

Universidade Federal de Uberlândia
Faculdade Matemática
Programa de Licenciatura em Matemática à Distância

**Estudo sobre o Método dos Mínimos Quadrados:
Aspectos Teóricos e Aplicações na Física**

Renato José Fernandes
orientando

Alessandro Alves Santana
orientador



Uberlândia-MG
2021

Aos que vêm na Matemática não apenas uma ciência para ajudar a entender o universo, mas sim para buscar entender, mesmo com grandes limitações, a infinita inteligência que emana de Deus.

Agradecimentos

O desenvolvimento de um trabalho de conclusão de curso é o ponto máximo de uma trajetória de 4 anos de empenho e dedicação. Nesse momento é necessário agradecer aos que contribuíram para esse momento. A minha esposa Juliana e minha filha, Maria Emília, pela compreensão e apoio em tantas madrugadas de estudo, provas e atividades no final de semana. Ao professor Alessandro Santana, quem dedicou várias horas, fora de seu horário de trabalho, pela valiosa orientação e motivação. Aos colegas Pedro Ernesto, Rejane Paula, Cíntia Kumamoto que por várias vezes foram motivadores pela continuidade do curso. Aos tutores, Carla, Danilo, Lóren e Helenice pela presteza e suporte nas várias etapas desse curso. Aos professores orientadores de estágio, Sérgio Bispo e Edna Maria, que contribuíram com experiências e vivências de sala de aula. Ao professor Thiago Costa Caetano (UNIFEI - Universidade Federal de Itajubá) pela atenção e valiosas considerações sobre os experimentos do Laboratório Remoto de Física. Aos coordenadores de curso, professores Janser e Juliano, que trabalharam nestes 4 anos auxiliando os alunos, organizando e zelando pelo funcionamento do curso.

Resumo

O presente trabalho consiste no estudo do chamado *Método dos Mínimos Quadrados* (MMQ), uma técnica de aproximação de funções muito utilizada em experimentos de laboratório para estimar valores funcionais de propriedades físicas, químicas e/ou biológicas cuja lei associada, de forma contínua, não é conhecida, tendo-se apenas conhecimento de valores discretos obtidos através de experimentos laboratoriais. Esse estudo aborda duas formas de dedução do MMQ, sendo que uma utiliza bastante embasamento teórico da álgebra linear e outro através do cálculo diferencial e integral, além de questões referentes à sensibilidade computacional, aos erros de arredondamento inerentes às operações aritméticas do computador, dos métodos numéricos empregados na resolução de sistemas lineares, os quais são necessários no processo de obtenção da solução de um problema de mínimos quadrados. Ressalta-se ainda que o trabalho ilustra questões práticas de ordem computacional através da apresentação de resultados gerados através de códigos desenvolvidos com o GNU Octave. Salienta-se um fato de grande importância nesse estudo que foi a utilização de dados, nas simulações numéricas, obtidos remotamente do LABREMOTO (Laboratório Remoto de Física) da UNIFEI - Universidade Federal de Itajubá - um laboratório de física que permite a realização de experimentos à distância, gerando dados, os quais foram empregados nesse estudo.

Palavras-chave: Método dos Mínimos Quadrados, Equações Normais, Transformações de Householder.

Abstract

The present work consists in the study of the so-called *Least Squares Method* (MMQ), a technique of function approximation widely used in laboratory experiments to estimate values functional properties of physical, chemical and/or biological organisms whose associated law, in a continuous way, is not known, having only knowledge of discrete values obtained through experiments laboratory tests. this study addresses two ways of deducting the MMQ, where one uses a lot of theoretical foundation from linear algebra and other through calculus differential and integral, in addition to questions regarding computational sensitivity, the rounding errors inherent in the computer arithmetic operations, of the numerical methods used in the resolution of linear systems, which are necessary in the process of obtaining a solution to a problem. least squares. It is noteworthy that the work illustrates issues computational practices through presentation of results generated through codes developed with the GNU Octave. stand out a fact of great importance in this study that was the use of data, in numerical simulations, obtained remotely from LABREMOTO (Remote Physics Laboratory) of UNIFEI - Federal University of Itajubá - a laboratory of physics that allows experiments to be carried out remotely, generating data, which were employed in this study.

Keywords: Least Square Method, Normal Equations, Householder Transformation.

Sumário

1	Introdução	1
2	Método dos Mínimos Quadrados	3
2.1	Via projeção ortogonal	3
2.1.1	Aspectos importantes sobre sistemas lineares	4
2.1.2	Sobre as soluções de sistemas lineares	6
2.1.3	Resolução numérica das Equações Normais	10
2.1.4	Resolução numérica via Transformações de Householder	10
2.2	Via Cálculo Diferencial e Integral	15
2.3	Aspectos computacionais sobre o MMQ	18
3	Aplicações do MMQ na Física	20
3.1	Aspectos computacionais dos experimentos	20
3.2	Movimento uniforme	20
3.3	Determinação da velocidade do som no ar	23
3.4	Curva de luz	24
3.5	Decaimento radioativo	28
4	Conclusões	30
A	Códigos Implementados em Octave	32
A.1	Ajuste de Curvas via Equações Normais	32
A.2	Ajuste de Curvas via Transformação de Householder	34
	Referências	37

Lista de Figuras

2.1 Ilustração da projeção ortogonal do vetor b no espaço coluna da matriz A associada ao sistema linear $Ax = b$	7
2.2 Ajuste de dados via polinômios de graus 1, 2 e 3.	19
3.1 Gráfico de dispersão dos dados: Experimento do trilho de ar para obtenção da velocidade.	21
3.2 Trilho de ar: Ajuste via polinômio de grau 1.	22
3.3 Gráfico de dispersão dos dados: Experimento do trilho de ar para obtenção da aceleração.	22
3.4 Trilho de ar: Ajuste via polinômio de grau 2.	23
3.5 Gráfico de dispersão dos dados: Experimento da velocidade do som no ar.	24
3.6 Gráfico de dispersão dos dados: Experimento da curva de luz.	25
3.7 Gráfico de dispersão dos dados: Ajuste via polinômios de grau 3.	26
3.8 Gráfico de dispersão dos dados: Ajuste via polinômios de grau 4.	26
3.9 Gráfico de dispersão dos dados: Ajuste via polinômios de grau 5.	26
3.10 Gráfico de dispersão dos dados: Ajuste via polinômios de grau 8.	27
3.11 Gráfico de dispersão dos dados: Ajuste via polinômios de grau 9.	27
3.12 Gráfico de dispersão dos dados: Ajuste via polinômios de grau 10.	27
3.13 Gráfico de dispersão dos dados: Decaimento radioativo.	29

Lista de Tabelas

2.1 Tabela com valores discretos de uma função com lei de formação desconhecida.	3
2.2 Ilustração de uma tabela de valores de uma função com lei de formação desconhecida.	15

CAPÍTULO 1

Introdução

As ciências da natureza se desenvolveram através da observação do meio e com busca por leis fundamentais que pudessem descrever com boa exatidão os fenômenos naturais. Essa exatidão, na maioria dos casos práticos não é possível, sendo necessário a aplicação de estratégias de aproximação para realizar estudos nesses campos. A matemática é a ferramenta mais ampla do conhecimento, pois possibilita realizar cálculos, previsões e resolver problemas através de técnicas analíticas e numéricas. O ajuste de curvas experimentais, a qual é uma técnica de aproximação funcional de propriedades físicas, químicas e/ou biológicas, possibilita a produção de resultados sistematizados e com valor amplo no estudo de situações reais em que seja necessário comprovar ou produzir uma relação de dependência dos dados.

Ocorre que em várias situações há possibilidades significativas de construir conhecimentos através da resolução de problemas da natureza através da modelagem matemática mediante o tratamento de dados experimentais.

Dentre as diversas formas de obter aproximações para propriedades físicas, por exemplo, existe o chamado *Método dos Mínimos Quadrados*, o qual representaremos ao longo do texto pelo acrônimo MMQ. Essa técnica na literatura é também conhecida pela terminologia *Ajuste de Curvas*. Essa técnica é extensivamente utilizada em laboratórios de várias áreas e produz boas aproximações para as mais variadas propriedades em estudo a partir de dados obtidos experimentalmente ou não, como os que ocorrem em livros didáticos de cálculo numérico cujo objetivo é a explanação do tema no processo de ensino e aprendizagem.

O Método dos Mínimos Quadrados pode ser utilizado em dados discretos ou contínuos [5, 9]. No presente texto a abordagem desse tema teve por foco o caso discreto, sendo um dos objetivos desse trabalho a apresentação do processo de dedução teórico do MMQ por duas abordagens distintas mas que conduzem ao mesmo fim do ponto de vista teórico, mas que na prática tem detalhamentos que precisam ser levados em consideração. Tal método normalmente é abordado nas universidades em livros de Cálculo Numérico abordando a dedução da técnica por apenas um dos caminhos possíveis, evitando no processo, ao que parece, pela maior demanda teórica exigida para o entendimento do MMQ. A apresentação do MMQ mais presente nos livros didáticos de Cálculo Numérico é dedução do método usando o Cálculo Diferencial e Integral, mediante a minimização de uma função escalar de várias variáveis, que são parâmetros utilizados no ajuste de curvas, empregando derivadas parciais do processo de obtenção do método. Um outro processo de dedução, que normalmente é mais visto em livros de Álgebra Linear, exige mais fundamentos dessa área para deduzir o método, mas que é de grande relevância para o conhecimento do assunto como um todo. Assim sendo, nesse trabalho foi realizado o estudo e apresentação do processo por essas duas vias, bem como o

detalhamento de aspectos muito importantes do ponto de vista computacional uma vez que é o computador que vai realizar na prática as operações aritméticas para obter os parâmetros de ajustes sendo que dependendo do problema pode ocorrer problemas de sensibilidade computacional a tais operações devido aos erros de arredondamento do computador.

A teoria do Método dos Mínimos Quadrados é apresentada no capítulo 2 sendo abordada em detalhes pelas duas vias supracitadas, levando-se também em considerações pontos relevantes nos aspectos computacionais para evitar problemas oriundos das operações aritméticas inerentes ao processamento computacional.

O capítulo 3 aborda experimentos realizados com programas desenvolvidos com o GNU Octave, um software livre com foco em Matemática Computacional, usando dados experimentais obtidos através do LABREMOTO, um laboratório de física remoto da Universidade Federal de Itajubá.

O capítulo 4 apresenta as considerações finais acerca do estudo desenvolvimento do ponto de vista teórico e do ponto de vista prático, bem como propostas futuras de implementação do que foi realizado.

