o conceito de função, tal como conhecemos hoje, tenha começado a ser definido de forma mais consistente a partir do surgimento do Cálculo Infinitesimal, nos trabalhos de Isaac Newton e Gottfried Leibniz, com noções bastante próximas das atuais da ideia de variáveis e relacionamento entre entidades matemáticas, alguns historiadores enxergam na Antiguidade alguns antecedentes de tal conceito (PITOMBEIRA; ROQUE, 2012). Eles alegam que, uma vez que a ideia de correspondência está intrinsecamente vinculada à ideia de função, pode-se citar as tabelas babilônias e egípcias, além das tabelas utilizadas pela astronomia grega e em toda a matemática antiga, como representantes de uma incipiente noção de função.

Uma contribuição considerada significativa é a que foi legada pelo bispo francês Nicolas Oresme (1323-1382), que utilizou segmentos de reta para representar, segundo suas palavras, "tudo o que varia"(BOYER, 2001). Em uma de suas obras, ele representou a variação de velocidade de um móvel ao longo do tempo, dispondo os instantes ao longo de um segmento horizontal e associando a cada um deles um segmento perpendicular cujo comprimento representava a velocidade no instante. É possível identificar no modelo uma tímida antecipação do que viria a ser o modelo cartesiano criado por Descartes em 1637.

Apesar da noção de variável só ter sido estabelecida de forma definitiva no século XVIII, a noção de variação estava fortemente presente na busca da representatividade matemática do movimento e no desenvolvimento do estudo da variação dos fenômenos naturais em relação ao tempo, movimento que se propagou a partir do surgimento do método científico elaborado por Galileu. Esse foi o ponto de partida para que se passasse a associar o movimento a uma curva representada por uma equação.

A representação simbólica introduzida por Viète também deve ser citada, uma vez que permitia representar quantidades desconhecidas por fórmulas algébricas que as relacionavam a outras quantidades. Porém, era um recurso matemático utilizado para resolver problemas específicos, desvinculado do conceito fundamental de relacionamento entre grandezas que variam. O primeiro matemático a introduzir esse tipo de abordagem foi Descartes (PITOMBEIRA; ROQUE, 2012), trabalhando com equações indeterminadas, nas quais especificando-se infinitos valores para x, pode-se obter infinitos valores para y. Em seus estudos sobre curvas, já havia, portanto, uma relação funcional entre quantidades indeterminadas, relação essa expressa por uma equação. Embora não houvesse estudos mais objetivos sobre relações entre variáveis em tal época, pairava sobre os matemáticos de então a ideia de que tais relações consistiam de expressões analíticas de curvas algébricas ou de séries infinitas.

Assim, com a expansão do estudo de curvas por meios algébricos, tornou-se necessário criar um termo que designasse a expressão analítica utilizada para relacionar quantidades dependentes e suas respectivas variáveis. Consta que foi Leibniz, matemático que tratava com grande rigor a linguagem matemática, quem primeiro utilizou a palavra função com esse sentido (ROONEY, 2012). É com esse enfoque que ela aparece em uma correspondência por ele enviada, em 1698, a Johann Bernoulli (1667-1748), que anos mais tarde (1718) a publicou em um artigo apresentado à Academia de Ciências de Paris. Nesse artigo, ele introduziu uma notação para funções nas quais a característica era expressa pela letra grega ϕ e o argumento não continha parênteses, ou seja, ϕx.

Caberia a um discípulo de Bernoulli, Leonard Euler (1707-1783), finalizar a definição de função segundo a ideia de expressão analítica, motivado pela crença de que o cálculo nada mais era do que o estudo da teoria das funções. Em seu importante trabalho *Introdução à Análise dos Infinitos*, de 1748, ele expõe a seguinte definição de função: *"Uma função de uma quantidade variável é uma expressão analítica composta de um modo qualquer desta quantidade e de números, ou de quantidades constantes."* (PITOMBEIRA; ROQUE, 2012, p. 302)

Tal definição, na verdade, como se vê no restante de sua obra, ampliava consideravelmente o conceito de função, uma vez que a dita expressão analítica citada em sua definição poderia ser composta não apenas pelas tradicionais operações algébricas (adição, subtração, multiplicação, divisão, potenciação, radiciação e a resolução de equações algébricas), mas também por aquelas conhecidas como funções transcendentes, entre as quais se incluem os logaritmos e as exponenciais. Além disso, a variável poderia receber valores irracionais e imaginários. Foi Euler quem primeiro utilizou a notação $f(x)$, adotada ainda nos dias atuais.

A vinculação da definição de função à noção de expressão analítica predominou e norteou estudos ao longo dos séculos XVIII e XIX, até que no ano de 1837, o alemão Johann Dirichlet (1805-1859) ampliou o conceito formulando aquela que foi considerada a mais completa definição de função produzida até então, com reflexos que se fazem presentes ainda nos dias atuais, enunciada do seguinte modo:

> "Uma variável é um símbolo que representa um qualquer dos elementos de um conjunto de números; se duas variáveis x e y estão relacionadas de maneira que, sempre que se atribui um valor a x, corresponde automaticamente, por alguma lei ou regra, uma valor a y, então se diz que y é uma função (unívoca) de x."(EVES, 2004, p. 659).

A disseminação da Teoria dos Conjuntos em fins do século XIX, conduziu à definição formal do conceito de função da maneira que assumimos hoje:

> "Sejam x e y duas variáveis representativas de conjuntos de números; diz-se que y é uma função de x e escreve-se $y = f(x)$, se entre as duas variáveis existe uma correspondência unívoca no sentido $x \to y$. A x chama-se variável independente e a y, variável dependente."(CARAÇA, 1975, p. 129).

1.2 Origem do Método dos Mínimos Quadrados

O **Método dos Mínimos Quadrados** surgiu como consequência da aplicação de métodos e estudos estatísticos às ciências sociais com o mesmo rigor com o qual já eram aplicados às ciências sociais, em meados do século 19. O método científico, criado por Galileu Galilei (1564-1642) e utilizado até hoje, baseado em observações e medições, também teve papel fundamental nesse evento (PITOMBEIRA; ROQUE, 2012).

Antes disso, o astrônomo dinamarquês Tycho Brahe (1546-1601) já utilizava métodos estatísticos primitivos para estimar as órbitas de determinados corpos celestes (BOYER, 2001). Sabe-se que ele eliminava, entre os valores coletados, aqueles que lhe pareciam discrepantes, obtinha médias entre os dados considerados válidos e as utilizava em suas estimativas, agregando a estas estimativas uma qualificação do seu rigor.

Levaria um bom tempo até que o estudo de conjunto de observações experimentais começasse a ter um tratamento matemático mais cuidadoso e sistemático. São nomes de grande relevância dessa fase, Roger Joseph Boscovich (1711-1787), Pierre de Laplace (1749-1827), Adrien Marie Legendre (1752-1833) e Carl Friedrich Gauss (1777-1855). Os dois últimos considerados os pais do **Método dos Mínimos Quadrados** (ROONEY, 2012).

Era uma época de intensas descobertas e debates sobre fenômenos astronômicos. Assim, era uma necessidade premente o cálculo dos parâmetros das órbitas dos cometas a partir de mediadas pontuais, conseguidas em momentos distintos.

Dentre as diversas soluções apresentadas para tal problema na transição do século XXIII para o século XIX, a que se mostrou mais eficaz e de maior alcance prático, além de maior respaldo teórico foi o **Método dos Mínimos Quadrados**, publicado pela primeira em 1805, na obra *Nouvelles Methodes pour la Determination des Orbites des Comètes*, de Legendre. Porém, credita-se a Gauss (figura 1.1) o desenvolvimento do método, por volta de 1795, embora ele só o tenha publicado em 1809. As conclusões contidas no trabalho de Gauss apontavam que o método fornecem a melhor estimativa quando se parte do princípio de que os erros cometidos nas medidas seguem a distribuição da curva normal (ROONEY, 2012).

Figura 1.1 – Carl Friedrich Gauss, por Christian Albrecht Jensen
Fonte: Rooney (2012, p. 126)

Segundo Crato (2000), o fato determinante da formatação definitiva do **Método dos Mínimos Quadrados** foi um problema levado até Gauss em 1808 por Giuseppi Piazzi (1746-1826), diretor do Observatório de Palermo na época. Depois de descobrir o asteroide Ceres e supor tratar-se de um planeta ainda desconhecido de nosso sistema solar, Palermo perdeu o astro de vista. Os métodos matemáticos para cálculos de órbitas dos corpos celestes na época requeriam um grande número de observações tomadas em um período de tempo razoavelmente longo. Dispondo de um número insuficiente de dados e ávido por preservar a prioridade da descoberta, Palermo recorreu aos préstimos do jovem Gauss, então com apenas vinte e quatro anos de idade.

Ainda segundo Crato (2000), o jovem gênio, que havia anos já se dedicava a estudar o problema do cálculo de órbitas de corpos celestes sem quaisquer pressupostos teóricos, tendo por base observações realizadas num período de tempo não necessariamente longo, conseguiu determinar a órbita de Ceres e, através de seus estudos, estabeleceu um método para manipular observações (dados) e daí estimar os parâmetros de uma função – o **Método dos Mínimos Quadrados**. Foi utilizando as estimativas baseadas nos cálculos de Gauss para a órbita de Ceres que o astrônomo húngaro Franz Xaver Von Zach(1754-182), então diretor do Observatório de Seeburgo, o redescobriu praticamente um ano após o primeiro avistamento. O asteroide estava a uma distância angular de meio grau da posição estimada por Gauss.

O **Método dos Mínimos Quadrados** continua a ter grande importância no campo da astronomia, notadamente na descoberta de planetas que orbitam outras estrelas que não o Sol. Além disso, o método foi aplicado a diversos outros ramos de atividades, tornando-se a principal ferramenta dos estatísticos no século XIX.

Capítulo 2

O Método dos Mínimos Quadrados

2.1 Considerações Preliminares

2.1.1 Motivação para o Surgimento e Utilização do Método

O **Método dos Mínimos Quadrados** tem sua origem no estudo dos valores máximos e mínimos de funções reais. Mais precisamente, na determinação do(s) ponto(s) mínimo(s) de uma função que representa o desvio estimado na busca pelo ajuste. Antes de abordarmos este e outros tópicos envolvidos na dedução e na aplicação do método, é conveniente nos debruçarmos sobre o problema que motivou a sua elaboração e seus desdobramentos.

No estudo de qualquer fenômeno, natural ou proveniente de qualquer área de atividade humana, a sua representação por equações e curvas produzidas pelo relacionamento entre as grandezas que o regem constitui, certamente, uma das ferramentas mais eficazes desse estudo. Impossível imaginar o entendimento e o alcance da Teoria de Relatividade de Einstein, das Leis da Mecânica Clássica elaboradas por Newton, dos juros simples e compostos na Matemática Financeira, sem o suporte das fórmulas matemáticas que expressam todos os princípios aí envolvidos.

Os temas acima citados são bastante conhecidos e têm suas bases matemáticas solidamente assentadas, porém há muitos outros fenômenos na vida cotidiana cuja análise será muito mais significativa e ampla se conseguirmos descrevê-los por meio de modelos matemáticos e com a inclusão de um termo que relaciona-se, em geral, com erros cometidos ou nas hipóteses do modelo ou na coleta de observações. Será por meio desses modelos que poderemos, por exemplo, estimar ocorrências futuras e direcionar tomadas de decisões.

Podemos classificar os inúmeros fenômenos que ocorrem à nossa volta em dois grupos:

- **Fenômenos determinísticos**: São aqueles em que, uma vez conhecidas suas causas, podemos identificar e descrever sua situação final.

- **Fenômenos aleatórios**: São aqueles não são passíveis de uma abordagem que nos permita descrever e identificar sua situação final, devido à própria natureza do fenômeno ou por conta das limitações de nossas observações.

Assim, o crescimento populacional de uma determinada região em função do tempo, a produtividade de uma fazenda em função da quantidade de adubo nela utilizada, a variação de volume de uma determinada substância em função da temperatura a partir de dados obtidos em laboratório, são alguns dos incontáveis exemplos de fenômenos que carecem de modelos determinísticos que lhes deem representatividade e suporte analítico.

O problema, portanto, consiste em, a partir de um conjunto de dados, obtidos por experimentação ou observação, relacionando grandezas (variáveis) envolvidas na ocorrência de determinado fenômeno, encontrar um modelo matemático que descreva e expresse de maneira satisfatória tal fenômeno.

As variáveis envolvidas em um determinado fenômeno podem estar relacionadas de acordo com uma das seguintes classificações:

- **Determinísticas**: Neste tipo de relação há uma fórmula matemática precisa vinculando as variáveis envolvidas, ou seja, é possível determinar-se com exatidão o valor de qualquer uma das variáveis quando se tem conhecimento dos valores das demais. Exemplo: a lei dos juros compostos: $M = C(1+i)^t$, onde M é o montante, C é o capital aplicado, i é a taxa de juros e t é o tempo de aplicação.

- **Semideterminísticas**: Neste tipo de relação há uma expressão matemática vinculando as variáveis envolvidas, porém desconhece-se algum ou alguns do(s) valor(es) dos parâmetros existentes na relação. Exemplo: a dilatação volumétrica de um corpo com a variação de temperatura pode ser expressa pela seguinte relação: $\Delta V = V_0.\gamma.\Delta\theta$. Neste caso, a variação de volume Δv e a variação de temperatura $\Delta\theta$ são as variáveis de modelo, enquanto o coeficiente de dilatação volumétrica do corpo, γ, é o seu parâmetro. O coeficiente de dilatação volumétrica é uma constante específica de cada material e pode ser obtida a partir de dados obtidos experimentalmente ou tentando-se ajustar esses dados à relação conhecida.

- **Empíricas**: Neste caso, a relação entre as variáveis envolvidas é desconhecida. Exemplo: A produtividade do solo de uma fazenda em função da quantidade de adubo nele introduzido. Dados obtidos experimentalmente, nos quais se busca ao máximo eliminar-se a influência de outros fatores, podem, depois do devido tratamento, levar à obtenção de uma relação satisfatória entre as duas grandezas.

Não é difícil deduzir que as análises apresentadas neste trabalho contemplarão os tipos de relacionamentos de variáveis descritos no segundo e, principalmente, no terceiro item.

Um dos métodos de obtenção de uma função representativa de um determinado fenômeno é a **interpolação**, que consiste em se obter uma curva suave que contenha um número reduzido de pontos, selecionados dentre um grupo obtido experimentalmente (SPERANDIO; MENDES; SILVA, 2003). A curva pode ser obtida por processo geométrico ou algébrico (**Interpolação de Lagrange**), devendo passar por todos os pontos selecionados. A seleção baseia-se no intervalo da função que é de nosso interesse e apenas nesse intervalo a função obtida dá representatividade ao fenômeno (SANCHES; JORDAN, 2007). Este método não é objeto de estudo deste trabalho.

O outro método, o **ajuste de curvas**, constitui uma alternativa mais adequada que a **interpolação** na obtenção de uma função matemática para representar um determinado fenômeno quando se deseja efetuar análises que vão além do intervalo de valores obtidos experimentalmente. O **Método dos Mínimos Quadrados**, que aqui será estudado, surgiu para dar suporte a esse recurso matemático.

2.1.2 O Ajuste de Curvas

Definição 2.1 *É o procedimento matemático que consiste em se determinar, a partir de uma série de pontos representativos das variáveis que compõem um determinado fenômeno, uma curva que o expresse matematicamente. A curva obtida deve permitir com satisfatória segurança a realização de análises e projeções sobre o fenômeno em questão.*

Uma boa maneira de demonstrar a utilidade e o alcance do Ajuste de Curvas é, mesmo sem ainda ter exposto os procedimentos que levam à sua execução, exibir um exemplo contendo o problema proposto e a solução obtida, sem exibir os passos intermediários que levaram a essa solução, para melhor entendimento e avaliação do método.

Exemplo 2.1 *Suponhamos que uma série de experimentos relacionando duas grandezas **A** e **B** foi realizada em laboratório, dando origem à tabela 2.1:*

A	$1,3$	$3,4$	$5,1$	$6,8$	$8,0$
B	$2,0$	$5,2$	$3,8$	$6,1$	$5,8$

Tabela 2.1 – Dados experimentais relacionando as grandezas **A** e **B**

Ao dispormos os pontos obtidos em um plano cartesiano obtemos o **diagrama de dispersão** do conjunto, exposto na figura 2.2:

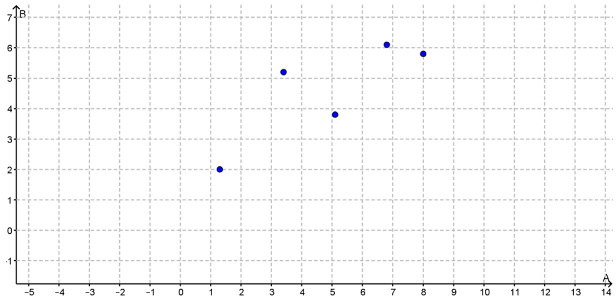

Figura 2.2 – Diagrama de dispersão para as grandezas **A** e **B**
Fonte: Elaboração própria

O objetivo consiste em encontrar uma função que represente, com boa aproximação, os valores tabelados a partir do experimento realizado e, além disso, nos permita determinar, com certa margem de segurança, valores que estão fora do intervalo inicialmente definido, ou seja, nos permita **extrapolar**.

O gráfico 2.3 apresenta tal curva, obtida após a aplicação do **Método dos Mínimos Quadrados** sobre os dados tabelados. A curva mostrada na figura 2.3 foi traçada após a

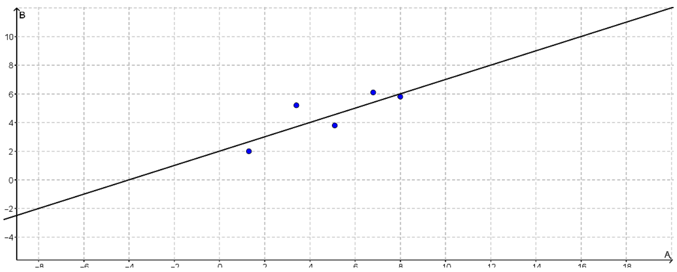

Figura 2.3 – Ajuste sobre dados da tabela 2.1
Fonte: Elaboração própria

determinação, pelo método, de uma função envolvendo as grandezas **A** e **B**. Esta função, acredita-se, representa com boa aproximação o fenômeno estudado e, assim sendo, de posse dela podemos estimar hipotéticas ocorrências do mesmo.

Neste exemplo buscou-se um polinômio que de antemão sabia-se ser de grau um, ou seja, efetuou-se um **Ajuste Linear Simples** ou uma **Regressão Linear** sobre o conjunto de dados, resultando na curva exibida no gráfico. Essa é uma das características do método, a necessidade de termos o tipo de curva a ser obtida previamente definido.

Uma pergunta que logo surge é sobre os critérios para escolha do tipo de função de ajuste. Obviamente, tal escolha não é aleatória. Segundo Guimarães (2001, p. 72):

> "A escolha da função f(x) a ser usada para ajustar o conjunto de pontos experimentais não é feita ao acaso. Ela envolve um processo de análise que passa, por exemplo, pela identificação das relações de causa e efeito que governam o sistema estudado. O fato dessas relações não serem conhecidas não é empecilho para que se intua a forma de f(x). Uma análise qualitativa do fenômeno e das condições experimentais em que a observação é feita muitas vezes é suficiente para que se possa estabelecer a forma dessa função."

Temos também a possibilidade de avaliar, por procedimento matemático que será exposto e utilizado neste trabalho, a qualidade do ajuste, o que poderá nos auxiliar a fazer a melhor escolha.

No caso do ajuste a um polinômio, um critério que se mostra bastante eficaz e será utilizado em uma das atividades propostas neste trabalho consiste em se aumentar sistematicamente o grau do polinômio procurado de modo a, baseados em uma avaliação matemática dos ajustes realizados, selecionarmos aquele que proporciona o melhor resultado. Isto significa implementar diversas vezes o método sobre os dados em mãos, o que não se mostra tão trabalhoso com os recursos computacionais e ferramentas de cálculo existentes nos dias atuais.

Outro bom exemplo de utilização do método ocorre quando, apesar de conhecermos a função não apenas em alguns pontos, mas em todo um intervalo, temos o interesse de aproximar essa função por funções de outra classe, mais adequada às nossas necessidades. Temos em Guimarães (2001) um exemplo no qual o objetivo é determinar qual a melhor reta que se ajusta a uma aproximação da função $\operatorname{sen}(x)$ no intervalo $[0, 1]$. Vejamos com mais detalhes.

A tabela 2.2 apresenta alguns valores da função $\operatorname{sen}(x)$ no intervalo $[0, \frac{\pi}{2}]$:

x	0	$\frac{\pi}{16}$	$\frac{\pi}{8}$	$\frac{\pi}{4}$	$\frac{\pi}{2}$
$\operatorname{sen}(x)$	$0,0000$	$0,1950$	$0,3827$	$0,7071$	$1,0000$

Tabela 2.2 – Valores de $\operatorname{sen}(x)$ no intervalo $[0, \frac{\pi}{2}]$

Aplicando o **Método dos Mínimos Quadrados** sobre os dados da tabela acima de modo a se obter uma função do segundo grau, obtém-se a curva representada no gráfico 2.4, em confronto com a curva original. Uma análise visual rápida nos permite atestar a boa aproximação fornecida pelo ajuste.

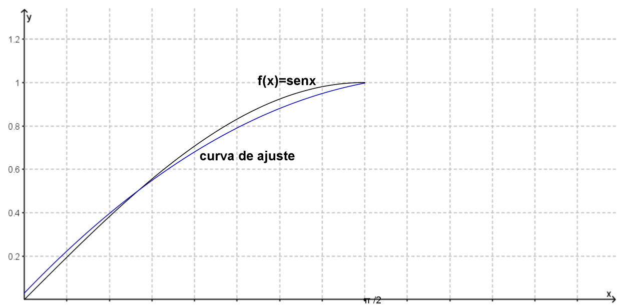

Figura 2.4 – Gráfico da função de juste em confronto com o gráfico de $\operatorname{sen}(x)$
Fonte: Elaboração própria

2.2 Definições Básicas

2.2.1 Máximos e Mínimos para Funções de uma Variável

Definição 2.2 *Seja $y = f(x)$ uma função definida num conjunto aberto $U \subset \mathbb{R}$. Dizemos f tem um **máximo local ou relativo** em $c \in U$ se existir um intervalo aberto I contendo c, de modo que $f(c) \geq f(x)$ para todo $x \in I$.*

Definição 2.3 *Seja $y = f(x)$ uma função definida num conjunto aberto $U \subset \mathbb{R}$. Dizemos f tem um **mínimo local ou relativo** em $c \in U$ se existir um intervalo aberto I contendo c, de modo que $f(c) \leq f(x)$ para todo $x \in I$.*

Como exemplo, a figura 2.5 exibe os pontos máximo local (ponto O) e mínimos locais (pontos P e Q) da função $f(x) = 3x^4 - 12x^2$.

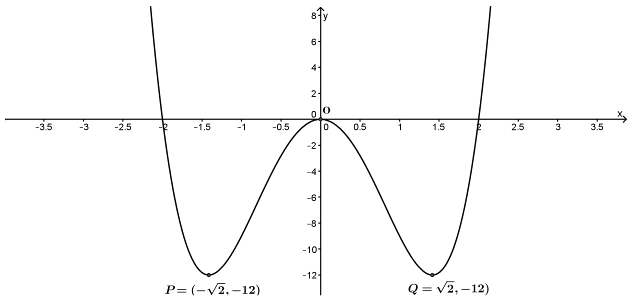

Figura 2.5 – Máximo e mínimos locais da função $f(x) = 3x^4 - 12x^2$
Fonte: Elaboração própria

Definição 2.4 *Seja $y = f(x)$ uma função definida num conjunto aberto $U \subset \mathbb{R}$. Dizemos que f tem um **máximo absoluto** em $x_0 \in U$ se, para todo $x \in U$, tivermos $f(x_0) \geq f(x)$.*

Definição 2.5 *Seja $y = f(x)$ uma função definida num conjunto aberto $U \subset \mathbb{R}$. Dizemos que f tem um **mínimo absoluto** em $x_0 \in U$ se, para todo $x \in U$, tivermos $f(x_0) \leq f(x)$.*

A figura 2.6 exemplifica, reúne e resume as definições de máximo e mínimo para funções de uma variável real vistas até aqui.

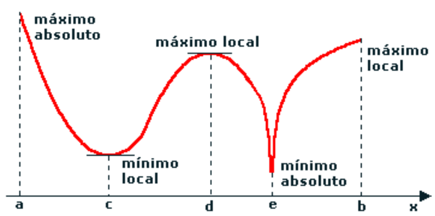

Figura 2.6 – Máximos e mínimos de uma hipotética função de uma variável
Fonte: UEL (s. d.)

Teorema 2.1 *Seja $f : U \to \mathbb{R}$ uma função derivável em $x_0 \in U$. Uma condição necessária para que x_0 seja ponto de máximo ou de mínimo local é que se tenha $f'(x_0) = 0$.*

Definição 2.6 *Sejam $f : U \to \mathbb{R}$ e $x_0 \in U$. Diz-se que x_0 é* ***ponto crítico*** *de f se $f'(p) = 0$.*

Ou seja, o teorema 2.1 nos diz que se $x_0 \in U$ e f é derivável em x_0, então uma condição necessária para que x_0 seja ponto de máximo ou de mínimo local de f é que x_0 seja ponto crítico de f. Vejamos agora a condição suficiente para que x_0 seja ponto de máximo ou de mínimo local de f.

Teorema 2.2 *Sejam $f : U \to \mathbb{R}$ uma função que admite derivada de 2ª ordem contínua no intervalo aberto I e $x_0 \in I$:*

1. *$f'(x_0) = 0$ e $f''(x_0) > 0 \Rightarrow p$ é ponto de mínimo local.*
2. *$f'(x_0) = 0$ e $f''(x_0) < 0 \Rightarrow p$ é ponto de máximo local.*

2.2.2 Derivadas Parciais

2.2.2.1 Derivadas Parciais em $\mathbb{R}^2$

Num estudo aprofundado sobre derivadas de funções reais, recomenda-se que a abordagem se inicie por funções de uma única variável. Porém, tendo o estudo da derivada, neste trabalho, um caráter auxiliar e introdutório, iremos nos ater aos conceitos e propriedades principais contidos no estudo das derivadas de funções de várias variáveis reais, situando as funções de uma única variável como um caso particular desse contexto.

O estudo de derivadas parciais de funções reais de várias variáveis reais vai muito além do que o exposto neste trabalho. Aqui será priorizada a abordagem sucinta dos tópicos nos quais se baseiam as premissas matemáticas que deram origem ao **Método dos Mínimos Quadrados**. Neste breve estudo sobre derivadas parciais de uma função real de várias variáveis reais, $f : Dom(f) \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$, vamos inicialmente definir as derivadas parciais para funções reais de duas variáveis e, em seguida, expandir o conceito a domínios além do $\mathbb{R}^2$, buscando a generalização do conceito.

Sejam

$$f : Dom(f) \subseteq \mathbb{R}^2 \longrightarrow \mathbb{R}$$

$$(x, y) \to z = f(x, y).$$

Fixemos $y = y_0$ e, a partir disso, criemos uma função de tal reta em $\mathbb{R}$, g_{y_0}, definida como $g_{y_0} = f(x, y_0)$. A função assim obtida é uma função real de variável real, sobre a qual podemos aplicar o conceito de derivabilidade. Podemos então dizer que, se g_{y_0} é derivável em x_0, existe $g'_{y_0}(x_0)$, definida da seguinte forma:

$$g'_{y_0}(x_0) = \lim_{h \to 0} \frac{g_{y0}(x_0 + h) - g_{y0}(x_0)}{h}$$

Tal limite define a *derivada parcial de f com ralação a x no ponto* (x_0, y_0), a qual representamos por $\frac{\partial f}{\partial x}(x_0, y_0)$. Ou seja:

$$\frac{\partial f}{\partial x}(x_0, y_0) = g'_{y0}(x_0) = \lim_{h \to 0} \frac{g_{y_0}(x_0 + h) - g_{y_0}(x_0)}{h}$$

Do mesmo modo, fixamos a variável x, fazendo $x = x_0$ e criamos a função de tal reta em $\mathbb{R}$, h_{x_0}, definida como $h_{x0} = f(x_0, y)$. Esta também é uma função real de variável real, sujeita ao conceito de derivabilidade. Assim, se h_{x_0} é derivável em y_0, existe $h'_{x_0}(y_0)$, a qual definimos da seguinte forma:

$$h'_{x_0}(y_0) = \lim_{t \to 0} \frac{h_{x0}(y_0 + t) - h_{x0}(y_0)}{t}$$

Este limite define a *derivada parcial de f com ralação a y no ponto* (x_0, y_0), representada por $\frac{\partial f}{\partial y}(x_0, y_0)$. Ou seja:

$$\frac{\partial f}{\partial y}(x_0, y_0) = h'_{x0}(y_0) = \lim_{t \to 0} \frac{h_{x_0}(y_0 + h) - h_{x_0}(y_0)}{t}$$

2.2.2.2 Derivadas Parciais em $\mathbb{R}^n$

Sejam

$$f : Dom(f) \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$$

$$X = (x_1, x_2, ..., x_n) \to f(X) = f(x_1, x_2, ..., x_n)$$

e seja $X_0 = (x_{1\,0}, x_{2\,0}, ..., x_{n\,0}) \in Dom(f)$. Criemos uma função g_{iX0}, definida da seguinte forma:

$$g_{iX0} =: \; Dom(g_{iX_0}) \subseteq \mathbb{R} \longrightarrow \mathbb{R}$$

$$x_i \to g_i x_0(x_i) = f(x_{1\,0}, x_{2\,0}, ..., x_{(i-1)\,0}, x_i, x_{(i+1)\,0} ..., x_{n\,0})$$

Ou seja, fixamos todas as variáveis, exceto x_i. Logo, g_{iX_0} é uma função da reta em $\mathbb{R}$. Se g_{iX_0} é derivável em $x_{i\,0}$, podemos definir a *derivada parcial de f em relação à variável x_i no ponto* $X_0 = (x_{1\,0}, x_{2\,0}, ..., x_{n\,0})$ como $g'_{iX_0}(x_0)$, cuja notação é $\frac{\partial f}{\partial x_1}(X_0)$.

Assim sendo, podemos escrever:

$$\frac{\partial f}{\partial x_1}(X_0) = g'_{iX_0}(x_0) = \lim_{h \to 0} \frac{g_{iX_0}(x_{i0} + h) - g_{iX_0}(x_{i0})}{h}$$

$$= \lim_{h \to 0} \frac{f(x_{1\,0}, ..., x_{i\,0} + h, ..., x_{n\,0}) - f(x_{1\,0}, ..., x_{i\,0}, ..., x_{n\,0})}{h}$$

Todas as regras de derivação de funções de uma variável se aplicam ao cálculo de derivadas de funções de n variáveis. Essas regras atuam apenas sobre a variável tomada como referência, sendo as demais tratadas como constantes.

2.2.2.3 Interpretação Geométrica das Derivadas Parciais

Analisaremos, somente o caso de funções reais de duas variáveis reais, pois é o que basta para um bom entendimento do assunto.

O gráfico gerado pela função de duas variáveis, $z = f(x, y)$, é, em geral, uma superfície no espaço $\mathbb{R}^3$. A interseção dessa superfície com um plano paralelo ao plano xz, passando por um ponto (x_0, y_0, z) é uma curva plana (ou um ponto, no caso de haver tangência), definida da seguinte forma:

$$C_1 : \begin{cases} z = & f(x, y) \\ y = & y_0 \end{cases}$$

Podemos tratar a curva assim definida como uma função de uma variável. Chamando essa função de g, escrevemos: $g(x) = f(x, y_0)$. Portanto, o coeficiente angular α da curva no ponto $P = (x_0, y_0, z)$, em relação ao plano x_0z, é dado por:

$$g'(x_0) = \frac{\partial f}{\partial x}(x_0, y_0)$$

A figura 2.7 representa geometricamente a *derivada parcial da função f em relação à variável x em um ponto P*.

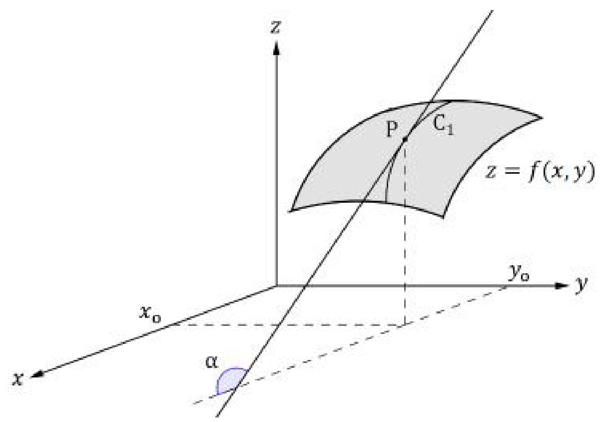

Figura 2.7 – Derivada parcial de f em relação a x
Fonte: Pereira (2007, p. 23)

Procedendo de modo análogo e tomando a curva definida pela interseção do gráfico de f com o plano que passa por um ponto (x_0, y, z), paralelo ao plano yz, definida da seguinte forma:

$$C_2 : \begin{cases} z = & f(x, y) \\ x = & x_0 \end{cases}$$

obtemos, como se nota, uma função de uma variável. Chamando essa função de h, escrevemos $h(x) = f(x_0, y)$. Assim sendo, o coeficiente angular β da curva no ponto $P = (x_0, y_0, z)$, em relação ao plano y_0z, é dado por:

$$g'(y_0) = \frac{\partial f}{\partial y}(x_0, y_0)$$

A figura 2.8 representa geometricamente a *derivada parcial da função f em relação à variável y em um ponto P.*

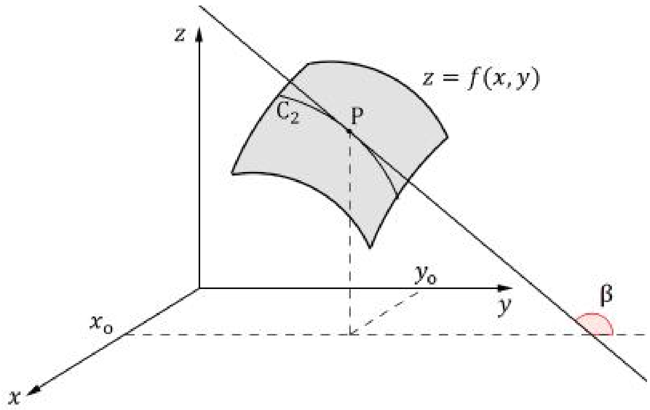

Figura 2.8 – Derivada parcial de f em relação a y
Fonte: Pereira (2007, p. 24)

2.2.2.4 Derivadas Parciais de Segunda Ordem

Definição 2.7 *Ao derivarmos uma função $z = f(x, y)$ duas vezes, obtemos suas derivadas de segunda ordem, assim representadas:*

$$\begin{aligned}
(f_x)_x &= f_{xx} = f_{11} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial x^2} = \frac{\partial^2 z}{\partial x^2} \\
(f_x)_y &= f_{xy} = f_{12} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial y \partial x} = \frac{\partial^2 z}{\partial y \partial x} \\
(f_y)_x &= f_{yx} = f_{21} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 z}{\partial x \partial y} \\
(f_y)_y &= f_{yy} = f_{11} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial^2 f}{\partial y^2} = \frac{\partial^2 z}{\partial y^2}
\end{aligned}$$

A notação f_{xx} nos diz que derivamos a função f duas vezes em relação a x. A notação f_{xy} nos diz que primeiro derivamos em relação a x e, em seguida, derivamos a função resultante em relação a y. A notação f_{yx} nos diz que derivamos primeiro em relação a y e, em seguida, em relação a x. Por fim, a notação f_{yy} significa duas derivações sucessivas em relação a y.

2.2.2.5 Derivadas Parciais de Ordem Superior

Teoricamente, não há limite para o número de vezes que podemos derivar uma função, desde que a haja garantia de existência de suas derivadas parciais. Exemplificando, podemos escrever:

$$\frac{\partial^3}{\partial x^2 \partial y} = f_{yxx}, \quad \frac{\partial^4 f}{\partial x^2 \partial y^2} = f_{yyxx}, \quad \dots$$

2.2.3 Máximos e Mínimos em Funções de Duas Variáveis

Definição 2.8 *Seja $z = f(x, y)$ uma função de duas variáveis, definida num conjunto aberto $U \subset \mathbb{R}^2$. Dizemos que um ponto (x_0, y_0) pertencente ao domínio de f é ponto de **mínimo relativo ou local** de f se existe uma bola aberta B contida em U, de raio r e centro em (x_0, y_0), tal que, para todo ponto $P(x, y)$ pertencente a esse domínio e situado no interior da bola, tenhamos $f(x, y) \geq f(x_0, y_0)$. Se esta inequação vale para todos os pontos (x, y) do domínio de f, então f tem um **mínimo absoluto** em f.*

A figura 2.9 mostra uma função de duas variáveis com mínimo absoluto em $P(x_0, y_0, z_0)$.