CAPÍTULO 2

Método dos Mínimos Quadrados

O *Método dos Mínimos Quadrados*, o qual será representado ao longo do texto pelo acrônimo MMQ, é uma técnica de aproximação de funções. Existe na literatura basicamente duas formas de abordar esse método, sendo uma baseada em projeção ortogonal, que utiliza muita álgebra linear em seu processo de dedução, e a outra que tem por base o cálculo diferencial e integral. Ambos os processos supracitados acabam por conduzir à necessidade de resolver um sistema linear e, no que tange aos métodos numéricos para resolver sistemas lineares, é necessário também levar em consideração a sensibilidade computacional quanto aos erros de arredondamento dos computadores. Em essência, a forma como se resolve um problema de aproximação funcional via MMQ deve levar em consideração essas características sendo que existem métodos bem eficientes, com baixa sensibilidade computacional, para resolver os sistemas lineares associados a um problema de mínimos quadrados. Esse capítulo aborda os aspectos teóricos e computacionais e da dedução do MMQ usando cada uma dessas formas, sendo as fontes utilizadas para esse estudo os livros dos autores Anton, Franco e Leon [1, 5–7].

2.1 Via projeção ortogonal

Seja f uma função, com lei de formação desconhecida, com m valores tabelados

x_k	x_1	x_2	x_3	...	x_{m-1}	x_m
$f(x_k)$	$f(x_1)$	$f(x_2)$	$f(x_3)$...	$f(x_{m-1})$	$f(x_m)$

Tabela 2.1: Tabela com valores discretos de uma função com lei de formação desconhecida.

e suponhamos uma função aproximadora para essa função dada pela expressão (2.1)

$$f_{\text{ap}}(x) = \alpha_1 \phi_1(x) + \alpha_2 \phi_2(x) + \alpha_3 \phi_3(x) + \dots + \alpha_n \phi_n(x) \quad (2.1)$$

com $m \geq n$, sendo ϕ_i , $i = 1, 2, \dots, n$, funções previamente escolhidas de acordo o gráfico de dispersão da função tabelada.

Analisando a função aproximadora f_{ap} , nota-se que o que a determina são os parâmetros α_i , $i = 1, 2, \dots, n$. Para obter esses parâmetros, vamos primeiramente considerar a possibilidade de que os mesmos possam ser determinados através da montagem e resolução do sistema linear (2.2) considerando principalmente que tal sistema é sobredeterminado, isto é,

exemplo, uma base do espaço vetorial $V = \mathbb{R}^3$, segue que dado um vetor v qualquer desse espaço, existem escalares a_1, a_2 e a_3 tais que

$$v = a_1 v_1 + a_2 v_2 + a_3 v_3 \quad (2.6)$$

sendo que tais escalares são únicos para esse vetor v e, nesse caso, diz-se que as coordenadas do vetor v em relação à base \mathcal{A} é dado e denotado por

$$[v]_{\mathcal{A}} = (a_1, a_2, a_3).$$

Seguindo esse princípio, concluímos que resolver um sistema linear $Ax = b$, com m equações e n incógnitas,

$$\underbrace{\begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{bmatrix}}_A \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix}}_x = \underbrace{\begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_m \end{bmatrix}}_b$$

no fundo significa determinar as coordenadas do vetor b (termos independentes) em relação à base do subespaço vetorial gerado pelas colunas da matriz dos coeficientes A . Em outras palavras, resolver o sistema linear $Ax = b$ acima significa determinar $x_1, x_2, x_3, \dots, x_n$ de tal forma que

$$(b_1, b_2, \dots, b_m) = x_1(a_{11}, a_{21}, \dots, a_{m1}) + x_2(a_{12}, a_{22}, \dots, a_{m2}) + \cdots + x_n(a_{1n}, a_{2n}, \dots, a_{mn})$$

sendo que $[b]_{\mathcal{A}} = (x_1, x_2, \dots, x_n)$, onde \mathcal{A} é o conjunto de vetores constituído pelas colunas da matriz A , isto é,

$$\mathcal{A} = \{(a_{11}, a_{21}, \dots, a_{m1}), (a_{12}, a_{22}, \dots, a_{m2}), \dots, (a_{1n}, a_{2n}, \dots, a_{mn})\}.$$

Como exemplo, para determinar as coordenadas do vetor $v = (4, -2, 8)$ em relação à base

$$\mathcal{A} = \{(2, 1, -4), (-4, 5, 2), (5, -3, 5)\},$$

é necessário determinar x_1, x_2 e x_3 de tal modo que

$$(4, -2, 8) = x_1(2, 1, -4) + x_2(-4, 5, 2) + x_3(5, -3, 5)$$

$$(2x_1 - 4x_2 + 5x_3, x_1 + 5x_2 - 3x_3, -4x_1 + 2x_2 + 5x_3) = (4, -2, 8)$$

sendo que para obter esses parâmetros segue que é necessário resolver o sistema linear

$$\begin{bmatrix} 2 & -4 & 5 \\ 1 & 5 & -3 \\ -4 & 2 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 4 \\ -2 \\ 8 \end{bmatrix} \xrightarrow{\text{solução}} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -5/18 \\ 7/18 \\ 11/9 \end{bmatrix}$$

e portanto

$$[(4, -2, 8)]_{\mathcal{A}} = \left(-\frac{5}{18}, \frac{7}{18}, \frac{11}{9}\right)$$

com a observação importante de que

$$(4, -2, 8) = -\frac{5}{18}(2, 1, -4) + \frac{7}{18}(-4, 5, 2) + \frac{11}{9}(5, -3, 5)$$

sendo $\mathcal{A} = \{(2, 1, -4), (-4, 5, 2), (5, -3, 5)\}$, um conjunto cujos vetores são as colunas da matriz A e constituem, nesse caso, uma base do \mathbb{R}^3 , sendo que a solução do sistema linear são as coordenadas do vetor b em relação aos vetores dessa base.

Resta agora a discussão com relação ao número de equações m e incógnitas n presentes em um sistema linear $Ax = b$ e o que se pode obter com a solução do mesmo.

2.1.2 Sobre as soluções de sistemas lineares

Em um sistema linear $Ax = b$ onde o número de equações é igual ao número de incógnitas, isto é, $m = n$, e as colunas da matriz dos coeficientes são linearmente independentes, o que implica $\det(A) \neq 0$, a solução existe e é única pois nessa situação o sistema é possível e determinado. Usando terminologias da Álgebra Linear, isso significa que o vetor b pertence ao subespaço gerado pelas colunas da matriz A . Se o número de equações m é menor que o número de incógnitas n pode acontecer do sistema linear ser possível e indeterminado (infinitas soluções) ou impossível (não tem solução).

Agora, se o sistema linear for sobredeterminado, que são sistemas onde o número de equações m é maior que o número de incógnitas n ($m > n$), a solução exata pode não ser obtida mas sim aproximada. Esse tema é o assunto da próxima subsecção.

Resolução de sistemas lineares sobredeterminados

Um sistema linear $Ax = b$ do tipo

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_m \end{bmatrix}$$

onde $m > n$, isto é, um sistema linear sobredeterminado, o qual de um modo geral, não é possível obter a solução exata, mas sim uma solução aproximada em determinado sentido. A ideia por trás disso do processo de obtenção dessa solução aproximada reside em minimizar o quadrado da norma do *vetor resíduo*, onde esse vetor e sua norma ao quadrado são dados por

$$\begin{cases} r = b - Ax \\ \|r\|^2 = r_1^2 + r_2^2 + r_3^2 + \cdots + r_m^2 \end{cases} \quad (2.7)$$

o que ilustra que o sentido do princípio da referida aproximação consiste em obter a solução aproximada no *sentido dos mínimos quadrados*, princípio este que dá nome ao método de aproximação em apresentação nesse texto.

Note que se x é solução exata do sistema linear, segue que $\|r\|^2 = 0$. É importante aqui ressaltar que o vetor r não é vetor dos erros na solução do sistema linear mas sim um parâmetro que tem relação com o vetor dos erros, mas é um parâmetro que ajuda a medir a qualidade da aproximação.

Para entender o processo de obtenção da solução aproximada do sistema linear sobredeterminado $Ax = b$ é necessário apresentar alguns teoremas.

Teorema 2.1

Seja W um subespaço vetorial do \mathbb{R}^m . Então, para todo vetor $v \in \mathbb{R}^m$, o vetor em W mais próximo de v é a projeção ortogonal deste vetor no subespaço W .

Uma maneira de interpretar o teorema 2.1 é que a norma $\|v - v_p\|$, onde o vetor $v \in \mathbb{R}^m$ e v_p é a projeção de v no subespaço W , é minimizada.

Esse teorema é muito importante no processo de resolução do sistema linear sobredeterminado $Ax = b$. Para entender essa importância, considere os seguintes apontamentos:

- As colunas da matriz dos coeficientes de um sistema linear $Ax = b$ geram um subespaço vetorial do \mathbb{R}^m , chamado *espaço coluna* da matriz A . Como já observado anteriormente, se o sistema linear é possível e determinado ($m = n$) segue que o vetor b pertence ao espaço coluna da matriz A e isso significa que b pode ser escrito exatamente como uma combinação linear das colunas da matriz A . Em outras palavras, as colunas da matriz A formam uma base do \mathbb{R}^m e como $b \in \mathbb{R}^m$ segue que b pode ser escrito como uma combinação linear de uma única forma como combinação linear das colunas da matriz dos coeficientes A . E como b nesse caso pertence ao espaço coluna da matriz A , a projeção ortogonal b_p deste vetor nesse espaço é tal que $\|b - b_p\| = 0$;
- Quando $Ax = b$ é sobredeterminado, isto é, $m > n$, segue que, de um modo geral embora haja exceções, b não pertence ao espaço coluna da matriz A . Mas uma solução aproximada é possível resolvendo o sistema linear $Ax = b_p$ onde b_p é a projeção ortogonal do vetor b no espaço coluna gerado pela matriz A . Essa projeção b_p é o vetor mais próximo do vetor b que pertence ao espaço coluna da matriz A conforme ilustrado na figura 2.2.