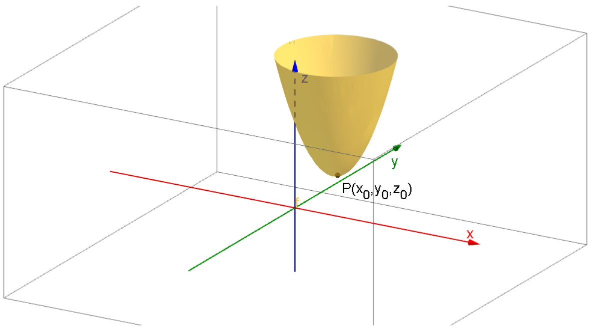

Figura 2.9 – Ponto mínimo absoluto de uma função de duas variáveis reais
Fonte: Elaboração própria

Definição 2.9 *Seja $z = f(x, y)$ uma função de duas variáveis, definida num conjunto aberto $U \subset \mathbb{R}^2$. Dizemos que um ponto (x_0, y_0) pertencente ao domínio de f é ponto de **máximo relativo ou local** de f se existe uma bola aberta B contida em U, de raio r e centro em (x_0, y_0), tal que, para todo ponto $P(x, y)$ pertencente a esse domínio e situado no interior da bola, tenhamos $f(x, y) \leq f(x_0, y_0)$. Se esta inequação vale para todos os pontos (x, y) do domínio de f, então f tem um **máximo absoluto** em f.*

A figura 2.10 mostra uma função de duas variáveis com máximo absoluto em $P(x_0, y_0, z_0)$.

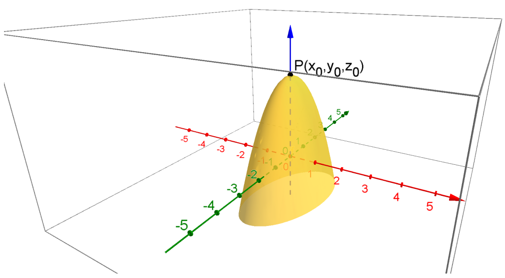

Figura 2.10 – Ponto máximo absoluto de uma função de duas variáveis reais
Fonte: Elaboração própria

Teorema 2.3 *Seja a função $f : U \subset \mathbb{R}^2 \to \mathbb{R}$, onde $z = f(x, y)$ e seja o ponto $(x_0, y_0) \in U$. Se f é diferenciável no intervalo aberto U e tem um máximo local ou um mínimo local em (x_0, y_0), então $\frac{\partial f}{\partial x}(x_0, y_0) = \frac{\partial f}{\partial y}(x_0, y_0) = 0$.*

Definição 2.10 *Seja uma função f, definida num conjunto aberto $U \subset \mathbb{R}^2$, onde $z = f(x, y)$. Um ponto $(x_0, y_0) \in U$ é dito um ponto crítico de f se as derivadas parciais $\frac{\partial f}{\partial x}(x_0, y_0)$ e $\frac{\partial f}{\partial y}(x_0, y_0)$ são nulas ou se f não é diferenciável em (x_0, y_0).*

Pela definição 2.10, concluímos que pontos máximos e mínimos são pontos críticos. Porém, nem todo ponto crítico é um máximo ou um mínimo. A função pode ter, em um ponto crítico, um máximo local ou um mínimo local ou, ainda, nenhum deles. Um ponto crítico que não é nem máximo local nem mínimo local é chamado de **ponto de sela**.

Geometricamente, um ponto é ponto crítico de uma função quando o plano tangente ao gráfico da função nesse ponto é um plano horizontal ou não existe.

2.2.4 Máximos e Mínimos em $\mathbb{R}^n$

Embora a interpretação geométrica para máximos e mínimos em $\mathbb{R}^n$, quando $n > 2$, não exista, o conceito pode ser estendido e aplicado. Será a partir dele que iremos chegar ao sistema de equações lineares que nos permitirá aplicar um caso particular do **Método dos Mínimos Quadrados**, o **Ajuste Linear Múltiplo**. Mais precisamente, será através da generalização do teorema 2.3 que obteremos o suporte matemático do método nesse caso.

Teorema 2.4 *Seja a função $f : U \subset \mathbb{R}^n \to \mathbb{R}$, onde $z = f(x_1, x_2, ..., x_n)$ e seja o ponto $(x_{1\,0}, x_{2\,0}, ..., x_{n\,0}) \in U$. Se f é diferenciável no intervalo aberto U e tem um máximo local ou um mínimo local em $(x_{1\,0}, x_{2\,0}, ..., x_{n\,0})$, então*

$$\frac{\partial f}{\partial x_1}(x_{1\,0}, x_{2\,0}, ..., x_{n\,0}) = \frac{\partial f}{\partial x_2}(x_{1\,0}, x_{2\,0}, ..., x_{n\,0}) = ... = \frac{\partial f}{\partial x_n}(x_{1\,0}, x_{2\,0}, ..., x_{n\,0}) = 0$$

.

2.3 Fundamentação Matemática do Método dos Mínimos Quadrados

Como já visto, o **Ajuste de Curvas** consiste na busca por uma função matemática que represente com boa aproximação uma relação que se aplique a um conjunto de dados. Com a aplicação do método, tenciona-se obter uma função cuja curva, embora não passe necessariamente pelos pontos que representam os dados em mãos, seja resultado de uma processo de otimização cujo objetivo é minimizar a distância entre essa curva e o conjunto de pontos que representam o fenômeno em questão, isto é, o erro entre a função que adotamos como a que define a curva e verdadeira equação da curva, que é desconhecida.

Suponhamos que tenhamos um conjunto de n pontos que associem duas grandezas x e y envolvidas em um determinado fenômeno, por exemplo um experimento realizado em laboratório. Seja f a função que se deseja definir para dar representatividade ao fenômeno, de modo que $y = f(x)$, e seja $(x_i, y_i),\ i \leq 1 \leq n$, o par ordenado que representa o i-ésimo ponto do conjunto experimental. Para cada um desses pares, podemos definir como **resíduo** a quantidade:

$$\Delta y_i = y_i - f(x_i),$$

que nos fornece a diferença entre a ordenada do i-ésimo ponto experimental e a ordenada obtida pela aplicação de f sobre a abscissa $x = x_i$. Podemos considerar que essa quantidade representa o erro h na aproximação de y_i através de $f(x_i)$.

Consideremos algumas possibilidades que se apresentam na busca pelo melhor ajuste:

- **Fazer com que os resíduos tendam a zero**: Tal procedimento corresponderia à **interpolação**, cujo estudo, como já visto, não é propósito deste trabalho.

- **Anular a soma dos resíduos**: Matematicamente, isto significa que devemos proceder de modo que o **desvio simples total**, dado por $\xi = \sum_{i=1}^{n}(y_i - f(x_i))$ venha a ser igual a zero. A experiência nos mostra que tal procedimento não é conveniente, uma vez que pode conduzir a resultados incorretos. Por exemplo, consideremos os resíduos para o i-ésimo e o j-ésimo pontos, Δ_i e Δ_j, respectivamente, e suponhamos que temos $\Delta_i = -\Delta_j$. Isto levaria a uma redução no somatório dos erros que não corresponde à realidade. Tal argumento leva a cogitar como melhores opções o quadrado e o módulo dos resíduos. O módulo, porém, tem a desvantagem de não ser derivável na origem.

- **Minimizar a soma das distâncias dos pontos experimentais à curva de ajuste**: Sendo distâncias quantidades positivas, o pensamento inicial nos leva a crer numa facilitação na busca da curva de ajuste. Se levarmos em conta, porém, que em Geometria Analítica a definição de distância recai em uma raiz quadrada de uma soma de quadrados, fica evidente que a bagagem matemática necessária para aplicar tal procedimento é bem mais complexa e trabalhosa do que se supunha inicialmente. Daí, mais uma vez, pensar-se na soma dos quadrados dos resíduos, artifício matemático que levará à eliminação das raízes quadradas existentes.

- **Minimizar o quadrado da soma dos desvios**: É a base do **Método dos Mínimos Quadrados.** Trabalhando com a soma dos quadrados dos resíduos, eliminamos incoerências no somatório desses resíduos e obtemos, para o erro total, uma função mais fácil de manipular na busca pela minimização. O **Método dos Mínimo Quadrados** é, assim, o procedimento mais eficiente e versátil na busca da função de ajuste.

O método parte do **erro quadrático** Δy_i^2, que ocorre ao substituirmos o valor da ordenada yi pelo valor aproximado correspondente $f(x_i)$, definido como o quadrado do **desvio simples**, para se obter uma expressão que otimize o ajuste. Podemos escrever que ao i-ésimo ponto, associa-se o desvio quadrático

$$\Delta y_i^2 = (y_i - f(x_i))^2$$

Além disso, definimos o erro total S associado ao ajuste como a soma dos desvios quadráticos individuais referentes a cada ponto (x_i, y_i), expressando-o da seguinte forma:

$$S = \sum_{i=1}^{n} (y_i - f(x_i))^2 \tag{2.1}$$

O valor de S representa, portanto, a soma dos quadrados das distâncias, na vertical, de cada ponto (x_i, y_i) à curva, ou seja, o **desvio quadrático total** provocado pelo ajuste. Assim sendo, a **incerteza média** σ associada a tal ajuste nos é dada pela equação

$$\sigma = \sqrt{\frac{S}{n}}$$

.

A otimização do ajuste consiste em minimizar-se o erro quadrático médio σ, o que, obviamente, equivale a minimizar-se a função S, dada na equação 2.1.

Vale ressaltar que neste trabalho, na aplicação do método, cada ponto utilizado tem a mesma relevância dos demais, o que nos leva a deduzir que nele supõe-se que uma dada variável independente envolvida (x) possa ser determinada com precisão absoluta. Mais objetivamente, estaremos considerando nulas ou desprezíveis as incertezas na determinação de tal variável. Ou seja, a aplicação do método dá-se a partir de pontos devidamente coletados e tratados segundo o grau de incerteza existente na observação do fenômeno.

Em geral, o ajuste conduz a um único tipo de função, exceto por um (ou mais) parâmetro(s), que pode(m) estar sujeito(s) à variação.

Analisemos, primeiramente, o caso em que tal função depende de apenas um único parâmetro α, ou seja, $f(x) = f(\alpha; x)$. Neste caso, a equação 2.1 terá a seguinte forma:

$$\S = S(\alpha) = \sum_{i=1}^{n} (y_i - f(\alpha; x_i))^2 \tag{2.2}$$

Para que o erro cometido no ajuste seja mínimo, temos que determinar o valor de α que satisfaça a equação

$$\frac{d}{d\alpha} S(\alpha) = 0 \tag{2.3}$$

Sabe-se, a princípio, que tal condição, se satisfeita, nos garante apenas a existência de um ponto crítico de S. Porém, o fato de S ser definida como a soma de termos quadráticos, nos permite demonstrar que esse ponto crítico é, na verdade, um mínimo da função $S(\alpha)$. Na verdade, a condição matemática que nos garante a existência de tal mínimo é dada por

$$\frac{d^2}{d\alpha^2}S(\alpha) > 0.$$

Aplicando tal condição ao último termo da equação 2.2, reescrevemos a condição:

$$\sum_{i=1}^{n}\left(\left(\frac{\partial}{\partial\alpha}f(\alpha;x_i)\right)^2 - (y_i - f(\alpha;x_i))\,\frac{\partial^2}{\partial\alpha^2}f(\alpha;x_i)\right) > 0 \tag{2.4}$$

A existência dessa desigualdade nos garante que o extremo da função $S(\alpha)$ é um mínimo.

Portanto, o processo envolvido no **Método dos Mínimos Quadrados** para obtenção de uma curva que dê representatividade a um conjunto de pontos, consiste em minimizar o erro quadrático médio cometido ao substituirmos o valor real y_i pelo valor $f(x_i)$, aproximado, que é o valor da função de ajuste f no ponto de abscissa $x = x_i$.

Ao executar tal procedimento, recairemos em um conjunto de equações lineares cujo número de equações dependerá do número de parâmetros a se determinar. De posse desse sistema de equações, deveremos trabalhar de modo a obtermos os valores dos parâmetros procurados. Como se sabe, se essas equações formarem um conjunto linearmente independente, o sistema terá solução única, o que é equivalente a dizer que a função S tem um único extremo.

Neste trabalho contemplaremos os casos em que cada parâmetro buscado desempenha na função de ajuste o papel de fator multiplicativo independente das variáveis envolvidas. Segundo Bassanezi (s. d.), ao contrário do que possa parecer, esta situação modela com satisfatória precisão e eficácia diversos fenômenos do mundo real.

Assim, prosseguindo com a suposição de uma função de ajuste composta de uma única variável e um único parâmetro, escrevemos:

$$f(\alpha;x) = \alpha\mu(x), \tag{2.5}$$

onde μ é uma função analítica na variável x perfeitamente definida. Ajustes com esta caraterística recebem o nome de **ajustes lineares**. A equação 2.2, rescrita segundo essas condições (parâmetro α como fator multiplicativo da função μ de x, que por sua vez independe de α), assume a seguinte forma:

$$S = S(\alpha) = \sum_{i=1}^{n}(y_i - \alpha\mu(x_i))^2 \tag{2.6}$$

Aplicando a condição expressa na equação 2.3 à equação 2.6, de modo a obtermos o valor do parâmetro α, temos:

$$\frac{d}{d\alpha}\left(\sum_{i=1}^{n}(y_i-\alpha\mu(x_i))^2\right)=0\Rightarrow\sum_{i=1}^{n}\mu(x_i)y_i-\sum_{i=1}^{n}\mu(x_i)^2=0\ \Rightarrow\alpha=\frac{\sum_{i=1}^{n}\mu(x_i)y_i}{\sum_{i=1}^{n}\mu(x_i)^2} \tag{2.7}$$

No caso da função de ajuste depender de mais de um parâmetro, ou seja, se $f(x) = f(\alpha_1, \alpha_2, ..., \alpha_m), m > 1$, a busca pela determinação desses parâmetros será baseada na ideia de que a condição de mínimo deve ser satisfeita por cada um desses m parâmetros, isto é,

$$\frac{\partial}{\partial\alpha_j}S(\alpha_1,\alpha_2,...,\alpha_m)=0,\ j=0,1,2,...,m.$$

O conjunto de equações assim obtido dará origem a um sistema de equações, no qual cada solução (lembrando que uma solução é um conjunto de parâmetros $\alpha_1, \alpha_2, ..., \alpha_m$ que satisfaça a todas as equações), supondo-se que haja soluções, representará um mínimo da função S. Deve-se ressaltar que, dependendo da forma que a função $f(x)$ possua, métodos analíticos serão insuficientes para se chegar a uma solução do sistema, devendo-se recorrer à métodos numéricos de cálculo para tal, o que foge do escopo deste trabalho.

Se a função $f(\alpha_1, \alpha_2, ..., \alpha_m)$ puder ser escrita como uma combinação linear de m funções da variável x, linearmente independentes entre si, onde os parâmetros $\alpha_1, \alpha_2, ..., \alpha_m$ são os coeficientes da combinação linear, dizemos que o ajuste é linear nos parâmetros buscados, e escrevemos:

$$f(x)=\sum_{j=1}^{m}\alpha_i\mu_i(x)=\alpha_1\mu_1(x)+\alpha_2\mu_2(x)+...+\alpha_m\mu_m(x) \tag{2.8}$$

Tal modelo simplifica bastante o procedimento matemático para se chegar aos parâmetros procurados. Além do mais, uma função da forma exibida na equação 2.8 satisfaz diretamente a condição para existência de mínimo expressa na equação 2.4

Façamos uma análise do caso em que há duas funções $\mu_1(x)$ e $\mu_2(x)$, cujas formas funcionais são previamente conhecidas, compondo a função de ajuste $f(x)$. Neste caso, o erro S, é dado, conforme a equação 2.2 por

$$S(\alpha)=\sum_{i=1}^{n}(y_i-\alpha_1\mu_1(x)-\alpha_2\mu_2(x))^2,$$

com a condição de mínimo

$$\frac{\partial}{\partial\alpha_1}S=\frac{\partial}{\partial\alpha_2}S=0. \tag{2.9}$$

Isto equivale a dizer que o erro cometido na busca da função de ajuste, expresso pela função S, será mínimo quando tivermos, simultaneamente:

$$\left(\sum_{i=1}^{n}\mu_1(x_i)\right)^2\alpha_1+\left(\sum_{i=1}^{n}\mu_1(x_i)\mu_2(x_i)\right)\alpha_2=\sum_{i=1}^{n}y_i\mu_1(x_i) \tag{2.10}$$

e

$$\left(\sum_{i=1}^{n}\mu_1(x_i)\mu_2(x_i)\right)\alpha_1+\left(\sum_{i=1}^{n}\mu_2(x_i)\right)^2\alpha_2+=\sum_{i=1}^{n}y_i\mu_2(x_i) \tag{2.11}$$

Como se vê, as equações 2.10 e 2.11 formam um conjunto de duas equações algébricas, cuja solução retorna os valores dos parâmetros α_1 e α_2. Uma vez que as funções $\mu_1(x)$ e $\mu_2(x)$ são linearmente independentes entre si, tais equações também o serão e o sistema composto por elas pode ser resolvido de forma relativamente simples.

Vale ressaltar mais uma vez que os funções $\mu(x_i)$ que compõem a função de ajuste, além de terem suas formas previamente conhecidas, também não dependem dos parâmetros envolvidos no ajuste. Esta modelagem simplifica de modo considerável os procedimentos na busca da função de ajuste. Modelos que fujam desse padrão recairão em sistemas que, caso tenham solução, deverão, como já dito, ser resolvidos por métodos numéricos que não serão tratados neste trabalho.

2.4 Ajuste Linear Simples

As relações mais simples envolvendo duas variáveis são as relações lineares (RUGGIERO; LOPES, 1996). Muitos fenômenos da natureza e outros oriundos de atividades humanas podem ser modelados conforme uma relação linear entre variáveis. Por exemplo: a função P que relaciona o peso de um corpo e a sua massa m, dada por $P(m) = gm$, onde g representa o valor, constante nas proximidades da superfície da Terra, da aceleração gravitacional; e a função V que retorna o valor de uma prestação, no regime de juros simples, em função de d dias de atraso, a uma taxa fixa i, dada por $V(d) = V_0 + diV_0$, onde V_0 é o valor da prestação original.

Há casos também em que uma adequada manipulação matemática de relações não lineares conduz a novas variáveis que compartilham entre si uma relação linear . Tal artifício justifica-se plenamente pela facilidade de manipulação e representação desse tipo de função.

Vamos aplicar os conceitos matemáticos vistos na seção 2.3 para obtermos o procedimento matemático a ser utilizado na realização do ajuste, através do **Métodos dos Mínimos Quadrados**, a um modelo linear. Inicialmente, trataremos do ajuste a modelos onde temos apenas uma variável independente, chamado por isso de **Ajuste Linear Simples** ou **Regressão Linear Simples**.

Uma relação de linearidade simples pode ser completamente definida pela expressão:

$$f(x) = \alpha_0 + \alpha_1 x, \tag{2.12}$$

Observemos a existência de uma única variável independente (x). Neste caso, num procedimento de ajuste, o esforço consiste na determinação de valores adequados para os parâmetros α_0 e α_1.

Assim sendo, suponhamos novamente um conjunto de n pontos como referência para a obtenção de uma função de ajuste linear, expressa na forma da equação 2.12. O desvio quadrático S, conforme visto na equação 2.1, é dado por:

$$S(\alpha_0, \alpha_1) = \sum_{i=1}^{n} (y_i - \alpha_0 - \alpha_1 x_i)^2$$

Em consonância com o teorema 2.3, podemos dizer que para que o desvio S seja mínimo, os parâmetros α_0 e α_1 devem ser escolhidos de modo a satisfazerem às seguintes condições:

$$\frac{\partial S(\alpha_0, \alpha_1)}{\partial \alpha_0} = 0$$

e

$$\frac{\partial S(\alpha_0, \alpha_1)}{\partial \alpha_1} = 0$$

Utilizando o conceito de derivadas parciais visto na seção 2.2.2 e aplicando as regras de derivação a estas equações, elas transformam-se, respectivamente, em:

$$\sum_{i=1}^{n} (y_i - \alpha_0 - \alpha_1 x_i) = 0$$

e

$$\sum_{i=1}^{n} (x_i y_i - \alpha_0 x_i - \alpha_1 x_i^2) = 0$$

Que, por sua vez, podem ser reescritas da seguinte forma:

$$n\alpha_0 + \left(\sum_{i=1}^{n} x_i\right)\alpha_1 = \sum_{i=1}^{n} y_i$$

e

$$\left(\sum_{i=1}^{n} x_i\right)\alpha_0 + \left(\sum_{i=1}^{n} x_i^2\right)\alpha_1 = \sum_{i=1}^{n} x_i y_i$$

Como se vê, essas duas equações compõem um sistema de duas equações lineares com duas incógnitas, α_0 e α_1. Elas são chamadas de **equações normais** do sistema construído. Dispondo na forma matricial, obtemos:

$$\begin{bmatrix} n & \sum_{i=1}^{n} x_i \\ \sum_{i=1}^{n} x_i & \sum_{i=1}^{n} x_i^2 \end{bmatrix} \begin{bmatrix} \alpha_0 \\ \alpha_1 \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{n} y_i \\ \sum_{i=1}^{n} x_i y_i \end{bmatrix} \tag{2.13}$$

A solução de tal sistema, segundo a regra de Cramer, é dada por:

$$\alpha_0 = \frac{\sum_{i=1}^{n} y_i \sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i y_i\right) \sum_{i=1}^{n} x_i}{n \sum_{i=1}^{n} x_i^2 - \left(\sum_{n=i}^{n} x_i\right)^2} \tag{2.14}$$

$$\alpha_1 = \frac{n \left(\sum_{i=1}^{n} x_i y_i\right) - \sum_{i=1}^{n} y_i \sum_{i=1}^{n} x_i}{n \sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2} \tag{2.15}$$

Vejamos um exemplo completo de obtenção de ajuste linear pelo **Método dos Mínimos Quadrados**.

Exemplo 2.2 ***Retirado de Burden e Faires (2003)****. Calcule o polinômio linear de mínimos quadrados para os dados da tabela 2.3:*

x	y
$0,3$	$1,8$
$2,75$	$1,9$
$4,5$	$3,1$
$5,95$	$3,9$
$7,8$	$3,3$

Tabela 2.3 – Dados referentes ao exemplo 2.2

Solução: *A partir dos valores contidos na tabela dada, construímos a tabela 2.4, contendo os valores dos somatórios contidos nas equações 2.14 e 2.15*

i	x_i	y_i	x_i^2	$x_i y_i$	y_i^2
1	$0,3$	$1,8$	$0,09$	$0,54$	$3,24$
2	$2,7$	$1,9$	$7,29$	$5,13$	$3,61$
3	$4,5$	$3,1$	$20,25$	$13,95$	$9,61$
4	$5,9$	$3,9$	$34,81$	$23,01$	$15,21$
5	$7,8$	$3,3$	$60,84$	$25,74$	$10,89$
$\sum$	$21,2$	$14,0$	$123,28$	$68,37$	$42,56$

Tabela 2.4 – Somatórios referentes ao exemplo 2.2

Levando os valores obtidos às equações 2.14 e 2.15, temos:

$$\alpha_0 = \frac{68,37 \cdot 21,20 - 14,00 \cdot 123,28}{(5 \cdot 123,28 - (21,20)^2} \Rightarrow \alpha_0 = 1,6560$$

$$\alpha_1 = \frac{5 \cdot 68,37 - 21,2 \cdot 14,00}{5 \cdot 123,38 - (21,20)^2} \Rightarrow \alpha_1 = 0,2698$$

Portanto, a função de ajuste é dada por $f(x) = 1,6560 + 0,2698x$. A figura 2.11 exibe o gráfico da função de ajuste obtida neste exemplo e os pontos que compõem o diagrama de dispersão do conjunto de pares ordenados dado.

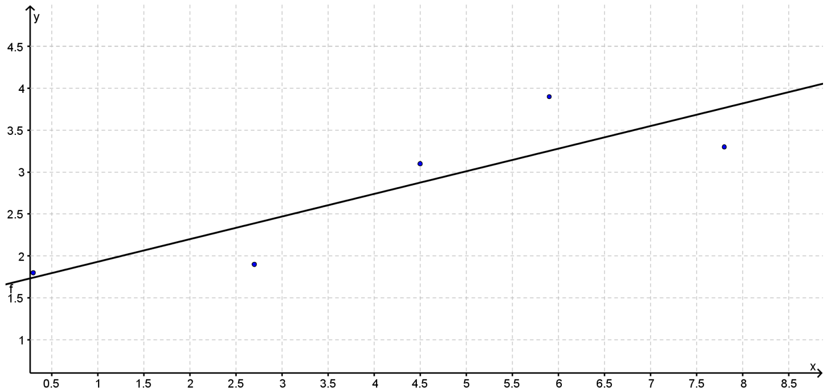

Figura 2.11 – Gráfico da função de ajuste $f(x) = 1,6560 + 0,2698x$
Fonte: Elaboração própria

2.5 Ajuste Linear Múltiplo

Trata-se do modelo, teoricamente mais completo, em que o ajuste (ainda linear) dispõe-se a representar um fenômeno em que há mais de uma variável independente envolvida, ou seja, um modelo expresso matematicamente por $y = f(x_1, \ldots, x_p), p \geq 2$. Ou ainda, $y = f_1(x_1) + ... + f_p(x_p), p \geq 2$, onde $f_1, \ldots, f_p$ são p funções lineares nas variáveis $x_1, \ldots, x_p$, respectivamente.

Levando em consideração a natureza linear dessas p funções e tomando como referência a equação 2.12, escrevemos:

$$f_1(x_1) = \alpha_{0\,1} + \alpha_1 x_1, \ldots, f_p(x_p) = \alpha_{0\,p} + \alpha_p x_p.$$

Fazendo $\alpha_{0\,1} + \cdots + \alpha_{0\,p} = \beta$, teremos:

$$f(x) = \beta + \alpha_1 x_1 + \cdots + \alpha_p x_p \tag{2.16}$$

Assumindo, como em todo este trabalho, a independência entre os p parâmetros $\alpha_1, \ldots, \alpha_p$ e as respectivas p variáveis $x_1, \ldots, x_p$, suponhamos a existência de um fenômeno com características que permitam tal modelagem matemática e suponhamos também a existência de n pontos $(y, x_{1i}, \ldots, x_{pi})$ que representem tal fenômeno. Assim sendo, o desvio quadrático S oriundo do erro cometido ao se obter uma função de ajuste para o fenômeno será dado por $S = \displaystyle\sum_{i=1}^{n}(y_i - f(x_{1i}, \ldots, x_{pi}))^2$, o que nos permite escrever :

$$S(\beta, \alpha_1, \ldots, \alpha_p) = \sum_{i=1}^{n}(y_i - \beta - \alpha_1 x_{1i} - \cdots - \alpha_p x_{pi})^2 \tag{2.17}$$

Como visto no teorema 2.3, o erro quadrático S será mínimo quando os parâmetros $\alpha_1, \ldots, \alpha_p$ forem tais que:

$$\frac{\partial S(\beta, \alpha_1, \ldots, \alpha_p)}{\partial \beta} = 0$$
$$\frac{\partial S(\beta, \alpha_1, \ldots, \alpha_p)}{\partial \alpha_1} = 0$$
$$\ldots$$
$$\frac{\partial S(\beta, \alpha_1, \ldots, \alpha_p)}{\partial \alpha_p} = 0$$

Depois de aplicar as regras de derivação para calcular as derivadas parciais da função S expressas nessas equações, obtemos:

$$\sum_{i=1}^{n}(y_i - \beta - \alpha_1 x_{1i} - \cdots - \alpha_p x_{pi}) = 0$$
$$\sum_{i=1}^{n}(y_i - \beta - \alpha_1 x_{1i} - \cdots - \alpha_p x_{pi})x_{1i} = 0$$
$$\ldots$$
$$\sum_{i=1}^{n}(y_i - \beta - \alpha_1 x_{1i} - \cdots - \alpha_p x_{pi})x_{pi} = 0$$

Desenvolvendo cada um desses produtos e rearranjando as equações, obtemos:

$$\sum_{i=1}^{n} \beta + \sum_{i=1}^{n} \alpha_1 x_{1i} + \cdots + \sum_{i=1}^{n} \alpha_p x_{pi} = \sum_{i=1}^{n} y_i$$
$$\sum_{i=1}^{n} \beta x_{1i} + \sum_{i=1}^{n} \alpha_1 x_{1i} x_{1i} + \cdots + \sum_{i=1}^{n} \alpha_p x_{1i} x_{pi} = \sum_{i=1} x_{1i} y_i$$
$$\ldots$$
$$\sum_{i=1}^{n} \beta x_{pi} + \sum_{i=1}^{n} \alpha_1 x_{1i} x_{pi} + \cdots + \sum_{i=1}^{n} \alpha_p x_{pi} x_{pi} = \sum_{i=1}^{n} x_{pi} y_i$$

O conjunto de equações assim obtidos é um sistema linear de $p+1$ equações com $p+1$ incógnitas que, quando solucionado, fornece os parâmetros da função de ajuste. Simplificando a notação e escrevendo na forma matricial, temos:

$$\begin{bmatrix} n & \sum x_{1i} & \cdots & \sum x_{pi} \\ \sum x_{1i} & \sum x_{1i} x_{1i} & \cdots & \sum x_{1i} x_{pi} \\ \vdots & \vdots & \ddots & \vdots \\ \sum x_{pi} & \sum x_{1i} x_{pi} & \cdots & \sum x_{pi} x_{pi} \end{bmatrix} \begin{bmatrix} \beta \\ \alpha_1 \\ \vdots \\ \alpha_p \end{bmatrix} = \begin{bmatrix} \sum y_i \\ \sum x_{1i} y_i \\ \vdots \\ \sum x_{pi} y_i \end{bmatrix} \tag{2.18}$$

Obviamente, quanto maior for o número de variáveis independentes envolvidas no fenômeno, maior será o número de equações a serem resolvidas simultaneamente, ou seja, o sistema de equações lineares resultante torna-se mais complexo, dificultando a sua resolução por métodos analíticos. Neste trabalho, iremos sugerir o auxílio de processos computacionais na resolução de sistemas mais complexos.

Exemplo 2.3 *Os dados contidos na tabela 2.5 expressam quantidades vendidas, preços e gastos com divulgação dos principais produtos de uma determinada empresa. Segundo um modelo linear e fazendo y = quantidade vendida, x_1 = preço e x_2 = investimento, determine a equação de regressão de y em função de x_1 e x_2.*

Quantidade(kg)	**Preço(R$)**	**Investimento(mil R$)**
70	90	630
90	80	720
100	70	700
90	70	625
105	70	735
80	70	560
110	65	715
125	60	750
115	60	690
130	55	715
130	50	650

Tabela 2.5 – Dados referentes ao exemplo 2.3

Solução: *A função de ajuste tem a forma $y = \beta + \alpha_1 x_1 + \alpha_2 x_2$ e a forma matricial do sistema de equações que nos dará os parâmetros procurados é a seguinte:*

$$\begin{bmatrix} n & \sum x_{1i} & \sum x_{2i} \\ \sum x_{1i} & \sum x_{1i}x_{1i} & \sum x_{1i}x_{2i} \\ \sum x_{2i} & \sum x_{1i}x_{2i} & \sum x_{2i}x_{2i} \end{bmatrix} \begin{bmatrix} \beta \\ \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} \sum y_i \\ \sum x_{1i}y_i \\ \sum x_{2i}y_i \end{bmatrix}$$

Tomando por base a equação 2.18, geramos a tabela 2.6.

i	x_1	x_2	y	x_1^2	x_2^2	x_1x_2	x_1y	x_2y
1	100	550	55	10000	302500	55000	5500	30250
2	90	630	70	8100	396900	56700	6300	44100
3	80	720	90	6400	518400	57600	7200	64800
4	70	700	100	4900	490000	49000	7000	70000
5	70	625	90	4900	390625	43750	6300	56250
6	70	735	105	4900	540225	51450	7350	77175
7	70	560	80	4900	313600	39200	5600	44800
8	65	715	110	4225	511225	46475	7150	78650
9	60	750	125	3600	562500	45000	7500	93750
10	60	690	115	3600	476100	41400	6900	79350
11	55	715	130	3025	511225	39325	7150	92950
12	50	650	130	2500	422500	32500	6500	84500
$\sum$	840	8040	1200	61050	5435800	557400	80450	816575

Tabela 2.6 – Somatórios referentes ao exemplo 2.3

Com os valores obtidos, escrevemos, na forma matricial, as equações normais do sistema:

$$\begin{bmatrix} 12 & 840 & 8040 \\ 840 & 61050 & 557400 \\ 8040 & 557400 & 5435800 \end{bmatrix} \begin{bmatrix} \beta \\ \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} 1200 \\ 80450 \\ 816575 \end{bmatrix}$$

E, utilizando a regra de Cramer:

$$\Delta = \det\left(\begin{bmatrix} 12 & 840 & 8040 \\ 840 & 61050 & 557400 \\ 8040 & 557400 & 5435800 \end{bmatrix}\right) = 973080000$$

$$\Delta_\beta = \det\left(\begin{bmatrix} 1200 & 840 & 8040 \\ 80450 & 61050 & 557400 \\ 816575 & 557400 & 5435800 \end{bmatrix}\right) = 113030850000$$

$$\Delta_1 = \det\left(\begin{bmatrix} 12 & 1200 & 8040 \\ 840 & 80450 & 557400 \\ 8040 & 816575 & 5435800 \end{bmatrix}\right) = -1272540000$$

$$\Delta_2 = \det\left(\begin{bmatrix} 12 & 840 & 1200 \\ 840 & 61050 & 80450 \\ 8040 & 557400 & 816575 \end{bmatrix}\right) = 109485000$$

$$\beta = \frac{\Delta_\beta}{\Delta} = \frac{113030850000}{973080000} = 116,1578$$

$$\alpha_1 = \frac{\Delta_1}{\Delta} = \frac{-1272540000}{973080000} = -1,3077$$

$$\alpha_2 = \frac{\Delta_1}{\Delta} = \frac{109485000}{973080000} = 0,11258$$

Finalmente, escrevemos a equação de ajuste: $y = 116,16 - 1,31x_1 + 0,11x_2$*.*

2.6 Ajuste Polinomial

É o tipo de ajuste no qual se busca uma curva associada a um polinômio de grau maior ou igual a dois. Ou seja, a função de ajuste tem a forma

$$f(x) = \alpha_0 + \alpha_1 x + \cdots + \alpha_p x^p, p \geq 2 \tag{2.19}$$

Observa-se que ao fazermos, na equação 2.19, $x = x_1, x^2 = x_2, \ldots, x^p = x_p$, a função de ajuste assume a forma $f(x) = \alpha_0 + \alpha_1 x_1 + \cdots + \alpha_p x_p$, que nada mais é que a forma buscada no ajuste linear múltiplo, visto na seção anterior. Assim sendo, todo procedimento matemático realizado para obtenção da equação 2.18 pode ser aqui aplicado. Fazendo então as substituições sugeridas e levando adiante tal procedimento, obtemos o sistema de equações que conduzem aos parâmetros do ajuste. Simplificando a notação, escrevemos:

$$\begin{bmatrix} n & \sum x_i & \cdots & \sum x_i^p \\ \sum x_i & \sum x_i^2 & \cdots & \sum x_i^{p+1} \\ \vdots & \vdots & \ddots & \vdots \\ \sum x_i^p & \sum x_i^{g+1} & \cdots & \sum x_i^{2p} \end{bmatrix} \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \vdots \\ \alpha_p \end{bmatrix} = \begin{bmatrix} \sum y_i \\ \sum x_i y_i \\ \vdots \\ \sum x_i^p y_i \end{bmatrix}. \tag{2.20}$$

Onde n é o número de pontos experimentais a partir dos quais se busca o ajuste.

Obviamente, assim como no caso do **ajuste linear múltiplo**, onde o aumento do número de variáveis independentes aumenta a complexidade do problema, o mesmo se dá no **ajuste polinomial** com relação ao aumento do grau do polinômio de ajuste.