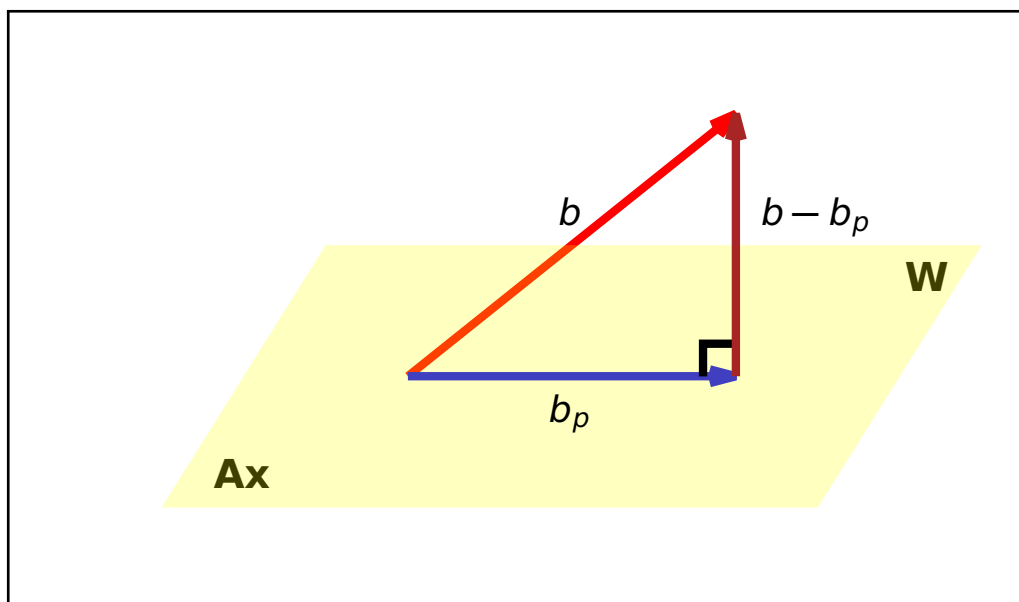


Figura 2.1: Ilustração da projeção ortogonal do vetor b no espaço coluna da matriz A associada ao sistema linear $Ax = b$.

Assim sendo, dentro do espaço coluna da matriz A precisamos encontrar \hat{x} tal que $A\hat{x} = b_p$. A solução deste sistema irá fornecer a solução que minimiza o quadrado da norma do resíduo $\|r\|^2 = \|b - A\hat{x}\|^2$. Observando a figura, pode-se notar que o vetor $b - b_p$ é ortogonal a qualquer vetor do espaço coluna da matriz A . Em particular, o vetor $b - b_p$ é ortogonal aos vetores da base do espaço coluna gerado pela matriz A . De posse dessas informações, podemos então

afirmar que

$$A^T(b - b_p) = 0 \xrightarrow{A\hat{x}=b_p} A^T(b - A\hat{x}) = 0 \Rightarrow A^T b - A^T A \hat{x} = 0 \Rightarrow A^T A \hat{x} = A^T b. \quad (2.8)$$

Note na equação (2.8) que para obter \hat{x} que minimiza o quadrado da norma do resíduo $\|b - A\hat{x}\|^2$ basta então resolver um sistema linear onde o número de equações e incógnitas é o mesmo, no caso, dado por $n \times n$.

O sistema linear $A^T A x = A^T b$ recebe o nome de *sistema normal* associado ao sistema linear $Ax = b$. As equações do sistema linear $A^T A x = A^T b$ também são chamados *equações normais*. O teorema a seguir estabelece as condições para existência e unicidade da solução do sistema normal.

Teorema 2.2

Se A é uma matriz $m \times n$ onde o posto de A é igual a n , então $A^T A$ é inversível e o sistema linear $Ax = b$ tem uma única aproximação por mínimos quadrados dada por $\hat{x} = (A^T A)^{-1} A^T b$, isto é, o sistema normal tem uma única solução.

Para exemplificar, considere o sistema linear sobredeterminado $Ax = b$ onde

$$A = \begin{bmatrix} 1 & 2 & -1 & 3 \\ 2 & 1 & 1 & 2 \\ -2 & 3 & 4 & 1 \\ 4 & 2 & 1 & 0 \\ 0 & 2 & 1 & 3 \\ 1 & -1 & 2 & 0 \end{bmatrix} \quad \text{e} \quad b = \begin{bmatrix} 1 \\ 5 \\ -2 \\ 1 \\ 3 \\ 5 \end{bmatrix} \quad (2.9)$$

com $m = 6$ equações e $n = 4$ incógnitas. Determinando $A^T A$ e $A^T b$ montamos o sistema normal

$$\underbrace{\begin{bmatrix} 26 & 5 & -1 & 5 \\ 5 & 23 & 13 & 17 \\ -1 & 13 & 24 & 6 \\ 5 & 17 & 6 & 23 \end{bmatrix}}_{A^T A} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}}_x = \underbrace{\begin{bmatrix} 24 \\ 4 \\ 10 \\ 20 \end{bmatrix}}_{A^T b} \xrightarrow{\text{solução}} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0.9990 \\ -2.0643 \\ 1.1039 \\ 1.8902 \end{bmatrix}. \quad (2.10)$$

Com a solução (2.10) podemos determinar a projeção ortogonal via $b_p = Ax$,

$$b_p = \underbrace{\begin{bmatrix} 1 & 2 & -1 & 3 \\ 2 & 1 & 1 & 2 \\ -2 & 3 & 4 & 1 \\ 4 & 2 & 1 & 0 \\ 0 & 2 & 1 & 3 \\ 1 & -1 & 2 & 0 \end{bmatrix}}_A \underbrace{\begin{bmatrix} 0.9990 \\ -2.0643 \\ 1.1039 \\ 1.8902 \end{bmatrix}}_x \Rightarrow b_p = \begin{bmatrix} 1.4371 \\ 4.8181 \\ -1.8852 \\ 0.9713 \\ 2.6459 \\ 5.2712 \end{bmatrix}. \quad (2.11)$$

sendo o vetor resíduo dado por

$$r = b - Ax = \begin{bmatrix} 1 \\ 5 \\ -2 \\ 1 \\ 3 \\ 5 \end{bmatrix} - \begin{bmatrix} 1 & 2 & -1 & 3 \\ 2 & 1 & 1 & 2 \\ -2 & 3 & 4 & 1 \\ 4 & 2 & 1 & 0 \\ 0 & 2 & 1 & 3 \\ 1 & -1 & 2 & 0 \end{bmatrix} \begin{bmatrix} 0.9990 \\ -2.0643 \\ 1.1039 \\ 1.8902 \end{bmatrix} \Rightarrow r = \begin{bmatrix} -0.4371 \\ 0.1819 \\ -0.1148 \\ 0.0287 \\ 0.3541 \\ -0.2712 \end{bmatrix} \quad (2.12)$$

O quadrado da norma do vetor resíduo é dado por

$$\|r\|^2 = \|b - Ax\|^2 = 0.437083. \quad (2.13)$$

É importante ressaltar que determinados sistemas lineares sobredeterminados $Ax = b$ podem gerar equações normais onde a matriz dos coeficientes $A^T A$ não é inversível. Um exemplo disso é o sistema linear $Ax = b$ sobredeterminado onde

$$A = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix} \quad \text{e} \quad b = \begin{bmatrix} -3 \\ -1 \\ 0 \\ 2 \\ 5 \\ 1 \end{bmatrix}.$$

Nesse caso temos que a matriz dos coeficientes do sistema normal é dada por

$$A^T A = \begin{bmatrix} 6 & 2 & 2 & 2 \\ 2 & 2 & 0 & 0 \\ 2 & 0 & 2 & 0 \\ 2 & 0 & 0 & 2 \end{bmatrix}$$

onde pode-se observar que essa matriz não é inversível pois a primeira linha é combinação linear das demais linhas (a primeira linha é igual a soma das demais linhas), ou seja, o determinante dessa matriz é nulo ($\det(A^T A) = 0$). Pode-se observar também que a primeira coluna é combinação linear das demais colunas (a primeira coluna é igual a soma das demais colunas). Além disso, a resolução do sistema linear $A^T Ax = A^T b$ conduz a seguinte solução

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 3 \\ -5 \\ -2 \\ 0 \end{bmatrix} + \lambda \begin{bmatrix} -1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \quad \text{onde} \quad \lambda \in \mathbb{R}.$$

ou seja, o sistema sobredeterminado tem infinitas soluções. Diante de situações como essa, o teorema a seguir estabelece condições sobre a matriz $A^T A$ para que a solução do problema de mínimos quadrados via equações normais exista e seja única.

Teorema 2.3

A matriz $A^T A$ é inversível se e somente se as colunas da matriz A são linearmente independentes. Nesse caso, o sistema linear $Ax = b$ tem uma única solução em termos de mínimos quadrados \hat{x} dado por

$$\hat{x} = (A^T A)^{-1} A^T b.$$

O que foi exemplificado acima ilustra um problema quando se está resolvendo um problema de mínimos quadrados. Se as colunas da matriz A do sistema linear $Ax = b$ formarem um conjunto de vetores quase linearmente dependentes a matriz $A^T A$ do sistema normal tende a ser singular ($\det(A^T A)$ tende a zero) e a matriz $A^T A$ tende a ser mal condicionada, isto é, sensíveis aos erros de arredondamento inerentes ao processo de resolução numérica via meios computacional, o que normalmente ocorre em situações práticas (reais). Devido a

isso os softwares utilizados para resolver sistemas desse tipo normalmente não utilizam as equações normais para resolver tais sistemas.

Os problemas de condicionamento podem ser contornados resolvendo o sistema linear $Ax = b$ através da transformação deste em um sistema linear da forma

$$Rx = c$$

utilizando transformações de Householder. As duas próximas subseções abordam o processo de resolução numérica do sistema linear sobredeterminado $Ax = b$.

2.1.3 Resolução numérica das Equações Normais

A resolução do problema de mínimos quadrados via equações normais só é viável se $A^T A$ for inversível e, mesmo assim, se o número de condição da matriz $A^T A$, dado por

$$\text{cond}(A^T A) = \|(A^T A)^{-1}\| \|A^T A\|,$$

for menor que 10^4 pois se o número de condição da matriz em questão for maior que esse valor a mesma é considerada mal condicionada e, portanto, consideravelmente sensível aos erros de arredondamento nas operações aritméticas realizadas pelo computador. Supondo que a matriz dos coeficientes é bem condicionada, a obtenção da solução do problema de mínimos quadrados é viável. O sistema normal $A^T Ax = A^T b$ pode ser resolvido utilizando o *Método de Eliminação de Gauss com Pivoteamento Parcial (MEGPP)*. Esse método numérico tem por base transformar, através da aplicação das chamadas operações elementares em uma ordem bem específica, um sistema linear $Ax = b$ em um sistema linear equivalente $\tilde{A}x = \tilde{b}$ onde \tilde{A} é uma matriz triangular superior conforme esquematizado na matriz ampliada do sistema abaixo e executada através do algoritmo do MEGPP.

$$\left[\begin{array}{cccc|c} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} & b_2 \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} & b_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} & b_n \end{array} \right] \sim \left[\begin{array}{cccc|c} \tilde{a}_{11} & \tilde{a}_{12} & \tilde{a}_{13} & \cdots & \tilde{a}_{1n} & \tilde{b}_1 \\ 0 & \tilde{a}_{22} & \tilde{a}_{23} & \cdots & \tilde{a}_{2n} & \tilde{b}_2 \\ 0 & 0 & \tilde{a}_{33} & \cdots & \tilde{a}_{3n} & \tilde{b}_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \tilde{a}_{nn} & \tilde{b}_n \end{array} \right]$$

Para resolver esse sistema linear simplificado onde a matriz dos coeficientes é uma matriz triangular superior, isto é, obter a solução do sistema, basta aplicar o chamado *Algoritmo da Substituição Regressiva*.

Mas como já comentado anteriormente, dependendo do problema de mínimos quadrados, é necessário aplicar métodos mais sofisticados, que exige um número maior bem de operações se comparado ao MEGPPP, porém mais precisos. Um dos métodos mais robustos nessa linha utiliza as Transformações de Householder, tema da próxima subsecção.

2.1.4 Resolução numérica via Transformações de Householder

A resolução de sistemas lineares sobredeterminados, associados a problemas de mínimos quadrados, via transformações de Householder tem por base a multiplicação sucessiva à esquerda de A e b do sistema linear $Ax = b$ por matrizes de Householder até transformar a matriz A em uma matriz triangular superior conforme esquematizado na matriz ampliada do

sistema linear $Ax = b$ abaixo.

$$\underbrace{\begin{bmatrix} X & X & X & \dots & X & | & X \\ X & X & X & \dots & X & | & X \\ X & X & X & \dots & X & | & X \\ X & X & X & \dots & X & | & X \\ X & X & X & \dots & X & | & X \\ X & X & X & \dots & X & | & X \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ X & X & X & \dots & X & | & X \end{bmatrix}}_{Ax=b} \sim \underbrace{\begin{bmatrix} \tilde{X} & \tilde{X} & \tilde{X} & \dots & \tilde{X} & | & \tilde{X} \\ 0 & \tilde{X} & \tilde{X} & \dots & \tilde{X} & | & \tilde{X} \\ 0 & 0 & \tilde{X} & \dots & \tilde{X} & | & \tilde{X} \\ 0 & 0 & 0 & \dots & \tilde{X} & | & \tilde{X} \\ 0 & 0 & 0 & \dots & \tilde{X} & | & \tilde{X} \\ 0 & 0 & 0 & \dots & 0 & | & \tilde{X} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & | & \tilde{X} \end{bmatrix}}_{\tilde{A}x=\tilde{b}}$$

Transformações de Householder são operações baseadas em transformações ortogonais, as quais são bem estáveis para operações computacionais.

As matrizes de Householder H são obtidas conforme os parâmetros apresentados abaixo.

$$\begin{cases} \alpha = \|x\|_2 \\ \beta = \alpha(\alpha - x_1) \\ v = (x_1 - \alpha, x_2, x_3, \dots, x_m)^T \\ H = I - \frac{1}{\beta}vv^T \\ y = Hx \end{cases}$$

Nesse esquema, x é um vetor de dimensão m o qual se tem interesse em anular todos as componentes desse vetor a partir da segunda posição. O exemplo a seguir irá ilustrar sua aplicação em um vetor x e em seguida será explicado como se aplica essas transformações em um sistema linear $Ax = b$ sobredeterminado.

Se $x = (-5, -1, -3, 9)^T$ o qual se tem interesse em anular os números que se encontram da segunda posição em diante, aplicando a Transformação de Householder descrita nos passos a seguir.

Passo 1: Cálculo de α :

$$\alpha = \|x\|_2 = \sqrt{x_1^2 + x_2^2 + x_3^2 + x_4^2} = \sqrt{(-5)^2 + (-1)^2 + (-3)^2 + (9)^2} = 10.7703.$$

Passo 2: Cálculo de β :

$$\beta = \alpha(\alpha - x_1) = 10.7703(10.7703 - (-5)) = 169.8509.$$

Passo 3: Vetor v :

$$v = (x_1 - \alpha, x_2, x_3, x_4) = (-15.7703, -1, -3, 9)^T.$$

Passo 4: Matriz H :

$$H = I - \frac{1}{\beta}vv^T = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} - \frac{1}{169.8509} \begin{bmatrix} -15.7703 \\ -1 \\ -3 \\ 9 \end{bmatrix} \begin{bmatrix} -15.7703 & -1 & -3 & 9 \end{bmatrix}$$

$$H = \begin{bmatrix} -0.464238 & -0.092848 & -0.278543 & 0.835629 \\ -0.092848 & 0.994113 & -0.017662 & 0.052987 \\ -0.278543 & -0.017662 & 0.947013 & 0.158962 \\ 0.835629 & 0.052987 & 0.158962 & 0.523113 \end{bmatrix}$$

Passo 5: Multiplicação Hx :

$$y = Hx = \begin{bmatrix} -0.464238 & -0.092848 & -0.278543 & 0.835629 \\ -0.092848 & 0.994113 & -0.017662 & 0.052987 \\ -0.278543 & -0.017662 & 0.947013 & 0.158962 \\ 0.835629 & 0.052987 & 0.158962 & 0.523113 \end{bmatrix} \begin{bmatrix} -5 \\ -1 \\ -3 \\ 9 \end{bmatrix}$$

$$y = \begin{bmatrix} 10.7703 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

O que foi apresentado acima se refere a um vetor. E é importante ressaltar a carga de operações exigidas pelo métodos para zerar as componentes do vetor x da segunda posição em diante. É uma característica desse método mas ganha-se em estabilidade computacional, isto é, evita-se amplificação dos erros inerentes às operações aritméticas computacionais.

A aplicação dessas transformações para resolver um sistema linear $Ax = b$, sendo A com ordem $m \times n$, sobredeterminado consiste em multiplicar à esquerda da matriz ampliada desse sistema pelas matrizes

$$H_k = H \text{ se } k = 1 \quad \text{e} \quad H_k = \begin{bmatrix} I & 0 \\ \dots & \dots \\ 0 & H \end{bmatrix} \text{ se } k = 2, 3, \dots, m.$$

com a observação de que todas essas matrizes são quadradas e tem ordem igual a m . A matriz I que aparece acima, para $k \geq 2$, são matrizes identidades de ordem $k - 1$ e H tem ordem $m - 1$. As matrizes H são sempre formadas utilizando vetores colunas extraídos da matriz A . Por exemplo, quando $k = 1$ o vetor x extraído para formar a matriz H é a primeira coluna da matriz A . Depois de multiplicar à esquerda a matriz ampliada do sistema por H_1 obtemos um outro sistema linear equivalente onde todos os números abaixo do elemento na primeira linha e primeira coluna são iguais a zero. Em seguida, para formar H_2 o vetor que teremos que utilizar para formar a matriz H é a segunda coluna da matriz ampliada a contar da segunda linha. Esse processo deve ser repetido sucessivamente até transformar a matriz A em uma matriz triangular superior. Para um maior entendimento, o exemplo a seguir ilustrará o processo.

Exemplo 2.1

Resolva o sistema linear a seguir em termos de mínimos quadrados

$$\begin{bmatrix} 5 & 3 & 6 \\ -5 & -8 & 7 \\ 7 & 3 & 6 \\ 7 & 3 & -7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \\ -4 \\ 3 \end{bmatrix}.$$

Resolução: A matriz dos coeficientes é uma matriz 4×3 , isto é, tem $m = 4$ linhas (equações) e $n = 3$ colunas (incógnitas). As matrizes de Householder serão então matrizes 4×4 . Teremos que multiplicar à esquerda a matriz ampliada associada a esse sistema linear $n = 3$ vezes pois a matriz dos coeficientes tem $n = 3$ colunas. A matriz ampliada do sistema linear, a qual chamaremos por M , é dada por

$$\left[\begin{array}{ccc|c} 5 & 3 & 6 & 1 \\ -5 & -8 & 7 & -2 \\ 7 & 3 & 6 & -4 \\ 7 & 3 & -7 & 3 \end{array} \right]$$

e a seguir as operações de triangularização via Transformações de Householder. Ressalta-se o observância quanto à alta carga de operações que exigidas nas operações que são apresentadas na resolução desse exemplo. Perde-se em volume de operações computacionais mas ganha-se em qualidade na obtenção da aproximação da solução do problema de mínimos quadrados.

1ª coluna: Vamos montar a matriz de Householder para zerar na primeira coluna todos os números abaixo do primeiro elemento dessa coluna. Ressalta-se aqui que esse primeiro elemento será alterado mas não irá zerar. Nesse caso, o vetor que iremos trabalhar é o vetor $x = (5, -5, 7, 7)^T$.

$$x = (5, -5, 7, 7)^T$$

$$\alpha = \|x\|_2 = 12.166$$

$$\beta = \alpha(\alpha - x_1) = 87.172$$

$$v = (-7.1655, -5, 7, 7)^T$$

$$\overbrace{\begin{bmatrix} 0.4110 & -0.4110 & 0.5754 & 0.5754 \\ -0.4110 & 0.7132 & 0.4015 & 0.4015 \\ 0.5754 & 0.4015 & 0.4379 & -0.5621 \\ 0.5754 & 0.4015 & -0.5621 & 0.4379 \end{bmatrix}}^H \overbrace{\begin{bmatrix} 5 & 3 & 6 & 1 \\ -5 & -8 & 7 & -2 \\ 7 & 3 & 6 & -4 \\ 7 & 3 & -7 & 3 \end{bmatrix}}^M = \begin{bmatrix} 12.166 & 7.9734 & -0.98639 & 0.6576 \\ 0 & -4.5297 & 2.1250 & -2.2389 \\ 0 & -1.8585 & 12.825 & -3.6655 \\ 0 & -1.8585 & -0.17499 & 3.3345 \end{bmatrix}$$

2ª coluna: Agora o vetor x a ser utilizado para montar a matriz de Householder diminui de dimensão.