Exemplo 2.4 *Ajuste um polinômio de grau dois aos dados contidos na tabela 2.7.*

x	$-2,0$	$-1,5$	$0,0$	$1,0$	$2,2$	$3,1$
y	$-30,5$	$-20,2$	$-3,3$	$8,9$	$16,8$	$21,4$

Tabela 2.7 – Dados referentes ao exemplo 2.4

Solução: *A função de ajuste terá a forma $\alpha_0 + \alpha_1 x + \alpha_2 x^2$ e a forma matricial do sistema linear que levará aos parâmetros procurados é*

$$\begin{bmatrix} n & \sum x_i & \sum x_i^2 \\ \sum x_i & \sum x_i^2 & \sum x_i^3 \\ \sum x_i^2 & \sum x_i^3 & \sum x_i^4 \end{bmatrix} \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} \sum y_i \\ \sum x_i y_i \\ \sum x_i^2 y_i \end{bmatrix}$$

A tabela 2.8 contém os valores dos somatórios necessários para gerar as equações normais do sistema.

i	x_i	y_i	x_i^2	x_i^3	x_i^4	$x_i y_i$	$x_i^2 y$
1	$-2,0$	$-30,5$	$4,00$	$-8,000$	$16,0000$	$61,00$	$-122,0000$
2	$-1,5$	$-20,2$	$2,25$	$-3,375$	$5,0625$	$30,30$	$-45,4500$
31	$0,0$	$-3,3$	$0,00$	$0,000$	$0,0000$	$0,00$	$0,0000$
4	$1,0$	$8,9$	$1,00$	$1,000$	$1,0000$	$8,90$	$8,9000$
5	$2,2$	$16,8$	$4,84$	$10,648$	$23,4256$	$36,96$	$81,3120$
6	$3,1$	$21,4$	$9,61$	$29,791$	$92,3521$	$66,34$	$205,6540$
$\sum$	$2,8$	$-6,9$	$21,70$	$30,064$	$137,8402$	$203,50$	$128,4160$

Tabela 2.8 – Somatórios referentes ao exemplo 2.4

Levando os valores obtidos ao sistema linear que dará a solução do problema, temos:

$$\begin{bmatrix} 6 & 2,8 & 21,7 \\ 2,8 & 21,7 & 30,064 \\ 21,7 & 30,064 & 137,8402 \end{bmatrix} \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} -6,9 \\ 203,5 \\ 128,416 \end{bmatrix}$$

Resolvendo o sistema, obtemos: $\alpha_0 = -2,018$*,* $\alpha_1 = 11,33$ *e* $\alpha_2 = -1,222$*. Portanto, a função de ajuste é* $y = -2,018 + 11,33x - 1,222x^2$*.*

2.7 Qualidade do Ajuste

Um estudo detalhado sobre a qualidade do ajuste nos levaria por caminhos e temas que fogem ao objetivo deste trabalho. Por este motivo, iremos apresentar apenas dois indicadores, ambos de grande relevância e utilização, para permitirão efetuar uma avaliação satisfatória sobre a qualidade dos ajustes que serão realizados nas atividades propostas neste trabalho.

2.7.1 O Coeficiente de Determinação

Consideremos novamente o conjunto de n pontos representando n ocorrências de um determinado fenômeno. Sejam y a variável dependente, $\hat{y}$ a variável dependente fornecida pela função de ajuste e $\overline{y} = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} y_i$ o valor médio da grandeza y. Sem maiores dificuldades, podemos constatar que é verdadeira a seguinte equação:

$$y_i = (y_i - \hat{y}_i) + (\hat{y}_i - \overline{y}) + \overline{y} \Rightarrow y_i - \overline{y} = (y_i - \hat{y}_i) + (\hat{y}_i - \overline{y}) \tag{2.21}$$

Elevando ao quadrado ambos os lados da equação:

$$(y_i - \overline{y})^2 = (y_i - \hat{y}_i)^2 + 2(y_i - \hat{y}_i)(\hat{y}_i - \overline{y}) + (\hat{y}_i - \overline{y})^2$$

Em seguida, fazendo o somatório para $i = 1, 2, \cdots, n$:

$$\sum_{i=1}^{n}(y_i - \overline{y})^2 = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2 + 2\sum_{i=1}^{n}(y_i - \hat{y}_i)(\hat{y}_i - \overline{y}) + \sum_{i=1}^{n}(\hat{y}_i - \overline{y})^2$$

É possível demonstrar que $\displaystyle\sum_{i=1}^{n}(y_i - \hat{y}_i)(\hat{y}_i - \overline{y}) = 0$, o que nos permite escrever:

$$\sum_{i=1}^{n}(y_i - \overline{y})^2 = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2 + \sum_{i=1}^{n}(\hat{y}_i - \overline{y})^2 \tag{2.22}$$

Na equação 2.22 chamamos os termos $sum_{i=1}^{n}(y_i - \overline{y})^2$, $\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$ e $\sum_{i=1}^{n}(\hat{y}_i - \overline{y})^2$, respectivamente, de **soma de quadrados total (**SQ_{tot}**)**, **soma de quadrados residual (**SQ_{res}**)** e **soma de quadrados devido à regressão (**SQ_{reg}**)**, e escrevemos:

$$SQ_{tot} = SQ_{res} + SQ_{reg} \tag{2.23}$$

O **coeficiente de determinação**, que chamaremos de r^2, é definido como a razão entre a soma dos **quadrados devido à regressão** e a **soma dos quadrados total**, ou seja:

$$r^2 = \frac{SQ_{reg}}{SQ_{tot}} \tag{2.24}$$

$$\Rightarrow r^2 = \frac{\displaystyle\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\displaystyle\sum_{i=1}^{n}(y_i - \overline{y})^2} \tag{2.25}$$

Aplicando o resultado contido na equação 2.23 à equação 2.24, escrevemos:

$$r^2 = \frac{SQ_{reg}}{SQ_{tot}} \Rightarrow r^2 = \frac{SQ_{tot} - SQ_{res}}{SQ_{tot}} \Rightarrow r^2 = 1 - \frac{SQ_{res}}{SQ_{tot}} \tag{2.26}$$

Como, $SQ_{tot} = \displaystyle\sum_{i=1}^{n}(y_i - \overline{y})^2 = \sum_{i=1}^{2} y_i^2 - 2\overline{y}\sum_{i=1}^{n} y_i + n\overline{y}^2$ e, como já visto, $\overline{y} = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} y_i$, então $\displaystyle\sum_{i=1}^{n}(y_i - \overline{y})^2 = \sum_{i=1}^{n} y_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n} y_i\right)^2$. Aplicando este resultado à equação 2.26, obtemos:

$$r^2 = 1 - \frac{\displaystyle\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\displaystyle\sum y_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n} y_i\right)^2} \tag{2.27}$$

O **coeficiente de determinação** representa a dimensão da variação total dos dados em torno da média dos valores da variável dependente. De fato, analisando a equação 2.24, vemos que o numerador da fração existente no segundo membro representa a soma dos quadrados dos desvios dos valores da função de ajuste em relação à media dos valores experimentais, enquanto que o denominador representa a soma dos quadrados dos desvios de cada ponto experimental em relação à esta média. Pode ser interpretado como a indicação do quão preciso é o modelo para explicar a variável dependente y.

O **coeficiente de determinação** r^2 é sempre positivo e pertence ao intervalo $[0, 1]$. Este indicador é mais recomendado para o **Ajuste Linear Simples** e o motivo será exposto na subseção 2.7.2. Teoricamente, tratando-se do **Ajuste Linear Simples**, quanto mais próximo de 1 for o valor de r^2, melhor será a qualidade do ajuste. Um **coeficiente de determinação** igual a 1, neste caso, significa um (improvável) ajuste perfeito.

2.7.2 O Coeficiente de Determinação Ajustado ($r^2_{ajust.}$)

Nem sempre um valor alto do coeficiente de determinação r^2 implica uma boa qualidade do ajuste, já que adição de uma nova variável ao modelo sempre aumenta o valor desse indicador, pois a soma dos quadrados devido à regressão sempre aumenta com a introdução de uma nova variável, independentemente de sua significância para o modelo. Sendo assim, modelos com um elevado valor de r^2 podem conduzir a estimativas pouco confiáveis do valor da variável dependente, comprometendo a eficácia do ajuste.

Por este motivo, em modelos com mais de uma variável independente, dá-se preferência à utilização do **coeficiente de determinação ajustado ($r^2_{ajust.}$)**:

$$r^2_{ajust.} = 1 - \frac{\dfrac{SQ_{res}}{n-p}}{\dfrac{SQ_{tot}}{n-1}} = 1 - \left(\frac{n-1}{n-p}\right)(1-r^2), \tag{2.28}$$

onde p é o número de parâmetros (coeficientes) buscados no ajuste.

Este indicador dá uma ideia melhor da proporção de variação da variável dependente explicada pelo modelo, uma vez que leva em consideração o número de variáveis independentes envolvidas no ajuste. Ao contrário do que se observa com o **coeficiente de determinação** r^2, o **coeficiente de determinação ajustado ($r^2_{ajust.}$)** não aumenta sempre que uma nova variável é adicionada ao modelo, só apresentando aumento se de alguma maneira houver vantagem na adição de uma nova variável. De fato, se forem adicionadas variáveis não significativas, o valor de $r^2_{ajust.}$, na maioria das vezes, decresce. Quando a diferença entrer^2 e $r^2_{ajust.}$ é acentuada, há uma boa possibilidade de que tenham sido incluídos no modelo termos estatisticamente não significativos.

Capítulo 3

O Método dos Mínimos Quadrados no Ensino Médio

3.1 Apresentação do Método

Este capítulo destina-se a sugerir a apresentação do **Método dos Mínimos Quadrados** de uma forma acessível aos alunos do Ensino Médio e, em seguida, trabalhar sua aplicação por meio de atividades significativas. A série indicada para se aplicar o que será exposto neste capítulo é o 3º ano deste segmento, pois o aluno nesta etapa, espera-se, já deve ser detentor de alguns conteúdos e conceitos necessários para entendimento e aplicação do método.

Assim, a aplicação do método demanda o domínio básico de temas como determinantes, resolução de sistemas de equações lineares, polinômios e funções, principalmente as polinomiais, visto que os ajustes aqui apresentados contemplam, na maioria das vezes, esta categoria de funções. Sugere-se, então, uma boa revisão desses assuntos antes de iniciar-se a exposição do método propriamente dito.

Consultando outros trabalhos (Silva (2014), Oliveira e Vertuan (2009), Souza (2014), etc.), percebi que a abordagem ao método, quando direcionada a alunos do Ensino Médio, costuma ser feita quase que exclusivamente através de tratamento e procedimentos matriciais, priorizando a apresentação quase imediata das equações normais que possibilitarão a obtenção das funções de ajuste, com pouca ou nenhuma ênfase no elemento motivador do método - o processo de otimização que consiste na minimização do erro na busca do melhor ajuste.

Muito provavelmente, isto se deve ao fato de que alguns dos conteúdos envolvidos na dedução matemática do método ainda não são de domínio dos alunos do Ensino Médio, estando reservados para (alguns) cursos de graduação. Assim, constitui-se um verdadeiro desafio apresentar o método ao aluno do Ensino Médio sem tomar atalhos que ocultem sua essência, de modo a evitar prejuízo na apresentação do método como ferramenta eficaz e consistente para o propósito ao qual se destina.

Como visto na seção 2.3, o conceito de derivada tem um papel fundamental no artifício de otimização que dará origem ao **Método dos Mínimos Quadrados**. Portanto, introduzir este conceito, mesmo no Ensino Médio, pelo menos até um nível que torne o aluno apto a perceber a fundamentação do método me parece de grande valia.

Partindo de um assunto acessível a um aluno do Ensino Médio, a função quadrática, podemos criar a motivação necessária para a abordagem do conceito de derivadas. Comecemos tratando do problema que consiste em determinar o valor mínimo de uma função quadrática.

Seja

$$\begin{aligned} f : \mathbb{R} &\to \mathbb{R} \\ x &\mapsto ax^2 + bx + c, \quad a > 0 \end{aligned} \tag{3.1}$$

Sabemos que a função corresponde a uma parábola com a concavidade voltada para cima, admitindo, portanto, um ponto mínimo. Para todo $x \in \mathbb{R}$, temos que $f(x) = ax^2+bx+c$, que podemos escrever, completando o quadrado, como $f(x) = ax^2 + bx + \frac{b^2}{4a} - \frac{b^2}{4a} + c \Rightarrow$ $f(x) = \left(\sqrt{a}x - \frac{b}{2\sqrt{a}}\right)^2 - \frac{b^2}{4a} + c$.

Observemos agora que o termo $\left(\sqrt{a}x - \frac{b}{2\sqrt{a}}\right)^2$, $\forall x \in \mathbb{R}$, enquanto que o termo $-\frac{b^2}{4a} + c$, por sua vez, não depende da variável x. Concluímos, portanto, que o valor mínimo de f ocorrerá quando $\left(\sqrt{a}x - \frac{b}{2\sqrt{a}}\right)^2$ for mínimo, isto é, quando tivermos $\left(\sqrt{a}x - \frac{b}{2\sqrt{a}}\right)^2 = 0$. Temos, então:

$$\begin{aligned} \left(\sqrt{a}x - \frac{b}{2\sqrt{a}}\right)^2 &= 0 \\ \Rightarrow \sqrt{a}x - \frac{b}{2\sqrt{a}} &= 0 \\ \Rightarrow x &= -\frac{b}{2a} \end{aligned} \tag{3.2}$$

Da equação 3.2, concluímos que f tem um mínimo quando o termo quadrático $\left(\sqrt{a}x - \frac{b}{2\sqrt{a}}\right)^2$ é mínimo, isto é, igual a zero. Este é, essencialmente, o **Método dos Mínimos Quadrados**.

Vamos agora definir uma certa classe de funções já conhecida.

Definição 3.1 *(Funções polinomiais reais de grau n). Seja a função $f : \mathbb{R} \to \mathbb{R}$, definida* *por $f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$, tal que $a_n, a_{n-1}, \cdots, a_1, a_0 \in \mathbb{R}$. Então f é uma função polinomial de grau n.*

Vamos agora definir algumas operações com as funções polinomiais.

Definição 3.2 *(Derivadas de funções polinomiais reais de grau n). Seja a função $f : \mathbb{R} \to \mathbb{R}$, definida por $f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$ uma função polinomial real de grau n. Definimos a derivada de f com relação à variável x como a função $f' : \mathbb{R} \to \mathbb{R}$, dada por $f'(x) = na_n x^{n-1} + (n-1)a_{n-1}x^{n-2} + \cdots + a_1$.*

Exemplo 3.1 *Seja a função $f : \mathbb{R} \to \mathbb{R}$, definida por $f(x) = 3x^2$. Então, $\forall x \in \mathbb{R}, f'(x) = 6x$.*

Exemplo 3.2 *Seja a função $g : \mathbb{R} \to \mathbb{R}$, definida por $g(x) = ax^4 + bx^2 + a_0$. Então, $\forall x \in \mathbb{R}, g'(x) = 4ax^3 + 2bx$.*

Portanto, dada

$$\begin{aligned} f : \mathbb{R} &\to \mathbb{R} \\ x &\mapsto ax^2 + bx + c, \quad a > 0, \end{aligned}$$

$\forall x \in \mathbb{R}$ existe $f'(x)$, com $f'(x) = 2ax + b$. Se quisermos o ponto x tal que $f'(x) = 0$, encontraremos $x = -\dfrac{b}{2a}$, resultado idêntico à equação 3.2. Isto nos leva a concluir que o valor mínimo da função f ocorre no mesmo ponto em que sua derivada se anula. Este é um resultado muito importante e será o ponto de partida para a apresentação do método aos alunos do Ensino Médio.

Consideremos agora o seguinte problema: Suponhamos o estudo da produção trimestral de certo produto agrícola como função da quantidade de chuva no período e que coletamos os seguintes dados:$(20, 96)$,$(15, 77)$ e $(30, 143)$, onde a primeira coordenada é a quantidade de chuva em ml no período e a segunda, o número de sacas do produto obtido. Com isso queremos determinar qual a curva que se ajusta melhor a esses dados, de modo que se a equação da curva obtida for $y = f(x)$, o erro cometido entre o previsto pela curva e o valor real no trimestre, ao ser somado em todos os três trimestres, seja mínimo. Além disso, com base nos resultados dos três trimestres, desejamos estimar a produção de sacas no quarto trimestre.

Utilizando os argumentos expostos na seção 2.3, conclui-se pela conveniência da utilização do erro quadrático total no tratamento do problema e apresenta-se a equação 2.1.

Na busca pela natureza da função de ajuste, utilizamos os dados disponíveis para traçamos o diagrama de dispersão do conjunto, mostrado na figura 3.12, o qual nos conduz à hipótese que a quantidade de sacas produzidas (y) relaciona-se com o volume de chuva no trimestre (x) através de uma relação linear.

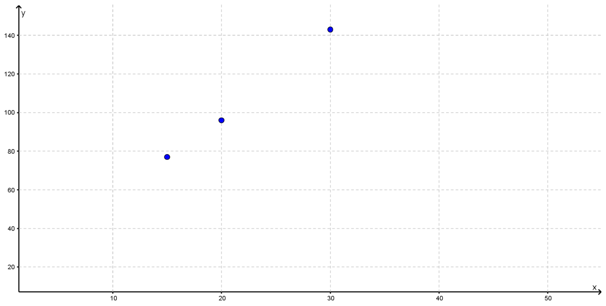

Figura 3.12 – Diagrama de dispersão para o conjunto de pontos $(20, 96)$, $(15, 77)$ e $(30, 143)$
Fonte: Elaboração própria

Escrevemos, então:

$$y = ax, \tag{3.3}$$

sendo $a \in \mathbb{R}$ o parâmetro a ser determinado para obtermos a relação entre produção e quantidade de chuva.

Para a determinação do valor de a, consideremos a soma dos desvios (ou erros) de cada ordenada y_i e a correspondente abscissa x_i, isto é:

$$(y_1 - ax_1)^2 + (y_2 - ax_2)^2 + (y_3 - ax_3)^2 = (96 - 20a)^2 + (77 - 15a)^{+}(143 - 30a)^2 =$$

$$= 1525a^2 - 14730a + 35594 = \left(a - \frac{1473}{350}\right)^2 + 35594 - \left(\frac{1473}{350}\right)^2$$

Portanto, a fim de minimizar o erro total, temos que $a = \dfrac{1470}{350} = 4,83$. Se chamarmos de S a expressão que representa o erro total, poderemos escrever $S(a) = 1525a^2 - 14730a + 3559$, com derivada, segundo a definição 3.2, $S'(a) = 3050a - 14730$ e valor mínimo quando $a = \dfrac{14730}{3050} = 4,82$, o que corrobora a afirmação de que o valor mínimo da função quadrática ocorre no ponto em que a sua derivada se anula.

A função de ajuste obtida é, portanto, $y = 4,83x$, cujo gráfico está na figura 3.13:

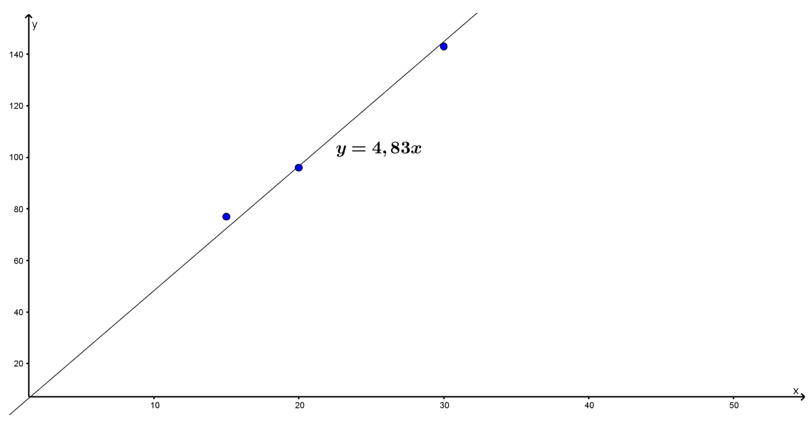

Figura 3.13 – Gráfico da função de ajuste $y = 4,83x$
Fonte: Elaboração própria

Vamos agora introduzir um novo conceito, necessário para a continuidade do estudo do método.

Definição 3.3 *(Derivadas parciais de funções polinomiais reais de grau n). Seja a função $f : \mathbb{R}^2 \to \mathbb{R}$, definida por $f(x, y) = a_{n,n}x^n y^n + a_{n-1,n}x^{n-1}y^n + \cdots + a_{1,n}xy^n + a_{0,n} + a_{n-1,n-1}x^{n-1}y^{n-1} + \cdots + a_{0,0}$, uma função polinomial real de grau n nas variáveis x e y. Definimos a derivada parcial de f com relação à variável x como a função*

$$\frac{\partial f}{\partial x} : \mathbb{R} \to \mathbb{R} = na_n x^{n-1}y^n + (n-1)a_{n-1}x^{n-2}y^n + \cdots + a_1 + (n-1)a_{n-1,n-1}x^{n-1}y^{n-2} + \cdots$$

e a derivada parcial de f em relação à variável y como

$$\frac{\partial f}{\partial y} : \mathbb{R} \to \mathbb{R} = na_n x^n y^{n-1} + (n-1)a_{n-1}x^n y^{n-2} + \cdots + a_1 + (n-1)a_{n-1,n-1}x^n y^{n-2} + \cdots$$

Em suma, quando derivamos f em relação à variável x, consideramos em f cada monômio $a_{i,j}y^i$ constante e, quando derivamos f em relação a y, $a_{i,j}x^i$ constante.

Vejamos agora um exemplo que nos ajudará a ampliar o entendimento do método e a dar uma ideia melhor do seu alcance.

Exemplo 3.3 *(O ajuste a uma função exponencial). Funções exponenciais são particularmente úteis em aplicações por expressarem o desenvolvimento temporal de uma grandeza A cuja variação depende da grandeza. Por exemplo, o crescimento de uma população depende do número de indivíduos na população (neste caso, A é o número de indivíduos na população); a remuneração de uma quantia aplicada a juros depende do valor aplicado (neste caso, A é o valor a ser aplicado); a emissão de partículas por uma material radioativo depende da massa do material (neste caso, A é a massa do material radioativo). Suponhamos, então um processo y que depende exponencialmente de uma variável x da seguinte maneira:*

$$y = ce^{ax}; c, a \in \mathbb{R},$$

sendo a a taxa de variação da massa do material em realçao à quantidade presente desse material radioativo, e c uma constante que, no caso, é a quantidade do material em $x = 0$ (observe que x representa o tempo decorrido a partir de um instante tomado como inicial).

Empregamos, inicialmente, a função logaritmo natural, inversa da função exponencial e, e suas propriedades para obter:

$$ln(y) = ln(c) + ax.$$

Fazendo $Y = ln(y)$ e $ln(c) = C \in \mathbb{R}$, obtemos:

$$Y = ax + C, \tag{3.4}$$

que é a equação de uma reta. Por meio do ***Método dos Mínimos Quadrados*** *iremos trabalhar de modo a determinar os valores dos coeficientes a e C.*

Consideremos os dados do decaimento de um material radioativo. À medida que ele emite partículas, sua massa diminui e a emissão, por sua vez, depende da massa presente do material. Assim, problemas envolvendo produtos radioativos tratam, em geral, da quantidade do material existente após um certo tempo. Suponhamos que obtivemos os dados contidos na tabela 3.9.

$tempo(x)$	0	1	2	3	4	5
$massa(y)$	8200	6540	5000	3200	2500	2000

Tabela 3.9 – Dados referentes ao exemplo 3.1

Calculando os logaritmos naturais das respectivas ordenadas do conjunto de dados fornecido, obtemos a tabela 3.10.

x	0	1	2	3	4	5
$Y = ln(y)$	$9,0$	$8,8$	$8,5$	$8,1$	$7,8$	$7,6$

Tabela 3.10 – Dados referentes ao exemplo 3.1, com $Y = ln(y)$

Para o problema em questão escrevemos, ponto a ponto, a expressão do respectivo erro quadrático:

$$S_1 = [(9,0)^2 - (0a + C)]^2 = (9,0)^2 - 2 \times 9,0 \times C + C^2$$
$$S_2 = [(8,8)^-(1a + C)]^2 = (8,8)^2 - 2 \times 8,8 \times a - 2 \times 8,8 \times C + a^2 + 2aC + C^2$$
$$S_3 = [(8,5)^2 - (2a + C)]^2 = (8,5)^2 - 2 \times 2 \times 8,5 \times a - 2 \times 8,5 \times C + 2^2a^2 + 2 \times 2aC + C^2$$
$$S_4 = [(8,1)^2 - (3a + C)]^2 = (8,1)^2 - 2 \times 3 \times 8,1 \times a - 2 \times 8,1 \times C + 3^2a^2 + 2 \times 3aC + C^2$$
$$S_5 = [(7,8)^2 - (4a + C)]^2 = (7,8)^2 - 2 \times 4 \times 7,8 \times a - 2 \times 7,8 \times C + 4^2a^2 + 2 \times 4aC + C^2$$
$$S_6 = [(7,6)^2 - (5a + C)]^2 = (7,6)^2 - 2 \times 5 \times 7,6 - 2 \times a \times 7,6 \times C + 5^2a^2 + 2 \times 5aC + C^2$$

Como vemos, cada desvio S_i é uma função nas variáveis a e C ($S_i = S_i(a, C)$). Chamando o erro quadrático total de S, escrevemos:

$$S = S_1 + S_2 + S_3 + S_4 + S_5 + S_6, \quad (3.5)$$

ou seja, $S = S(a, C)$. Utilizando a notação de somatório, escrevemos:

$$(9,0)^2 + (8,8)^2 + (8,5)^2 + (8,1)^2 + (7,8)^2 + (7,6)^2 = \sum_{i=1}^{6} Y_i^2$$
$$2(0 \times 9,0 + 1 \times 8,8 + 2 \times 8,5 + 3 \times 8,1 + 4 \times 7,8 + 5 \times 7,6) = 2\sum_{i=1}^{6} x_iY_i$$
$$2(9,0 + 8,8 + 8,8 + 8,1 + 7,8 + 7,6) = 2\sum_{i=1}^{6} Y_i$$
$$0^2 + 1^2 + 2^2 + 3^2 + 4^2 + 5^2 = \sum_{i=1}^{6} x_i^2$$
$$2(0 + 1 + 2 + 3 + 4 + 5) = 2\sum_{i=1}^{6} x_i$$
$$C^2 + C^2 + C^2 + C^2 + C^2 + C^2 = 6 \times C$$

Reescrevendo a equação 3.5, temos:

$$S = \sum_{i=1}^{6} Y_i^2 - 2a\sum_{i=1}^{6} x_iY_i + a^2\sum_{i=1}^{6} x_i^2 - 2C\sum_{i=1}^{6} Y_i + 2aC\sum_{i=1}^{6} x_i + 6C^2 \tag{3.6}$$

Analisemos a equação 3.6 em termos do parâmetro a, considerando para tal o parâmetro C como uma constante. Enxergaremos assim, no lado direito, um polinômio de grau dois em a, cujo coeficiente do termo de maior grau (a^2) é $\sum x_i^2 > 0$, o que lhe garante a existência de um mínimo. Aplicando o mesmo raciocínio em relação ao parâmetro C e considerando o parâmetro a constante, observamos que o coeficiente de C^2 é $6 > 0$, o que nos leva à mesma conclusão quanto à existência de um mínimo para o polinômio em C. Tal raciocínio, baseado em conteúdos ministrados no Ensino Médio, nos permite, juntamente com a relação já vista entre a derivada de uma função quadrática de concavidade voltada para cima e sua derivada, a introdução da ideia de que erro quadrático total será mínimo quando as derivadas da função $S(a, b)$ em relação a a e C forem nulas. Lançamos mão, assim, de forma suave, do conceito de derivadas parciais visto na definição 3.3. Representando-as por $\frac{\partial S}{\partial a}$ e $\frac{\partial S}{\partial C}$, escrevemos então a condição necessária para a ocorrência do mínimo da função S:

$$\frac{\partial S}{\partial a} = \frac{\partial S}{\partial C} = 0$$

.

Aplicando sobre a equação 3.6 a definição 3.3 para derivadas parciais de funções polinomiais de grau n e lembrando que ao derivarmos S em relação a a o parâmetro C é tratado como constante e ao derivarmos S em relação a C o parâmetro a é tratado como constante, obtemos:

$$\frac{\partial S}{\partial a} = -2\sum_{i=1}^{6} x_iY_i + 2a\sum_{i=1}^{6} x_i^2 + 2C\sum_{i=1}^{6} x_i$$

e

$$\frac{\partial S}{\partial C} = -2\sum_{i=1}^{6} Y_i + 2a\sum_{i=1}^{6} x_i + 2 \times 6C$$

Levando agora em consideração a condição para a ocorrência do mínimo da função S, temos:

$$\frac{\partial S}{\partial a} = -2\sum_{i=1}^{6} x_iY_i + 2a\sum_{i=1}^{6} x_i^2 + 2C\sum_{i=1}^{6} x_i = 0 \Rightarrow C\sum_{i=1}^{6} x_i + a\sum_{i=1}^{6} x_i^2 = \sum_{i=1}^{6} x_iY_i$$

e

$$\frac{\partial S}{\partial C} = -2\sum_{i=1}^{6} Y_i + 2a\sum_{i=1}^{6} x_i + 6C = 0 \Rightarrow 6C + a\sum_{i=1}^{6} x_i = \sum_{i=1}^{6} Y.$$

Organizamos e escrevemos na forma matricial para obter:

$$\begin{bmatrix} 6 & \sum_{i=1}^{6} x_i \\ \sum_{i=1}^{6} x_i & \sum_{i=1}^{6} x_i^2 \end{bmatrix} \begin{bmatrix} C \\ a \end{bmatrix} = \begin{bmatrix} \sum_{i=6}^{n} Y_i \\ \sum_{i=1}^{6} x_i Y_i \end{bmatrix}. \tag{3.7}$$

*A equação 3.7 nada mais é que o conjunto de equações normais do **Ajuste Linear Simples** para o caso em que se trabalha com um conjunto de 6 (seis) pontos. Buscando a generalização, considerando n pontos disponíveis para o ajuste, e substituindo $\sum_{i=1}^{n}$ por $\sum$ a título de simplificação, obtemos:*

$$\begin{bmatrix} n & \sum x_i \\ \sum x_i & \sum x_i^2 \end{bmatrix} \begin{bmatrix} C \\ a \end{bmatrix} = \begin{bmatrix} \sum Y_i \\ \sum x_i Y_i \end{bmatrix}. \tag{3.8}$$

*Tal resultado é idêntico à equação 2.13, exceto pelos parâmetros α_0 e α_1, aqui substituídos, respectivamente, por C e a, e pela variável y, aqui substituída por Y. Chegamos, portanto, à forma genérica das equações normais do **Ajuste Linear Simples**.*

Aplicando à equação 3.7 os dados do problema analisado, efetuamos os cálculo dos somatórios nela expressos e obtemos o seguinte sistema:

$$\begin{bmatrix} 6 & 15 \\ 15 & 55 \end{bmatrix} \begin{bmatrix} C \\ a \end{bmatrix} = \begin{bmatrix} 49,8 \\ 119,3 \end{bmatrix},$$

cuja solução é $a \cong -0,3$ e $C \cong 9$. Lembrando que os parâmetros originais do problema são a e c e que $ln(c) = C$, calculamos: $c = e^C \Rightarrow c = e^9 \cong 8103,1$. Portanto, a função de ajuste é $y = (8103,1)e^{-0,3x}$, cujo gráfio está representado na figura 3.14.

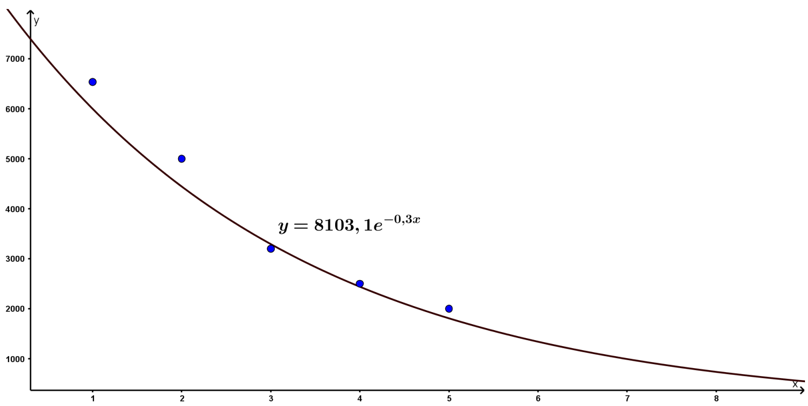

Figura 3.14 – Gráfico da função de ajuste $y = (8103,1)e^{-0,3x}$
Fonte: Elaboração própria

Assim, a partir de exemplos não muito complexos e da introdução de alguns conceitos novos de forma ajustada ao nível de conhecimento do aluno do segmento ao qual o trabalho se destina, espera-se chegar à motivação necessária para se expor a fundamentação matemática que compõe o **Método dos Mínimos Quadrados** e obter-se as equações normais para o **Ajuste Linear Simples**. A partir daí, pode-se discorrer sobre as demais modalidades de ajuste estudadas neste trabalho (**Ajuste Linear Múltiplo** e **Ajuste Polinomial**) e apresentar suas equações normais. No meu entendimento, apresentar as deduções matemáticas para se chegar às equações normais destas duas últimas modalidades de ajuste, devido ao grau de complexidade envolvido, não se faz necessário. Uma vez assimilada a essência e comprovada a utilidade do método, pode-se passar à realização das atividades propostas.

3.2 Atividades

As atividades contidas nesta seção objetivam, além de dotar o aluno do Ensino Médio de desenvoltura na utilização do método e auferir sua eficácia, fazê-lo confrontar-se com aplicações concretas de certos conteúdos por ele já estudados. Uma questão que surge é sobre a vantagem de se levar essa forma de abordagem ao aluno do Ensino Médio, ao invés de fazê-lo esperar pela chegada ao Ensino Superior. O objetivo do trabalho é justamente, tentando respeitar as limitações do aluno do segmento ao qual se destina, enquadrar-se em uma série de ações que visam a redução do impacto que o aluno da área tecnológica, ao ingressar em um curso de graduação, sofre ao confrontar-se com uma matemática com viés mais prático e aplicado.

Todas as atividades propostas envolvem resolução de sistemas de equações lineares. Como alguns dos sistemas que surgem nos exercícios que aqui serão expostos têm resoluções trabalhosas e a meta deste trabalho não é promover esta prática, recomendo a utilização de um software de computador para resolvê-los. Há diversas boas opções gratuitas na internet, todas de fácil utilização, obedecendo o padrão: digitação da ordem do sistema e dos coeficientes e termos independentes antes de um comando para efetuar a resolução. Na maioria dos exercícios deste trabalho utilizei o software **EqlinPlus** baixado em *www.ecivilnet.com/softwares/eqlin_sistemas_lineares_matrizes.htm*. Do mesmo modo, recomenda-se que os diagramas de dispersão e demais gráficos solicitados sejam gerados em um programa específico para isso. Recomendo o **GeoGebra**, que pode ser baixado em *www.geogebra.org/download*.

O profissional que se dispuser a utilizar este trabalho poderá aplicar as atividades aqui sugeridas de acordo com sua avaliação a respeito de obtenção de resultados e desempenho da turma a ser trabalhada, suprimindo itens ou atividades consideradas menos significativas ou muito complexas. O exercícios 4 e 5 ficam como sugestões para um estudo mais aprofundado sobre o tema.

Exercício 1 *(Sobre dilatação linear; uma aplicação no campo da Física). As medições dos comprimentos (y) de uma barra metálica em oito temperaturas (x) diferentes deram origem à tabela 3.11. Para o conjunto de pontos dados:*

a) Trace o diagrama de dispersão.

*b) Determine a curva de ajuste linear pelo **Método dos Mínimos Quadrados**.*

c) Calcule o coeficiente de determinação do ajuste.

d) O valor estimado do comprimento da barra para as temperaturas de 17^oC e 36^oC.

$x(^oC)$	$25,0$	$50,0$	$75,0$	$100,0$	$125,0$	$150,0$	$175,00$	$200,00$
$y(mm)$	$100,07$	$100,12$	$100,16$	$100,21$	$100,26$	$100,30$	$100,35$	$100,40$

Tabela 3.11 – Dados referentes ao exercício 1

Solução:

a) O diagrama de dispersão está representado na figura 3.15. Observe que a disposição dos pontos grafados sugere o ajuste linear.