$$x = (-4.5297, -1.8585, -1.8585)^T$$

$$\alpha = \|x\|_2 = 5.2370$$

$$\beta = \alpha(\alpha - x_1) = 5.2370$$

$$v = (-9.7666, -1.8585, -1.8585)^T$$

$$\overbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -0.8649 & -0.3549 & -0.3549 \\ 0 & -0.3549 & 0.9325 & -0.0675 \\ 0 & -0.3549 & -0.0675 & 0.9325 \end{bmatrix}}^H \overbrace{\begin{bmatrix} 12.166 & 7.9734 & -0.98639 & 0.6576 \\ 0 & -4.5297 & 2.1250 & -2.2389 \\ 0 & -1.8585 & 12.825 & -3.6655 \\ 0 & -1.8585 & -0.17499 & 3.3345 \end{bmatrix}}^M =$$

$$\left[\begin{array}{ccc|c} 12.1655 & 7.9734 & -0.9864 & 0.6576 \\ 0 & 5.2370 & -6.3272 & 2.0540 \\ 0 & 0 & 11.2167 & -2.8486 \\ 0 & 0 & -1.7833 & 4.1514 \end{array} \right]$$

3ª coluna: Novamente o vetor x a ser utilizado para montar a matriz de Householder diminui de dimensão.

$$x = (11.2167, -1.7833)^T$$

$$\beta = \alpha(\alpha - x_1) = 1.6001$$

$$\alpha = \|x\|_2 = 11.358$$

$$v = (-0.1409, -1.7833)^T$$

$$\overbrace{\left[\begin{array}{cccc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0.9876 & -0.1570 \\ 0 & 0 & -0.1570 & -0.9876 \end{array} \right]}^H \overbrace{\left[\begin{array}{cccc} 12.1655 & 7.9734 & -0.9864 & 0.6576 \\ 0 & 5.2370 & -6.3272 & 2.0540 \\ 0 & 0 & 11.2167 & -2.8486 \\ 0 & 0 & -1.7833 & 4.1514 \end{array} \right]}^M =$$

$$\left[\begin{array}{ccc|c} 12.166 & 7.9734 & -0.98639 & 0.65760 \\ 0 & 5.2370 & -6.3272 & 2.0540 \\ 0 & 0 & 11.358 & -3.4651 \\ 0 & 0 & 0 & -3.6526 \end{array} \right]$$

Agora, para completar o processo de resolução para obter a solução no sentido dos mínimos quadrados basta extrair a última matriz ampliada tudo o que estiver acima da 4 linha.

$$\left[\begin{array}{ccc|c} 12.166 & 7.9734 & -0.98639 & 0.65760 \\ 0 & 5.2370 & -6.3272 & 2.0540 \\ 0 & 0 & 11.358 & -3.4651 \end{array} \right]$$

A solução desse último sistema linear fornecerá a solução em termos de mínimos quadrados do sistema linear original. Essa solução pode ser obtida aplicando o algoritmo da substituição regressiva, que no caso é dado por

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0.013845 \\ 0.023606 \\ -0.305094 \end{bmatrix}.$$

O cálculo do quadrado da norma do resíduo pode ser feito via $\|b - Ax\|^2$ usando a solução obtida. Mas no caso da solução obtida mediante a aplicação de transformações de Householder o cálculo do quadrado da norma do resíduo fica bem simplificado, quase de modo direto, conforme explicação a seguir. As transformações de Householder constituem a base da chamada *Fatoração QR*. A matriz Q é o resultado da multiplicação das matrizes de Householder e R é a matriz triangular superior como a que foi obtida no exemplo acima. Esquemmatizando, depois de transformar o sistema linear sobredeterminado $Ax = b$ em um sistema linear equivalente $Rx = c$ através da aplicação de transformações de Householder, o resíduo desse último sistema pode ser escrito como

$$r = c - Rx = \begin{bmatrix} c_1 \\ \dots \\ c_2 \end{bmatrix} - \begin{bmatrix} R \\ \dots \\ 0 \end{bmatrix} \underbrace{\begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix}}_x = \begin{bmatrix} c_1 - Rx \\ \dots \\ c_2 - 0x \end{bmatrix} = \begin{bmatrix} 0 \\ \dots \\ c_2 \end{bmatrix} \Rightarrow \|r\|^2 = \|c_2\|^2.$$

Esse vetor c_2 é um vetor extraído do vetor c a partir da posição $n + 1$ desse vetor. Para ilustrar, considerando o problema de mínimos quadrados apresentado no exemplo anterior, o vetor c_2 tem na verdade apenas um elemento e nesse o quadrado da norma do resíduo é dado por $\|c_2\|^2 = (-3.6526)^2 = 13.3415$.

Com essa parte finaliza-se a apresentação dos aspectos teóricos sobre o processo de resolução de problemas de mínimos quadrados associados à resolução de sistemas lineares sobredeterminados $Ax = b$. Nesse processo foi apresentada toda uma metodologia baseada na Álgebra Linear, tais como base de espaço e subespaço, coordenadas em relação a uma determinada base formada pelas colunas da matriz dos coeficientes, projeção ortogonal de vetores, produtos internos e normas. Mas existe, além das deduções apresentadas, uma outra baseada em Cálculo Diferencial e Integral, a qual será apresentada na próxima seção.

2.2 Via Cálculo Diferencial e Integral

Nessa seção é apresentada na dedução da técnica de ajuste de curvas via Cálculo Diferencial e Integral, o qual conduz à resolução do problema de mínimos quadrados por meio das Equações Normais. O desenvolvimento inicial, com a finalidade de ambientação no tema, ilustra o processo do ajuste, de uma função tabelada, por meio de uma função afim, isto é, uma função de 1º grau. Após o desenvolvimento desse caso é apresentada a dedução do caso genérico envolvendo essa técnica de ajuste.

Considere então o problema de ajustar a função, com lei de formação desconhecida, apresentada na Tabela 2.2, a qual contém n valores da referida função.

x_k	x_1	x_2	x_3	...	x_{n-1}	x_n
$f_{\text{tab}}(x_k)$	$f_{\text{tab}}(x_1)$	$f_{\text{tab}}(x_2)$	$f_{\text{tab}}(x_3)$...	$f_{\text{tab}}(x_{n-1})$	$f_{\text{tab}}(x_n)$

Tabela 2.2: Ilustração de uma tabela de valores de uma função com lei de formação desconhecida.

Para ajustar a função tabelada acima, indicada por f_{tab} , por uma reta (função de grau um), a qual indicaremos por f_{aj} , considere a função de ajuste

$$f_{\text{aj}}(x) = a_1x + a_0. \tag{2.14}$$

Para obter a função de ajuste f_{aj} , função aproximadora da função tabelada f_{tab} , no sentido dos mínimos quadrados, é necessário determinar a_0 e a_1 de tal forma que o somatório das diferenças ao quadrado entre a função tabelada f_{tab} e a função de ajuste f_{aj} seja o menor possível, isto é, é necessário determinar a_0 e a_1 de tal forma que

$$F(a_0, a_1) = \sum_{k=1}^n [f_{\text{tab}}(x_k) - f_{\text{aj}}(x_k)]^2 \tag{2.15}$$

onde a expressão é chamada *resíduo*. A substituição da Equação (2.14) na Equação (2.15) produz

$$F(a_0, a_1) = \sum_{k=1}^n [f_{\text{tab}}(x_k) - a_1x_k - a_0]^2 \tag{2.16}$$

que expressa de modo claro uma função de duas variáveis. Para determinar a_0 e a_1 que minimiza F é necessário obter as coordenadas dos pontos críticos desta última função, os quais

são obtidos resolvendo um sistema linear, nesse caso, envolvendo duas equações: F_{a_0} e F_{a_1} . Essas derivadas parciais, a montagem do sistema linear e suas manipulações/simplificações são apresentadas nas eqs. (2.17a) e (2.17b) e eqs. (2.18) a (2.24).

Derivadas parciais:

$$\frac{\partial F}{\partial a_0} = -2 \sum_{k=1}^n [f_{\text{tab}}(x_k) - a_1 x_k - a_0] \tag{2.17a}$$

$$\frac{\partial F}{\partial a_1} = -2 \sum_{k=1}^n [f_{\text{tab}}(x_k) - a_1 x_k - a_0] x_k \tag{2.17b}$$

Montagem do sistema linear: A substituição das eqs. (2.17a) e (2.17b) no sistema linear envolvendo as equações $F_{a_0} = 0$ e $F_{a_1} = 0$ produz as operações abaixo descritas, cujo sistema linear final é o sistema linear cuja solução são os parâmetros a_0 e a_1 que minimizam o resíduo.

$$\begin{cases} \frac{\partial F}{\partial a_0} = 0 \\ \frac{\partial F}{\partial a_1} = 0 \end{cases} \tag{2.18}$$

$$\begin{cases} -2 \sum_{k=1}^n [f_{\text{tab}}(x_k) - a_1 x_k - a_0] = 0 \\ -2 \sum_{k=1}^n [f_{\text{tab}}(x_k) - a_1 x_k - a_0] x_k = 0 \end{cases} \tag{2.19}$$

$$\begin{cases} \sum_{k=1}^n [f_{\text{tab}}(x_k) - a_1 x_k - a_0] = 0 \\ \sum_{k=1}^n [f_{\text{tab}}(x_k) - a_1 x_k - a_0] x_k = 0 \end{cases} \tag{2.20}$$

$$\begin{cases} \sum_{k=1}^n f_{\text{tab}}(x_k) - \left(\sum_{k=1}^n x_k\right) a_1 - \left(\sum_{k=1}^n 1\right) a_0 = 0 \\ \sum_{k=1}^n x_k f_{\text{tab}}(x_k) - \left(\sum_{k=1}^n x_k^2\right) a_1 - \left(\sum_{k=1}^n x_k\right) a_0 = 0 \end{cases} \tag{2.21}$$

$$\begin{cases} \left(\sum_{k=1}^n 1\right) a_0 + \left(\sum_{k=1}^n x_k\right) a_1 = \sum_{k=1}^n f_{\text{tab}}(x_k) \\ \left(\sum_{k=1}^n x_k\right) a_0 + \left(\sum_{k=1}^n x_k^2\right) a_1 = \sum_{k=1}^n x_k f_{\text{tab}}(x_k) \end{cases} \tag{2.22}$$

$$\begin{bmatrix} \sum_{k=1}^n 1 & \sum_{k=1}^n x_k \\ \sum_{k=1}^n x_k & \sum_{k=1}^n x_k^2 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} = \begin{bmatrix} \sum_{k=1}^n f_{\text{tab}}(x_k) \\ \sum_{k=1}^n x_k f_{\text{tab}}(x_k) \end{bmatrix} \tag{2.23}$$

$$\begin{bmatrix} n & \sum_{k=1}^n x_k \\ \sum_{k=1}^n x_k & \sum_{k=1}^n x_k^2 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} = \begin{bmatrix} \sum_{k=1}^n f_{\text{tab}}(x_k) \\ \sum_{k=1}^n x_k f_{\text{tab}}(x_k) \end{bmatrix} \tag{2.24}$$