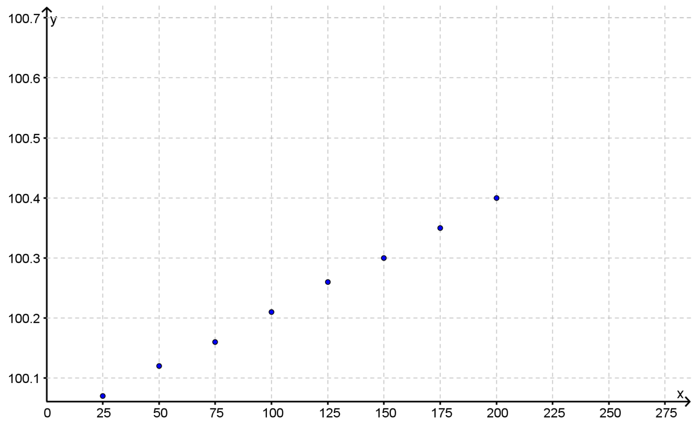

Figura 3.15 – Diagrama de dispersão referente ao exercício 1
Fonte: Elaboração própria

*b) Calculando os somatórios necessários para chegarmos às equações normais referentes ao problema, obtemos a tabela 3.12. Apesar de serem cálculos simples, que podem ser feitos sem maiores dificuldades com o auxílio de uma calculadora eletrônica, tendem a tornar-se um tanto trabalhosos. Algumas fórmulas básicas dispostas em uma planilha do software **Excel** ou similar agilizam bastante a tarefa, facilitando a manipulação e o reaproveitamento das informações.*

i	x_i	y_i	x_i^2	$x_i y_i$	y_i^2
1	$25,0$	$100,07$	$625,00$	$2501,750$	$10014,0049$
2	$50,0$	$100,12$	$2500,00$	$5006,000$	$10024,0144$
3	$75,0$	$100,16$	$5625,00$	$7512,000$	$10032,0256$
4	$100,0$	$100,21$	$10000,00$	$10021,000$	$10042,0441$
5	$125,0$	$100,26$	$15625,00$	$12532,500$	$10052,0676$
6	$150,0$	$100,3$	$22500,00$	$15045,000$	$10060,0900$
7	$175,0$	$100,35$	$30625,00$	$17561,250$	$10070,1225$
8	$200,0$	$100,4$	$40000,00$	$20080,000$	$10080,1600$
$\sum$	$900,0$	$801,9$	$127500,00$	$90259,500$	$80374,529$

Tabela 3.12 – Somatórios referentes ao exercício 1

Sendo linear o ajuste procurado, deverá ter a forma $y = \alpha_0 + \alpha_1 x$. As equações normais do problema são:

$$\begin{bmatrix} 8 & 900,0 \\ 900,0 & 127500,00 \end{bmatrix} \begin{bmatrix} \alpha_0 \\ \alpha_1 \end{bmatrix} = \begin{bmatrix} 801,870 \\ 90259,500 \end{bmatrix}$$

*Aqui apresentaremos a solução obtida através do software **EqlinPlus**, que, como visto, é uma boa opção a ser levada para a sala de aula. A figura 3.16 exibe a tela principal do programa, capturada no momento da obtenção da solução do sistema. Através dela pode-se verificar a simplicidade e a praticidade no uso do software.*

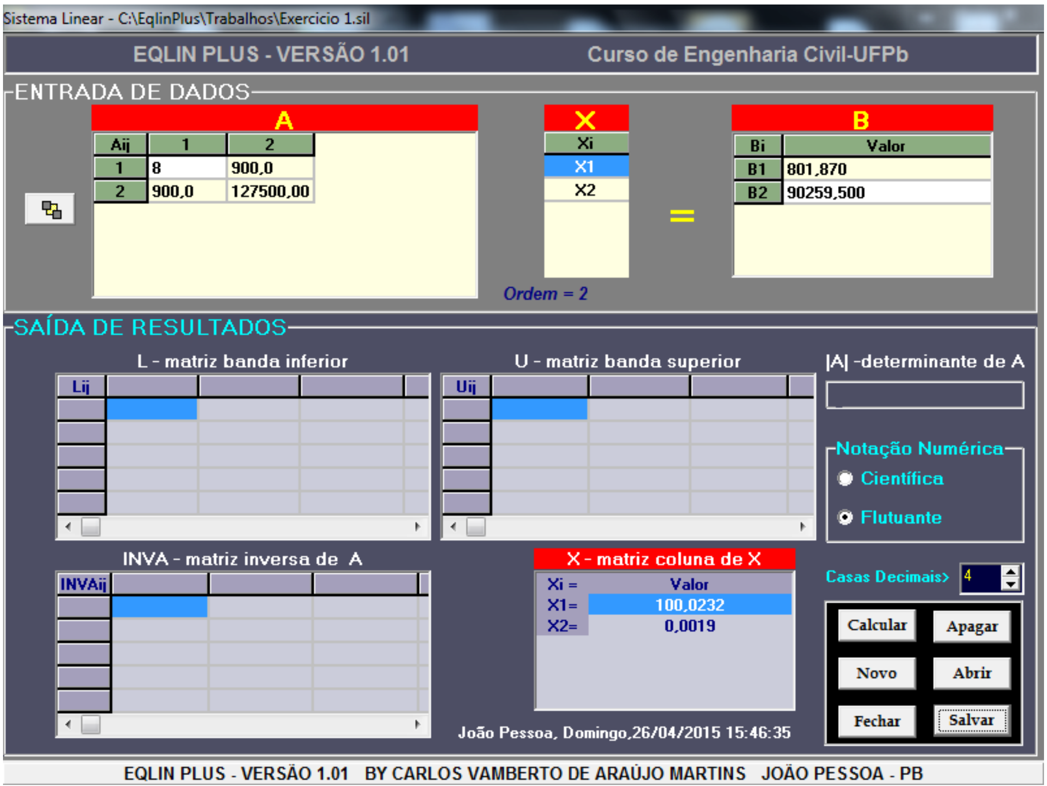

Figura 3.16 – Tela de trabalho do software EqlinPlus
Fonte: Elaboração própria

A solução do sistema, portanto, é dada por $\alpha_0 = 100,0232$ e $\alpha_1 = 0,0019$ e a equação de ajuste é $y = 100,0232 + 0,0019x$.

c) *Para o cálculo do **coeficiente de determinação**, chamando de $\hat{y}$ a variável resposta fornecida pelo ajuste, geramos a tabela 3.13:*

i	y_i	y^2	$\hat{y}_i$	$(y_i - \hat{y}_i)^2$
1	$100,07$	$10014,0049$	$100,0707$	$4,9 \times 10^{-7}$
2	$100,12$	$10024,0144$	$100,1182$	$3,24 \times 10^{-6}$
3	$100,16$	$10032,0256$	$100,1657$	$3,249 \times 10^{-5}$
4	$100,21$	$10042,0441$	$100,2132$	$1,024 \times 10^{-5}$
5	$100,26$	$10052,0676$	$100,2607$	$4,9 \times 10^{-7}$
6	$100,3$	$10060,0900$	$100,3082$	$6,724 \times 10-5$
7	$100,35$	$10070,1225$	$100,3557$	$3,249 \times 10^{-5}$
8	$100,40$	$10080,1600$	$100,4032$	$1,024 \times 10-5$
$\sum$	$801,870$	$80374,529$	$801,8956$	$0,00015692$

Tabela 3.13 – Somatórios referentes ao exercício 1 para cálculo de r^2

$$r^2 = 1 - \frac{\displaystyle\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{y_i^2 - \dfrac{1}{n}\left(\displaystyle\sum_{i=1}^{n} y_i\right)^2} \Rightarrow r^2 = \frac{0,00015692}{80374,529 - \dfrac{1}{8}(801,870)^2} = 0,9982 = 99,82\%.$$

d) *Levando os valores $x_1 = 17^oC$ e $x_2 = 36^oC$ à equação de ajuste, obtemos:*

$y_1 = 100,0232 + 0,0019 \times 17 = 100,05mm$ e $y_2 = 100,0232 + 0,0019 \times 36 = 100,09mm$

Exercício 2 *(Baseado em exemplo contido em Aguiar F. L.; Moreira Junior (s. d.); uma aplicação no campo da Estatística). A tabela 3.14 apresenta, segundo o Instituto Brasileiro de Geografia e Estatística (IBGE), alguns valores da população brasileira (em milhões de habitantes) (y) e seus respectivos anos (x) de referência.*

x	1872	1890	1900	1920	1940	1950	1960	1970	1980	1991	1996
y	$9,9$	$14,3$	$17,4$	$30,6$	$41,2$	$51,90$	$70,2$	$93,1$	$119,0$	$146,2$	$157,1$

Tabela 3.14 – Dados referentes ao exercício 2

a) *Trace o diagrama de dispersão do conjunto de dados.*

b) *Ajuste o conjunto de dados a uma função quadrática.*

c) *Ajuste o conjunto de dados a uma função exponencial*

d) *Estime o valor da população brasileira no anos de 2000, 2005 e 2014 segundo os modelos obtidos nos itens a e b e compare-os com os números oficiais, fornecidos pelo IBGE, que são $169,8$, $184,2$ e $202,7$ milhões, respectivamente.*

Solução:

a) *A figura 3.17 contém o diagrama de dispersão do conjunto. Aparentemente, a disposição dos pontos justifica a busca do ajuste nas formas quadrática e exponencial.*

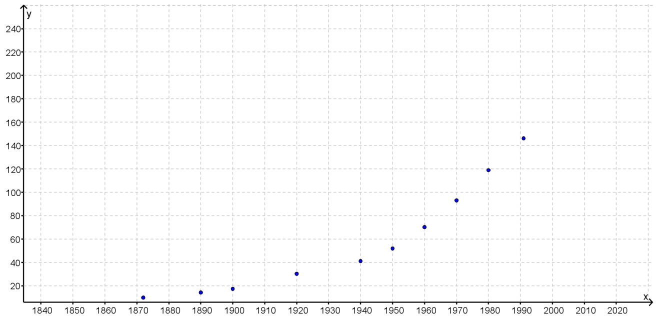

Figura 3.17 – Diagrama de dispersão referente ao exercício 2
Fonte: Elaboração própria

b) *A tabela 3.15 contém os somatórios que irão compor as equações normais do problema.*

i	x_i	y_i	x_i^2	x_i^3	x_i^4	x_iy_i	x_i^2y
1	1872	9,9	3504384	6560206848	12280707219456	18532,8	34693401,6
2	1890	14,3	3572100	6751269000	12759898410000	27027,0	51081030,0
3	1900	17,4	3610000	6859000000	13032100000000	33060,0	62814000,0
4	1920	30,6	3686400	7077888000	13589544960000	58752,0	112803840,0
5	1940	41,2	3763600	7301384000	14164684960000	79928,0	155060320,0
6	1950	51,9	3802500	7414875000	14459006250000	101205,0	197349750,0
7	1960	70,2	3841600	7529536000	14757890560000	137592,0	269680320,0
8	1970	93,1	3880900	7645373000	15061384810000	183407,0	361311790,0
9	1980	119,0	3920400	7762392000	15369536160000	235620,0	466527600,0
10	1991	146,2	3964081	7892485271	15713938174561	291084,2	579548642,2
11	1996	157,1	3984016	7952095936	15872383488256	313571,6	625888913,6
$\sum$	21369	750,9	41529981	80746505055	157061074992273	1479779,6	2916759607,4

Tabela 3.15 – Somatórios referentes ao exercício 2 - ajuste polinomial

A função de ajuste buscada é quadrática, tendo a forma $y = \alpha_0 + \alpha_1 x + \alpha_2 x^2$. As equações normais do problema são:

$$\begin{bmatrix} 11 & 21369 & 41529981 \\ 21369 & 41529981 & 80746505055 \\ 41529981 & 80746505055 & 157061074992273 \end{bmatrix} \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} 750,9 \\ 1479779,6 \\ 2916759607,4 \end{bmatrix}$$

Resolvendo o sistema, obtemos: $\alpha_0 = 46044,8$, $\alpha_1 = -48,7163$ *e* $\alpha_2 = 0,0128889$*. A equação de ajuste é, portanto,* $y = 46044,8 - 48,7163x + 0,0128889x^2$*.*

c) *Vamos tentar ajustar os pontos dados a uma função da forma* $y = \alpha_0 e^{\alpha_1 x}$*. Para isso, vamos linearizar tal função. Aplicando logaritmos aos dois lados da equação, temos:*

$$ln(y) = ln(\alpha_0 e^{\alpha_1 x}) \Rightarrow ln(y) = ln(\alpha_0) + \alpha_1 x\, ln(e) \Rightarrow ln(y) = \beta + \alpha_1 x, \text{ com } \beta = ln(\alpha_0).$$

Tendo linearizado a função inicial, podemos buscar o ***ajuste linear simples*** *através do* ***Método dos Mínimos Quadrados*** *e, finalmente, retroceder à função exponencial. O sistema linear contendo as equações normais do ajuste terá o seguinte aspecto:*

$$\begin{bmatrix} n & \sum x_i \\ \sum x_i & \sum x_i^2 \end{bmatrix} \begin{bmatrix} \beta \\ \alpha_1 \end{bmatrix} = \begin{bmatrix} \sum ln(y_i) \\ \sum x_i\, ln(y_i) \end{bmatrix}$$

A tabela 3.16 contém os dados necessários para compor o sistema.

i	x_i	y_i	x_i^2	$ln(y_i)$	$x_i\, ln(y_i)$
1	1872	9,9	3504384	2,29253	4291,62507
2	1890	14,3	3572100	2,66026	5027,89053
3	1900	17,4	3610000	2,85647	5427,29339
4	1920	30,6	3686400	3,42100	6568,32002
5	1940	41,2	3763600	3,71844	7213,77022
6	1950	51,9	3802500	3,94932	7701,17164
7	1960	70,2	3841600	4,25135	8332,64269
8	1970	93,1	3880900	4,53367	8931,33814
9	1980	119,0	3920400	4,77912	9462,66452
10	1991	146,2	3964081	4,98498	9925,08631
11	1996	157,1	3984016	5,05688	10093,53756
$\sum$	21369	750,9	41529981	42,50403	82975,34008

Tabela 3.16 – Somatórios referentes ao exercício 2- ajuste exponencial

Equações normais do problema:

$$\begin{bmatrix} 11 & 21369 \\ 21369 & 41529981 \end{bmatrix} \begin{bmatrix} \beta \\ \alpha_1 \end{bmatrix} = \begin{bmatrix} 42,50403 \\ 82975,34008 \end{bmatrix}$$

Resolvendo o sistema, obtemos $\beta = -40,42572$ *e* $\alpha_1 = 0,0227988$*. Como* $\beta = ln(\alpha_0) \Rightarrow \alpha_0 = e^{\beta} \Rightarrow \alpha_0 = 2,77545 \times 10^{-18}$*. A função de ajuste, portanto, é* $y = 2,77545 \times 10^{-18} e^{0,0227988x}$

d) *Seja* P_q *a população estimada pelo modelo quadrático e seja* P_e *a população estimada pelo modelo exponencial. Temos, então:*

Ano de 2005: $P_q = 46044,8 - 48,7163 \times 2005 + 0,0129 \times (2005)^2 = 167,8$ *milhões e* $P_e = 2,77545 \times 10^{-18} e^{0,0227988 \times 2005} = 182,32$ *milhões.*

Ano de 2014: $P_q = 46044,8 - 48,7163 \times 2014 + 0,0129 \times (2014)^2 = 210,1$ *milhões e* $P_e = 2,77545 \times 10^{-18} e^{0,0227988 \times 2014} = 242,32$ *milhões.*

Considerando-se a maior proximidade dos valores obtidos com os valores oficiais apresentada pelo modelo quadrático, este mostra-se mais adequado a estimativas futuras.

Exercício 3 *(Uma aplicação no estudo da função quadrática). Determine, pelo **Método dos Mínimos Quadrados**, a equação da parábola que passa pela origem e melhor se ajusta aos pontos* $(-1,3),\ (1,1)$ *e* $(2,5)$ *e trace o gráfico da curva encontrada.*

Solução:

Sendo uma parábola a função procurada, ela tem a forma $y = \alpha_0 + \alpha_1 x + \alpha_2 x^2$. *Para passar pela origem, a curva precisa satisfazer a condição* $\alpha_0 = 0$, *o que reduz a equação inicial a* $y = \alpha_1 x + \alpha_2 x^2$. *Sendo assim, as equações normais do problema terão a seguinte aspecto:*

$$\begin{bmatrix} 3 & \sum x_i & \sum x_i^2 \\ \sum x_i & \sum x_i^2 & \sum x_i^3 \\ \sum x_i^2 & \sum x_i^3 & \sum x_i^4 \end{bmatrix} \begin{bmatrix} 0 \\ \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} \sum y_i \\ \sum x_i y_i \\ \sum x_i^2 y_i \end{bmatrix}$$

Analisando o lado esquerdo da equação e lembrando do produto de matrizes, é fácil perceber que a primeira coluna da matriz dos coeficientes não interfere no produto final. Lembrando também que a primeira linha dessa matriz tem origem na derivação em relação ao parâmetro α, *que neste caso é nulo, concluímos que ela também não afeta o produto final. Logo, podemos escrever as equações normais desse problema da seguinte forma:*

$$\begin{bmatrix} \sum x_i^2 & \sum x_i^3 \\ \sum x_i^3 & \sum x_i^4 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} \sum x_i y_i \\ \sum x_i^2 y_i \end{bmatrix}$$

A tabela 3.17 contém os valores necessários para escrevermos as equações normais do problema.

i	x_i	y_i	x_i^2	x_i^3	x_i^4	$x_i y_i$	$x_i^2 y$
1	-1	3	1	-1	1	-3	3
2	1	1	1	1	1	1	1
3	2	5	4	8	16	10	20
$\sum$	2	9	6	8	18	8	24

Tabela 3.17 – Somatórios referentes ao exercício 3

As equações normais do ajuste são:

$$\begin{bmatrix} 6 & 8 \\ 8 & 18 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} 8 \\ 24 \end{bmatrix}$$

Resolvendo o sistema, obtemos $\alpha_1 = -1,1$ *e* $\alpha_2 = 1,8$. *A função de ajuste é, portanto,* $y = -1,1x + 1,8x^2$. *O gráfico da curva encontrada está representado na figura 3.18.*

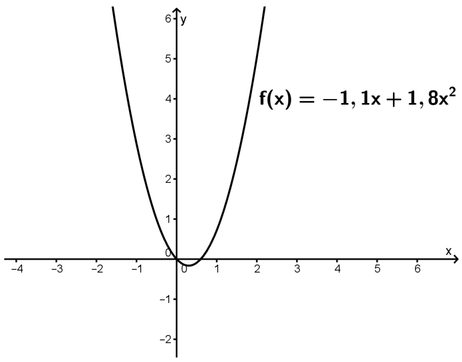

Figura 3.18 – Gráfico da função de ajuste $y = -1,1x + 1,8x^2$
Fonte: Elaboração própria

Exercício 4 *(Uma aplicação no campo da Química). Deseja-se avaliar a utilização de um modelo de regressão linear múltipla na tentativa de explicar a variação da viscosidade* (y) *de um polímero em função da temperatura* (x_1) *de reação e da taxa de adição* (x_2) *do catalisador. Através da realização de experimentos para diferentes valores de* x_1 *e* x_2*, obtiveram-se os valores de* y *mostrados na tabela 3.18. Encontre o ajuste desejado pelo* ***Método dos Mínimos Quadrados*** *e em seguida calcule e utilize o coeficiente de determinação ajustado* ($r^2_{ajust.}$) *para fazer uma avaliação do modelo.*

Nº da observação	**Viscosidade($y, Pa.s$)**	**Temperatura($x_1, ^oC$)**	**Catalisador($x_2, lb/h$)**
1	2256	80	8
2	2340	93	9
3	2426	100	10
4	2293	82	12
5	2330	90	11
6	2368	99	8
7	2250	81	8
8	2409	96	10
9	2364	94	12
10	2379	93	11
11	2440	97	13
12	2364	95	11
13	2404	100	8
14	2317	85	12
15	2309	86	9
16	2328	87	12

Tabela 3.18 – Dados referentes ao exercício 4

*Sendo o **ajuste linear múltiplo** com duas variáveis independentes o modelo proposto, a equação do ajuste é do tipo $y = \beta + \alpha_1 + \alpha_2 x^2$ e as equações normais que permitirão resolver o problema tem a seguinte forma:*

$$\begin{bmatrix} n & \sum x_{1i} & \sum x_{2i} \\ \sum x_{1i} & \sum x_{1i}x_{1i} & \sum x_{1i}x_{2i} \\ \sum x_{2i} & \sum x_{1i}x_{2i} & \sum x_{2i}x_{2i} \end{bmatrix} \begin{bmatrix} \beta \\ \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} \sum y_i \\ \sum x_{1i}y_i \\ \sum x_{2i}y_i \end{bmatrix}$$

A tabela 3.19 contém os valores que irão compor as equações normais do ajuste.

i	x_1	x_2	y	x_1^2	x_2^2	x_1x_2	x_1y	x_2y	
1	80	8	2256	5089536	6400	64	640	180480	18048
2	93	9	2340	5475600	8649	81	837	217620	21060
3	100	10	2426	5885476	10000	100	1000	242600	24260
4	82	12	2293	5257849	6724	144	984	188026	27516
5	90	11	2330	5428900	8100	121	990	209700	25630
6	99	8	2368	5607424	9801	64	792	234432	18944
7	81	8	2250	5062500	6561	64	648	182250	18000
8	96	10	2409	5803281	9216	100	960	231264	24090
9	94	12	2364	5588496	8836	144	1128	222216	28368
10	93	11	2379	5659641	8649	121	1023	221247	26169
11	97	13	2440	5953600	9409	169	1261	236680	31720
12	95	11	2364	5588496	9025	121	1045	224580	26004
13	100	8	2404	5779216	10000	64	800	240400	19232
14	85	12	2317	5368489	7225	144	1020	196945	27804
15	86	9	2309	5331481	7396	81	774	198574	20781
16	87	12	2328	5419584	7569	144	1044	202536	27936
$\sum$	1458	164	37577	88299569	133560	1726	14946	3429550	385562

Tabela 3.19 – Somatórios referentes ao exercício 4

As equações normais do sistema são as seguintes:

$$\begin{bmatrix} 16 & 1458 & 164 \\ 1458 & 133560 & 14946 \\ 164 & 14946 & 1726 \end{bmatrix} \begin{bmatrix} \beta \\ \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} 37577 \\ 3429550 \\ 385562 \end{bmatrix}$$

Os parâmetros procurados são: $\beta = 1566,0777$, $\alpha_1 = 7,6213$ e $\alpha_2 = 8,5848$ e a equação de ajuste é $1566,0777 + 7,6212x_1 + 8,5848x_2$.

Tratando-se de ajuste linear múltiplo, o coeficiente de determinação (r^2) não é o parâmetro mais indicado na avaliação do resultado, devendo ser substituído pelo coeficiente de determinação ajustado ($r^2_{ajust.}$). Pela equação 2.28, vemos que antes do cálculo de $r^2_{ajust.}$ devemos efetuar o cálculo de r^2. Sendo $\hat{y}$ o valor da variável dependente fornecida pelo ajuste, montamos a tabela 3.20 para efetuar tal cálculo.

i	y_i	$\hat{y}_i$	$(y_i - \hat{y}_i)^2$
1	2256	2244,4521	133,3539944
2	2340	2352,1125	146,7126562
3	2426	2414,0457	142,9052885
4	2293	2294,0337	1,06853569
5	2330	2346,4185	269,5671423
6	2368	2389,2549	451,770774
7	2250	2252,0733	4,29857289
8	2409	2383,5609	647,1478088
9	2364	2385,4881	461,7384416
10	2379	2369,2821	94,43758041
11	2440	2416,9365	531,9250323
12	2364	2384,5245	421,2551003
13	2404	2396,8761	50,74995121
14	2317	2316,8973	0,01054729
15	2309	2298,7641	104,7736488
16	2328	2332,1397	17,13711609
$\sum$	37577	37576,86	3478,852191

Tabela 3.20 – Somatórios referentes ao exercício 4 - cálculo de r^2 e $r^2_{ajust.}$

Temos, então:

$$r^2 = 1 - \frac{3478,852191}{88299569 - \dfrac{1}{16} \times (37577)^2} = 0,92697$$

e

$$r^2_{ajust.} = 1 - \left(\frac{16-1}{16-3}\right)(1 - 0,92697)^2 = 0,9157.$$

O próximo exercício tem por objetivo a escolha do melhor ajuste a partir da análise dos indicadores estudados (**coeficiente de determinação** e **coeficiente de determinação ajustado**). Para isso, serão efetuados diversos ajustes com os dados fornecidos, resultando em procedimento bastante trabalhoso. Recomenda-se fortemente, portanto, a utilização de recursos computacionais, principalmente software do tipo planilha eletrônica (recomenda-se o Excel), para cálculo dos somatórios que irão compor as equações normais, e software para resolução de sistemas de equações lineares (recomenda-se o EqlinPlus).

Exercício 5 *(Baseado em exercício contido em Cardeal (s. d.); uma aplicação no campo da Biologia). Os dados contidos na tabela 3.21 relacionam o peso (y) de embriões de frangos desidratados, em gramas, com sua idade (x) em dias.*

i	1	2	3	4	5	6	7	8	9	10	11
$x(dias)$	6	7	8	9	10	11	12	13	14	15	16
$y(g)$	0,029	0,052	0,079	0,125	0,181	0,261	0,425	0,738	1,130	1,882	2,812

Tabela 3.21 – Dados referentes ao exercício 5

a) Trace o diagrama de dispersão do conjunto.

*b) Efetue o **ajuste polinomial** para o conjunto. Para isso, encontre os ajustes até o polinômio de grau cinco e escolha o que melhor representa o conjunto.*

c) Trace o gráfico da curva que representa o melhor ajuste.

c) Estime, com o ajuste realizado, o peso de um embrião de frango no 20º dia de vida.

Solução:

a) O diagrama de dispersão do conjunto está na figura 3.19. Uma vez que o ajuste linear não parece adequado, não efetuaremos o ajuste ao polinômio de grau um.

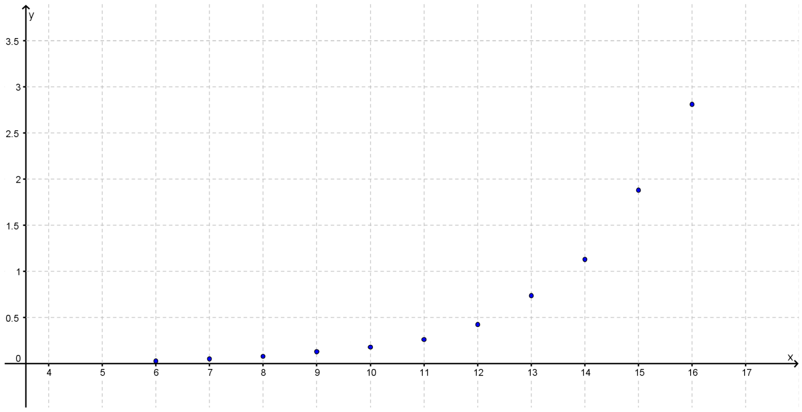

Figura 3.19 – Diagrama de dispersão referente ao exercício 5
Fonte: Elaboração própria

*b) Após realizarmos todos os procedimentos de aplicação do **Método dos Mínimos Quadrados**, obtemos as seguintes funções de ajuste:*

Grau 2: $y = 3,260284 - 0,78462x + 0,04635x^2$
Grau 3: $y = -3,768121 + 1,35263x - 0,157964x^2 + 0,006191x^3$
Grau 4: $y = 0,710235 - 0,476569x + 0,11063x^2 - 0,010668x^3 + 0,000383x^4$
Grau 5: $y = 5,635556 - 2,996244x + 0,609878x^2 - 0,058632x^3 + 0,002622x^4 - 0,000041x^5$

A tabela 3.22 contém os valores de r^2 e $r^2_{ajust.}$ para essas funções de ajuste.

Grau	r^2	$r^2_{ajust.}$
2	$0,969837739$	$0,962297174$
3	$0,998828458$	$0,998326368$
4	$0,999528956$	$0,999214926$
5	$0,976653$	$0,953305147$

Tabela 3.22 – Valores de r^2 e r^2_{ajust} para diferentes funções de ajuste - exercício 5

À medida que se aumenta o grau do polinômio de ajuste, aumenta-se também o número de parâmetros a serem determinados, o que torna recomendável a utilização de $r^2_{ajust.}$ na análise da qualidade do ajuste. Observando-se a tabela 3.22, verifica-se que dentre os ajustes efetuados, o que apresenta maior valor de $r^2_{ajust.}$ é o polinômio de grau quatro, nos levando a concluir que é esse o modelo que melhor explica a variável dependente y, motivo pelo qual recai sobre ele a nossa escolha como função de ajuste.

c) *A função de ajuste* $y = 0,710235 + 0,476569x - 0,11063x^2 - 0,010668x^3 + 0,000383x^4$ *tem a forma mostrada na figura 3.20.*

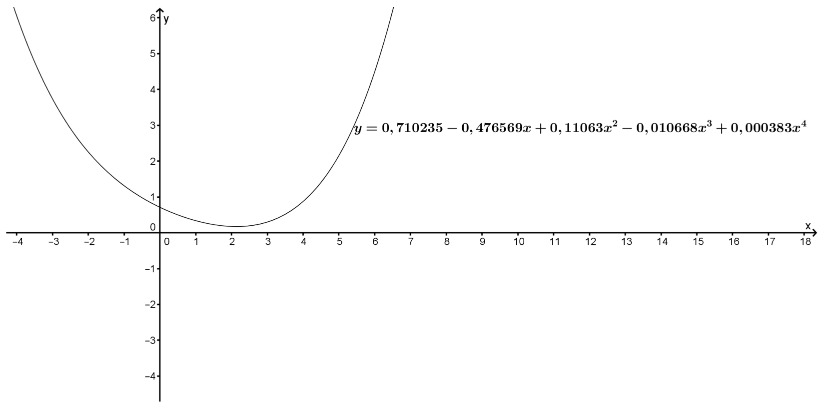

Figura 3.20 – Gráfico da função de ajuste $y = 0,710235 + 0,476569x - 0,11063x^2 - 0,010668x^3 + 0,000383x^4$

Fonte: Elaboração própria

d) *Fazendo* $x = 20$ *na função de ajuste encontrada, temos* $y = 0,710235 + 0,476569 \times 20 - 0,11063 \times 20^2 - 0,010668 \times 20^3 + 0,000383 \times 20^4 = 11,3678g$.

Considerações Finais

Foram expostos neste trabalho os conceitos e fundamentos considerados necessários para orientar quem se dispuser a levar adiante a empreitada de apresentar o **Método dos Mínimos Quadrados** a alunos do Ensino Médio. Buscou-se assim manter-se fiel à proposta inicial, que não priorizava um estudo aprofundado do tema, mas sim prover quem utiliza este trabalho de meios que permitam a abordagem do método para alunos do segmento em questão.

Foi por esse motivo também que conteúdos dos quais o método lança mão em sua implementação, tais como os sistemas de equações lineares e os polinômios, não receberam tratamento diferenciado dentro deste trabalho, embora, por exigência do método, se faça um uso considerável dos mesmos. Como frisado desde o início, um dos objetivos de introduzir-se ao aluno do Ensino Médio o **Método dos Mínimos Quadrados** era oferecer-lhe a possibilidade de utilizar de forma significativa conteúdos já razoavelmente sedimentados, reforçando seu conhecimento e ajudando-o a estimar a importância dos mesmos. Acredito que o trabalho contenha as diretrizes e os meios para a consecução dessa meta, além de proporcionar boas possibilidades de exploração desses conteúdos correlatos.

A utilização de recursos computacionais foi sugerida e incentivada, mas sem orientações rígidas, priorizando-se mais a ênfase à afinidade do método com tais recursos, deixando a cargo do professor explorá-los de acordo com seu projeto de ensino e o perfil de seus alunos.

Por fim, com a introdução do conceito de derivada de uma função polinomial e sua relação com a determinação dos extremos de uma função quadrática, objetivando o entendimento da fundamentação matemática do método sem deixar de evidenciar o processo de otimização que o originou, buscou-se a possibilidade um tratamento ao mesmo tempo acessível e de profundidade suficiente para fazer o aluno perceber a sua dimensão.

Caracterizando dessa forma o presente trabalho, espero ter cumprido a meta de apresentar uma proposta viável de introdução de um novo recurso matemático ao aluno do Ensino Médio, ao qual ele agregará o uso significativo de conteúdos matemáticos por ele já estudados, permitindo-lhe relacionar-se com o universo matemático de forma mais direta e íntima e preparando-o para a aquisição de novas competências e novos saberes.

Referências

AGUIAR F. L.; MOREIRA JUNIOR, W. I. *Ajuste de Curvas por Quadrados Mínimos Lineares*. Belo Horizonte-MG: [s.n.], s. d. UFMG. Disponível em: <http://www.mat.ufmg.br/gaal/aplicacoes/quadrados_minimos.pdf>. Citado na página 57.

BASSANEZI, R. C. *Temas e Modelos*. Santo-André-SP: UFABC, s. d. Citado na página 33.

BOYER, C. B. *História da Matemática*. 2^a. ed. São Paulo: Edgar Blucher, 2001. Citado 2 vezes nas páginas 15 e 17.

BRASIL. *Parâmetros Curriculares Nacionais: Ensino Médio*. Brasília, DF, 2000. Citado na página 12.

BURDEN, R. L.; FAIRES, J. D. *Análise Numérica*. 1^a. ed. São Paulo: Thomson, 2003. Citado na página 37.

CARACcA, B. J. d. *Conceitos Fundamentais da Matemática*. Lisboa-Portugal: Libraria Sá da Costa Editora, 1975. Citado na página 17.

CARDEAL, F. *Ajuste de Curvas*. Belo Horizonte-MG: [s.n.], s. d. CEFETMG. Disponível em: <http://www.decom.cefetmg.br/docentes/flavio_cardeal/Teaching/mnc/aula_ajuste.pdf>. Citado na página 63.

CRATO, N. O papel dos mínimos quadrados na descoberta dos planetas. *Boletim da Sociedade Portuguesa de Matemática*, 2000. Citado na página 18.

EVES, H. *Introdução à História da Matemática*. Campinas-SP: UNICAMP, 2004. Citado na página 17.

GUIMARÃES, P. S. *Ajuste de Curvas Experimentais*. Santa Maria -RS: UFSM, 2001. Citado 2 vezes nas páginas 22 e 23.

OLIVEIRA, A. C. d.; VERTUAN, R. S. *Modelagem Matemática no Ensino Médio*. Londrina-PR, 2009. Citado na página 46.

PEREIRA, M. S. *Cálculo II*. Manaus-AM: Universidade do Estado do Amazonas, 2007. Citado 2 vezes nas páginas 27 e 28.

PITOMBEIRA, J. B.; ROQUE, T. *Tópicos de História da Matemática*. 1^a. ed. Rio de Janeiro-RJ: SBM, 2012. Citado 3 vezes nas páginas 15, 16 e 17.

ROONEY, A. *A História da Matemática*. São Paulo-SP: Makon Books do Brasil Editora, 2012. Citado 3 vezes nas páginas 16, 17 e 18.

RUGGIERO, M.; LOPES, V. *Cálculo Numérico: Aspectos Teóricos e Computacionais*. São Paulo-SP: Makron Books, 1996. Citado na página 35.

SANCHES, J. S.; JORDAN, D. L. *Métodos Numéricos*. Curitiba-PR: [s.n.], 2007. Universidade Federal do Paraná. Disponível em: <http://www.ionildo.cjb.net/metodos/>. Citado na página 21.

SILVA, F. F. d. *O Método dos Mínimos Quadrados: Uma Proposta ao Ensino Médio Para o Ajuste Por Parábolas*. Rio de Janeiro-RJ, 2014. Citado na página 46.

SOUZA, S. A. d. *O Estudo do Produto Matricial Por Meio do Método dos Mínimos Quadrados: Uma Abordagem Destinada ao Ensino Médio*. Santa Maria-RS, 2014. Citado na página 46.

SPERANDIO, D.; MENDES, J. T.; SILVA, L. H. M. *Cálculo Numérico - Características Matemáticas e Computacionais dos Métodos Numéricos*. São Paulo-SP: Prentice-Hall, 2003. Citado na página 21.

UEL. *Matemática Essencial*. Londrina-PR: [s.n.], s. d. Universidade Estadual de Lonfrina. Disponível em: <http://www.uel.br/projetos/matessencial/superior/calculo/maxmin/mm02.htm>. Citado na página 24.