O modo mais simples para resolver o sistema linear na eq. (2.24) é via Regra de Cramer onde

$$\alpha_0 = \frac{\left(\sum_{k=1}^n x_k^2\right)\left(\sum_{k=1}^n f_{\text{tab}}(x_k)\right) - \left(\sum_{k=1}^n x_k\right)\left(\sum_{k=1}^n x_k f_{\text{tab}}(x_k)\right)}{n \sum_{k=1}^n x_k^2 - \left(\sum_{k=1}^n x_k\right)^2} \quad (2.25)$$

e

$$\alpha_1 = \frac{n \left(\sum_{k=1}^n x_k f_{\text{tab}}(x_k)\right) - \left(\sum_{k=1}^n x_k\right)\left(\sum_{k=1}^n f_{\text{tab}}(x_k)\right)}{n \sum_{k=1}^n x_k^2 - \left(\sum_{k=1}^n x_k\right)^2} \quad (2.26)$$

As equações (2.24) são as **equações normais** e podem ser resolvidas por vários métodos diretos, no caso 2 por 2 a resolução por regra de Cramer é rápida, porém para sistemas com várias variáveis o método mais comum é via Método de Eliminação de Gauss com Pivoteamento Parcial (MEGPP).

Generalizado o método para aplicação para uma função tabelada qualquer queremos obter uma função f_{aj} que produz a melhor aproximação dos dados tabelados x_i e f_{tab}

$$f_{aj}(x) = \alpha_1 \phi_1(x) + \alpha_2 \phi_2(x) + \alpha_2 \phi_2(x) + \alpha_3 \phi_3(x) + \dots + \alpha_m \phi_m(x) \quad (2.27)$$

com $\phi_i(x)$, $i = 1, 2, 3, \dots, m$ funções previamente escolhidas de modo a minimizar a função

$$F(\alpha_1, \alpha_2, \dots, \alpha_m) = \sum_{k=1}^n [f_{\text{tab}}(x_k) - f_{aj}(x_k)]^2. \quad (2.28)$$

Após escolhida as funções $\phi_i(x)$, $i = 1, \dots, m$ o próximo passo é determinar os valores de α_i que minimizam a função F acima. Para tanto é necessário derivar F em relação a cada parâmetro α_i e igualar cada derivada parcial a zero para determinar as coordenadas do ponto crítico da função, candidato a mínimo da função F . Esse procedimento é apresentado nas equações a seguir.

$$F(\alpha_1, \alpha_2, \dots, \alpha_m) = \sum_{k=1}^n [f_{\text{tab}}(x_k) - f_{\text{tab}}(x_k)]^2 \quad (2.29)$$

$$F(\alpha_1, \alpha_2, \dots, \alpha_m) = \sum_{k=1}^n [f_{\text{tab}}(x_k) - \alpha_1 \phi_1(x_k) - \alpha_2 \phi_2(x_k) - \alpha_2 \phi_2(x_k) - \dots - \alpha_m \phi_m(x_k)]^2 \quad (2.30)$$

A derivada parcial em relação a cada parâmetro α_i é dada a seguir.

$$\frac{\partial F}{\partial \alpha_i} = -2 \sum_{k=1}^n [f_{\text{tab}}(x_k) - \alpha_1 \phi_1(x_k) - \alpha_2 \phi_2(x_k) - \dots - \alpha_m \phi_m(x_k)] \phi_i(x_k) \quad (2.31)$$

Abrindo o somatório,

$$\frac{\partial F}{\partial \alpha_i} = -2 \sum_{k=1}^n f(x_k) \phi_i(x_k) - \left[\sum_{k=1}^n \phi_1(x_k) \phi_i(x_k) \right] \alpha_1 - \left[\sum_{k=1}^n \phi_2(x_k) \phi_i(x_k) \right] \alpha_2 - \dots - \left[\sum_{k=1}^n \phi_m(x_k) \phi_i(x_k) \right] \alpha_m \quad (2.32)$$

Agora, para determinar os parâmetros α_i , $i = 1, \dots, m$, que minimizam F , basta resolver um sistema linear montado a partir dessas derivadas parciais acima igualando-as a zero para determinar o ponto crítico da função F . Igualando essas derivadas parciais a zero, e depois deixando no primeiro membro apenas as parcelas que envolvem as incógnitas,

$$\left[\sum_{k=1}^n \phi_1(x_k) \phi_i(x_k) \right] \alpha_1 + \left[\sum_{k=1}^n \phi_2(x_k) \phi_i(x_k) \right] \alpha_2 + \dots + \left[\sum_{k=1}^n \phi_m(x_k) \phi_i(x_k) \right] \alpha_m = \sum_{k=1}^n f_{\text{tab}}(x_k) \phi_i(x_k)$$

Para simplificar as expressões que serão apresentadas a seguir, considere

$$a_{ij} = \sum_{k=1}^n \phi_i(x_k) \phi_j(x_k) \quad \text{e} \quad b_i = \sum_{k=1}^n f_{\text{tab}}(x_k) \phi_i(x_k)$$

onde $i = 1, 2, \dots, m$ e $j = 1, 2, \dots, m$. Fazendo i variar de 1 até m será produzido o seguinte sistema linear.

$$\begin{cases} a_{11}\alpha_1 + a_{12}\alpha_2 + a_{13}\alpha_3 + \dots + a_{1m}\alpha_m & = & b_1 \\ a_{21}\alpha_1 + a_{22}\alpha_2 + a_{23}\alpha_3 + \dots + a_{2m}\alpha_m & = & b_2 \\ a_{31}\alpha_1 + a_{32}\alpha_2 + a_{33}\alpha_3 + \dots + a_{3m}\alpha_m & = & b_3 \\ & \vdots & \\ a_{m1}\alpha_1 + a_{m2}\alpha_2 + a_{m3}\alpha_3 + \dots + a_{mm}\alpha_m & = & b_m \end{cases} \quad (2.33)$$

Esse sistema linear representado em notação matricial é dado por

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1m} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2m} \\ a_{31} & a_{32} & a_{33} & \dots & a_{3m} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \dots & a_{mm} \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \vdots \\ \alpha_m \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_m \end{bmatrix}. \quad (2.34)$$

Como a matriz dos coeficientes é simétrica, $a_{ij} = a_{ji}$. A resolução da mesma pode ser feita via Método de Eliminação de Gauss com Pivoteamento Parcial. Vale ressaltar aqui o sistema linear obtido é o mesmo sistema normal obtido no início desse capítulo, na parte abordagem matricial.

2.3 Aspectos computacionais sobre o MMQ

O método computacional mais eficiente e estável numericamente para resolver o problema de mínimos quadrados é resolvendo o sistema sobredeterminado usando operações baseadas em transformações ortogonais, isto é, transformações de Householder [6]. Uso das equações normais podem conduzir à formação de sistema lineares mal condicionados, isto é, sistemas lineares sensíveis às operações aritméticas computacionais inerentes as operações do computador, gerando soluções sem nenhuma ligação com o problema em análise. Mais adiante, no capítulo de testes computacionais com problemas reais, isso será ilustrado.

Ilustração Gráfica

Como ilustração considere o problema de ajustar a função tabelada

x_k	0.54	0.83	1.36	1.90	2.11	2.68	3.88	4.28
$f(x_k)$	2.9855	2.6488	2.0654	2.1553	2.3854	3.0487	1.4207	0.4432

usando polinômios de graus 1, 2 e 3. A figura a seguir mostra o que acontece à medida que se aumenta o grau do polinômio.

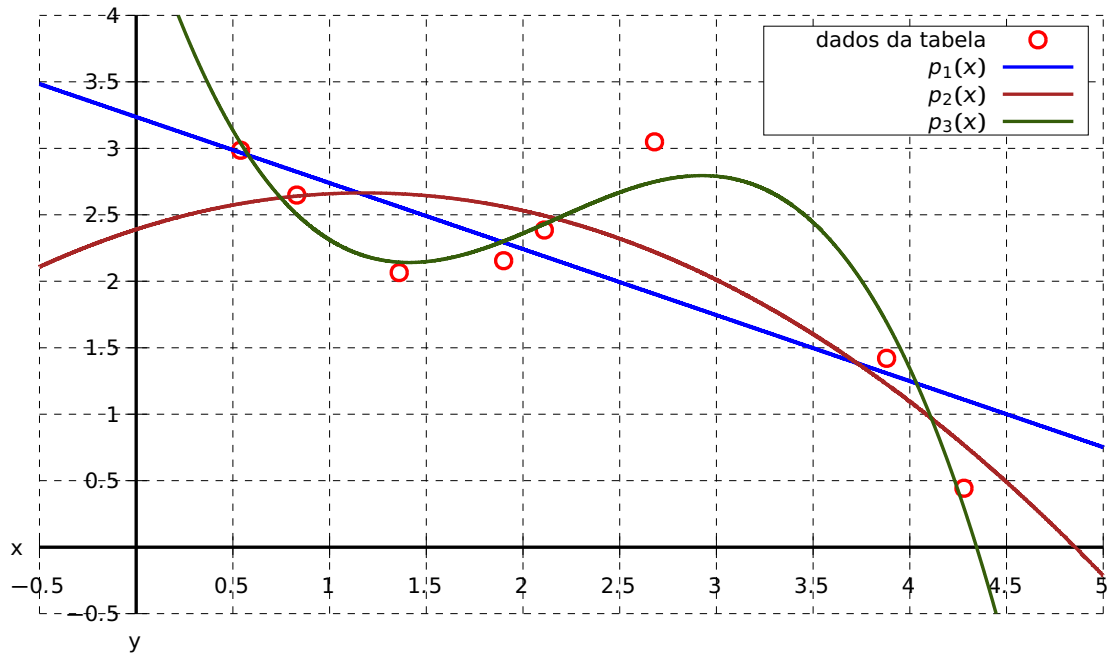


Figura 2.2: Ajuste de dados via polinômios de graus 1, 2 e 3.

Pode-se notar que o aumento do grau dos polinômios faz com que as curvas destes tendam a se aproximar cada vez mais dos pontos associados à tabela de valores utilizados no problema de mínimos quadrados. Ressalta-se ainda o princípio teórico, a ideia base do método, que é minimizar o somatório dos quadrados das diferenças entre os valores tabelados e os valores gerados pelos polinômios. É importante ainda levar em consideração que o processo pode ser aplicado considerando qualquer tipo de função elementar no que tange às funções $\phi_i(x)$ utilizadas.

No próximo capítulo serão apresentados problemas de ajuste de curvas com dados provenientes de laboratórios onde serão apresentados resultados dos ajustes quando se utiliza as equações normais e quando se resolve o problema de mínimos quadrados via Transformações de Householder.

CAPÍTULO 3

Aplicação do Método de Mínimos Quadrados na Física

Esse capítulo apresenta um conjunto de experimentos onde foram realizados ajustes de curvas a partir de dados obtidos de testes de laboratórios de física em modo remoto. Tais informações foram geradas no *Laboratório Remoto de Física* (LABREMOTO) do Instituto de Física e Química, unidade acadêmica da Universidade Federal de Itajubá [10]. Esse laboratório possui várias bancadas para testes de fenômenos físicos sendo que alguns dos mesmos geram dados numéricos, de determinadas propriedades físicas que serão descritas mais adiante, a partir de dados de entrada fornecidos remotamente. Além da exemplificação da aplicação do MMQ na aproximação de funções utilizando dados oriundos do laboratório supracitado, esse capítulo apresenta no final uma proposta de ajustes de dados, de um experimento associado ao decaimento radioativo de um elemento químico, para ser realizado em laboratórios de física no ensino médio, o que é importante como fonte motivacional para estudantes.

3.1 Aspectos computacionais dos experimentos

Os experimentos a serem descritos nas próximas seções foram desenvolvidos sob plataforma Linux, distribuição Debian [8], sendo que os códigos implementados, para a geração das aproximações dos parâmetros que definem as funções de ajuste, foram programados usando o GNU Octave [4], um software livre com foco em matemática computacional, sendo à sintaxe para o desenvolvimento de códigos similar a sintaxe utilizada pelo MATLAB, um outro software, proprietário, com a mesma finalidade. Nesse sentido, dois códigos foram desenvolvidos para obter os parâmetros de ajuste. Um desses códigos tem por base a geração das equações normais e a resolução do sistema linear associados a essas equações usando rotinas prontas do GNU Octave. O outro código foi implementado tendo por base a resolução dos sistemas lineares sobredeterminados associados aos problemas de ajuste utilizando transformações de Householder. Tais códigos são apresentados nos apêndice A.

Nas seções a seguir são apresentados problemas de física cujos dados utilizados, como anteriormente já comentado, foram obtidos através do *Laboratório de Física Remoto* [10].

3.2 Movimento uniforme

O movimento uniforme e uniformemente variado é amplamente discutido nas aulas de física do ensino médio, na matemática é comum que estes movimentos sejam apresentados

nos estudos das funções de 1º grau por fornecer uma situação prática do cotidiano em que a distância varia conforme o tempo passa, sendo assim d uma função de t com relação direta. No caso do movimento uniformemente variado a distância d tem relação de dependência de t^2 .

Na aplicação prática do estudo do movimento podem ocorrer aulas experimentais que buscam comprovar e exemplificar a relação entre d e t . O experimento remoto intitulado "Trilho de ar" realiza o movimento com baixa dissipação da energia por atrito de modo que o experimento seja considerado movimento uniforme. Os dados obtidos em formato .txt são distância em milímetros mm e tempo em milissegundos ms .

O processo de linearização pelo MMQ deve determinar uma função de ajuste de primeiro grau, do tipo $d(t) = v \cdot t$ o que produz uma reta em que o coeficiente angular é a medida da velocidade. O valor de v deve ser uma constante de proporcionalidade entre a distância percorrida e o tempo. Os dados experimentais devem produzir uma reta com inclinação positiva.

Neste experimento remoto é possível produzir um movimento acelerado, classificado como movimento uniformemente variado. Para que isso ocorra é necessário definir uma inclinação do plano em que o carrinho se desloca e executar uma nova medição de tempo e distância. A aceleração do carrinho, o que configura o movimento uniformemente variado é uma componente da aceleração da gravidade. Assim como no experimento anterior os dados obtidos são em .txt e produzem uma equação do tipo $d(t) = At^2$ em que a constante $A = \frac{a}{2}$ obtida pelo MMQ corresponde à metade da aceleração do movimento o que resulta na equação do MUV com d em função de t

$$d(t) = \frac{a}{2}t^2.$$

A figura 3.1 mostra o gráfico de dispersão dos dados gerados do Laboratório Remoto de Física. Nesse experimento não tem a necessidade de entrada de dados no sistema para gerar os dados. Ela exige apenas a inicialização do mecanismo para gerar os dados, o qual fornece a posição, em milímetros (mm), de um corpo em relação ao tempo, este em milissegundos (ms).

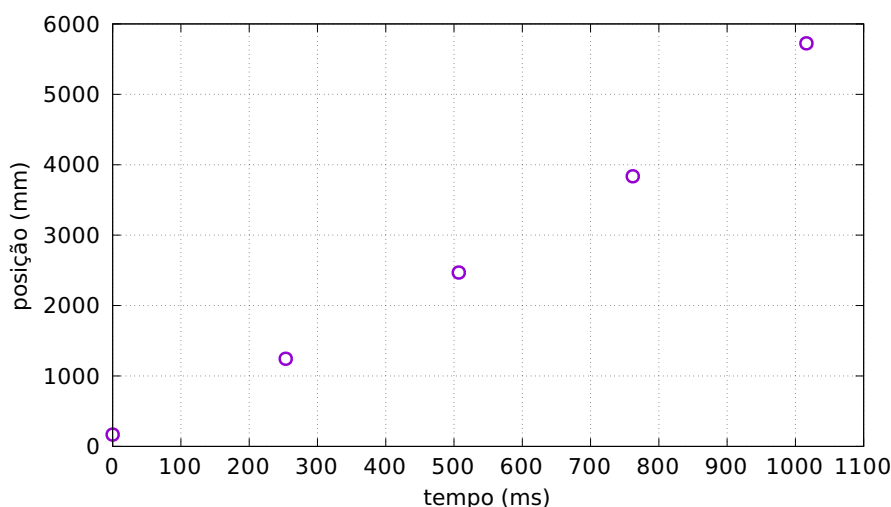


Figura 3.1: Gráfico de dispersão dos dados: Experimento do trilho de ar para obtenção da velocidade.

Esse gráfico de dispersão mostra que a função que melhor se ajusta aos dados é uma reta. Por ser uma reta não existem complicações de ordem computacional no processo de ajuste. A figura 3.2 a seguir apresenta os gráficos da função de ajuste juntamente com os

dados experimentais. Nessa figura pode-se observar que não existe diferença entre o ajuste obtido via resolução das equações normais e o ajuste obtido empregando Transformações de Householder na resolução do sistema linear sobredeterminado associado ao problema de mínimos quadrados.

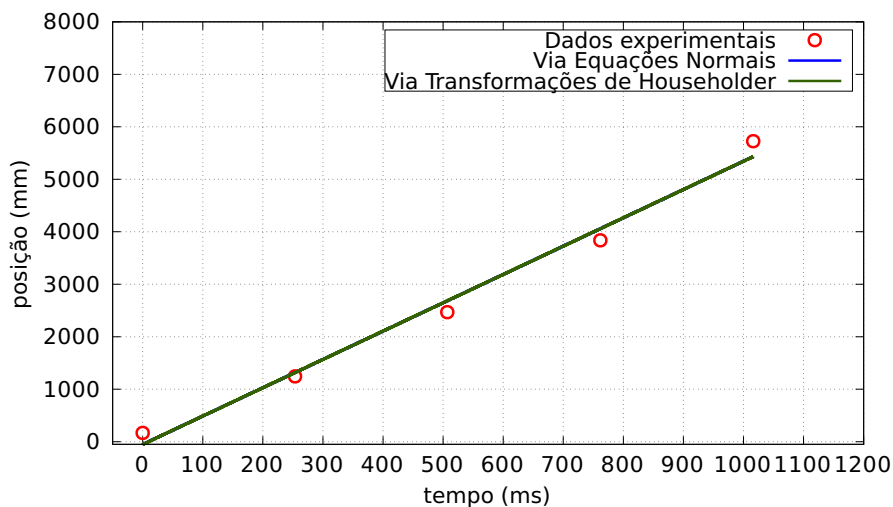


Figura 3.2: Trilho de ar: Ajuste via polinômio de grau 1.

A inclinação de ambas as retas de ajuste fornece a velocidade do corpo no trilho de ar, sendo a inclinação obtida valendo aproximadamente 5.3959 mm/ms .

O próximo gráfico de dispersão na figura 3.3 se refere ao mesmo problema só que o tempo está elevado ao quadrado, sendo que esse problema descreve o movimento de um corpo no chamado movimento retilíneo uniformemente variado, isto é, com aceleração constante. Pode se notar que devido a isso o gráfico de dispersão apresenta uma forma parabólica. E nesse caso, o tipo de função que melhor se ajusta é um polinômio de grau 2, conforme pode ser visto na figura 3.4. A utilização de polinômios de grau 2 pelas duas metodologias para esse conduziu à mesma parábola como pode ser observado no gráfico citado. A expressão do polinômio de ajuste é dado por

$$p_2(t) = 0.00064491t^2 + 0.22438t - 38.973$$

sendo que a metade do coeficiente do termo com t^2 fornece a aceleração do corpo, sendo que esta vale então aproximadamente $1.2898 \cdot 10^{-3} \text{ mm/ms}^2$.

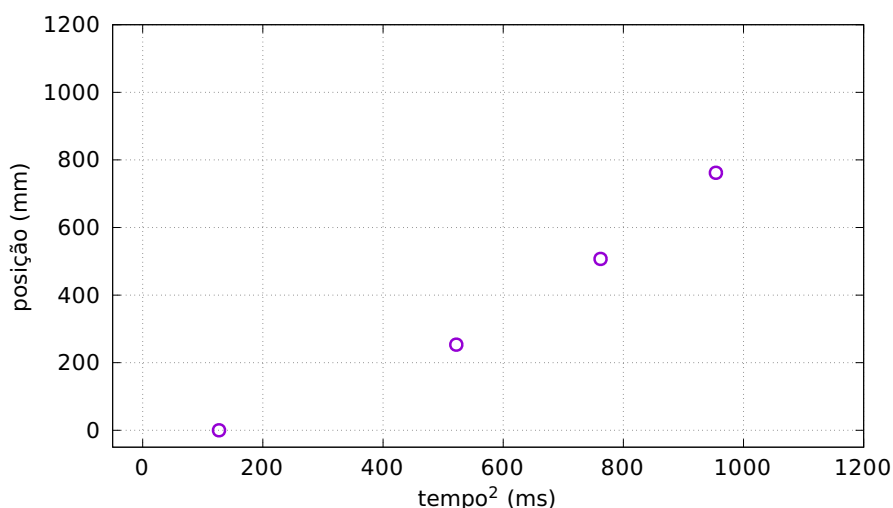


Figura 3.3: Gráfico de dispersão dos dados: Experimento do trilho de ar para obtenção da aceleração.

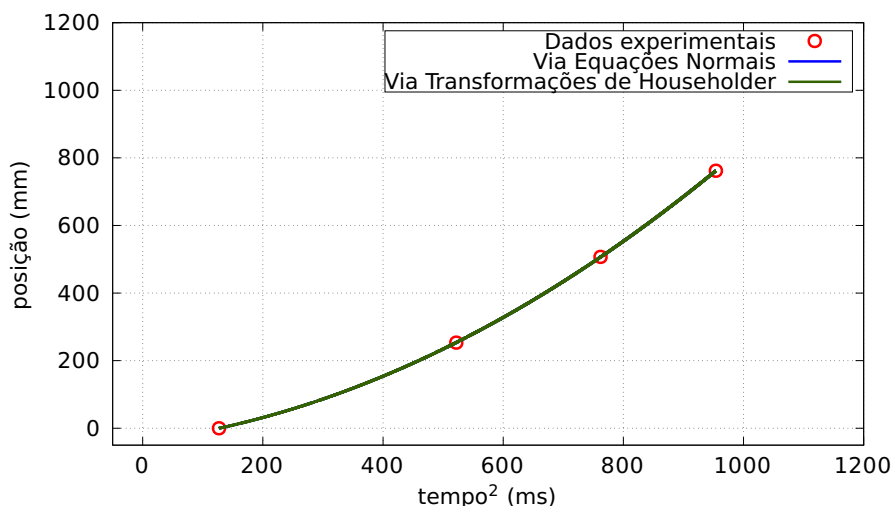


Figura 3.4: Trilho de ar: Ajuste via polinômio de grau 2.

3.3 Determinação da velocidade do som no ar

A acústica é a área da física que estuda as ondas sonoras. Com grande aplicabilidade no dia a dia a acústica explica e modela matematicamente diversos fenômenos. É comum que várias aplicações desse conteúdo recorram ou dependam da velocidade de propagação do som no ar. Esse valor é próximo de 340m/s e por vezes considerado como um valor constante. Ocorre que a temperatura do ar e outras condições climáticas podem alterar ligeiramente esse valor.

O experimento construído para determinação da velocidade de propagação do som no ar disponível no site Labremoto, laboratório remoto, é constituído de um tubo cilíndrico de vidro, uma fonte sonora de frequência fixa em 500 Hertz, válvulas de controle de água, sensores e câmeras. O operador aciona virtualmente o experimento determinando que o tubo fique cheio até o limite máximo, ativa a fonte sonora e clica no comando detectar harmônico, ao detectar o harmônico a válvula de controle do fluxo de água é fechada e nesse momento é possível determinar a distância entre a fonte sonora e a superfície da água. No experimento devem ser detectados os 3 primeiros harmônicos em tubo fechado. Como é conhecida a medida de distância entre harmônicos sucessivos é possível determinar o comprimento de onda para a onda sonora.

A velocidade do som pode ser determinada fazendo o produto entre a frequência e o comprimento de onda medido. O experimento realizado é em tubo fechado (essa denominação considera o tubo fechado em uma extremidade e aberto na outra). O laboratório remoto da UNIFEI possibilita realizar tal experimento considerando ou desprezando a temperatura ambiente, conforme o usuário definir como conveniente, uma vez que a velocidade de propagação do som tem variações devido à temperatura do ar.

O experimento possibilita ao aluno realizar medições de nós consecutivos na propagação de ondas sonoras em tubo fechado através da detecção de harmônicos consecutivos. Sabendo que no tubo fechado os harmônicos são múltiplos do comprimento de onda conforme a relação

$$L = \frac{\lambda}{2}n.$$

A detecção é feita pelo sensor identificando o harmônico sucessivo, assim ao realizar a medida de distância entre os dois nós o valor obtido indica meio comprimento de onda, que é uma medida necessária para determinar a velocidade de propagação do som, uma vez que

a frequência é conhecida.

É possível construir um gráfico que indica a posição em centímetros ou metros e o número do harmônico respectivo e assim determinar com maior precisão o comprimento de onda da onda sonora no meio. Para realizar o cálculo da velocidade de propagação do som no meio é necessário que o valor do comprimento de onda λ esteja em metros, bastando apenas dividir o resultado obtido no gráfico por 100. A frequência selecionada no experimento é de 500 Hz, ao obter o comprimento de onda médio, a determinação da velocidade do som é dada por

$$v = \lambda f.$$

O gráfico de dispersão na figura 3.5 mostra que para ajustar esses dados basta uma reta. Essa reta de ajuste é apresentada no mesmo gráfico, sendo que a obtenção dessa função aproximadora dos dados é a mesma independente do processo de resolução do problema de mínimos quadrados. O coeficiente de inclinação dessa reta corresponde a metade de um comprimento de onda.

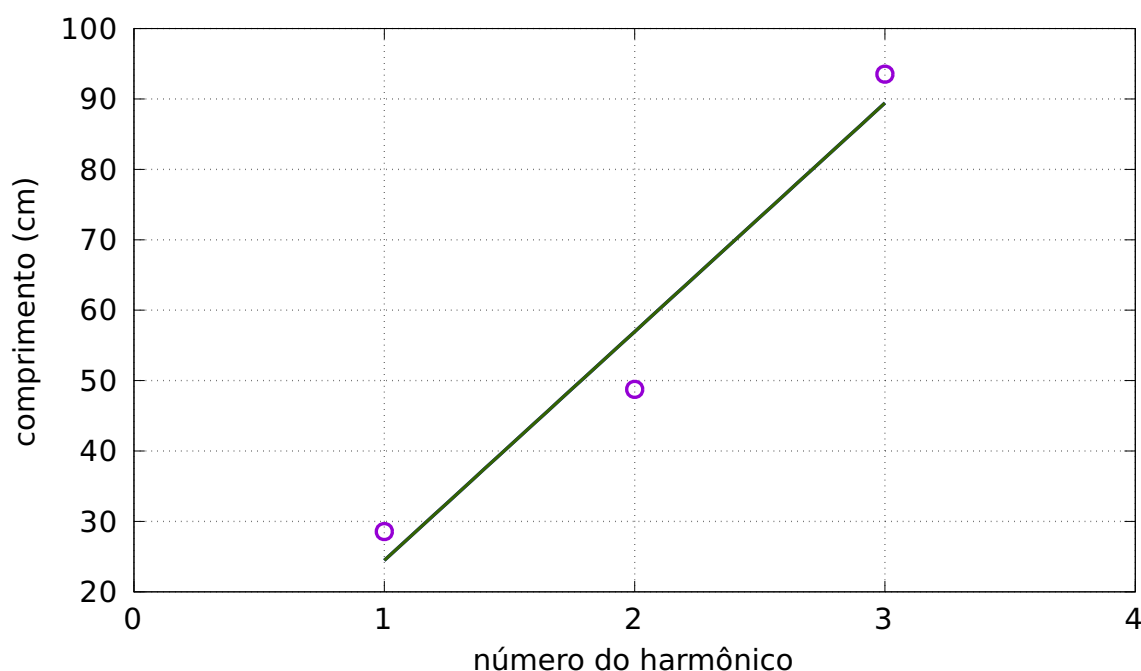


Figura 3.5: Gráfico de dispersão dos dados: Experimento da velocidade do som no ar.

3.4 Curva de luz

A curva de luz é um experimento que descreve o comportamento de uma estrela variável Cefeida, sendo uma estrela de taxa de brilho variável. O experimento em questão é relacionado à astrofísica e faz parte de um conjunto de experimentos remotos disponíveis para acesso online. A iniciativa é da Universidade Federal de Itajubá - UNIFEI, que tem o curso de licenciatura em física com forte direcionamento para a astrofísica.

Esse comportamento da estrela Cefeida fornece dados que podem determinar a distância em função do período de pulsação da estrela. O que é útil para oferecer uma mediação indireta da distância e apresentar noções sobre medidas de distâncias em astrofísica.

Os dados relativos à sua magnitude em função do tempo, para facilitar a obtenção dos dados o tempo é definido em dias. O assunto é trabalhado em uma disciplina de introdução à Astronomia no curso de licenciatura em Física. O experimento pretende apresentar aos alunos noções sobre técnicas de medições e grandezas relacionadas à astrofísica.

Neste experimento o tempo, que na prática é medido em dias, é medido em segundos e indicado em dias. A lâmpada utilizada no experimento tem comportamento, em termos de emissão de luz, semelhante a uma estrela Cefeida de brilho variável. Observações reais dependem de condições climáticas favoráveis e disponibilidade de equipamentos adequados.

Em um experimento real a obtenção desses dados demoraria semanas e seria dependente de boas condições climáticas. A magnitude é dada em função do tempo em dias, sendo esta um relação inversamente proporcional ao brilho emitido pela estrela. Os dados obtidos no laboratório remoto são medidas do fluxo luminoso em função do tempo.

Os dados obtidos no experimento da curva de luz podem ser baixados em formato .txt o que facilita sua inserção no programa para realização dos testes para busca da melhor função de ajuste. A teoria mostra que a função associada ao gráfico de dispersão dos dados é uma *série de Fourier* dada pela função

$$\omega(t) = A + B \sum_{k=1}^n \frac{1}{k} \text{sen}(n\omega_L t)$$

em que ω_L é a frequência angular e equivale a $2\pi P_L^{-1}$. O parâmetro B da equação tem relação com a amplitude da onda e A, com o seu deslocamento ao longo do eixo y. As informações descritas foram baseadas em Thiago [2].

O gráfico de dispersão na figura 3.6 representa um extrato dos dados gerados pela bancada desse experimento no intervalo de tempo igual a 20 segundos que corresponde a 20 dias no experimento, informação que consta no eixo x do gráfico. No eixo y temos a magnitude $\omega(t)$, a qual fornece relação do brilho em função do tempo t.

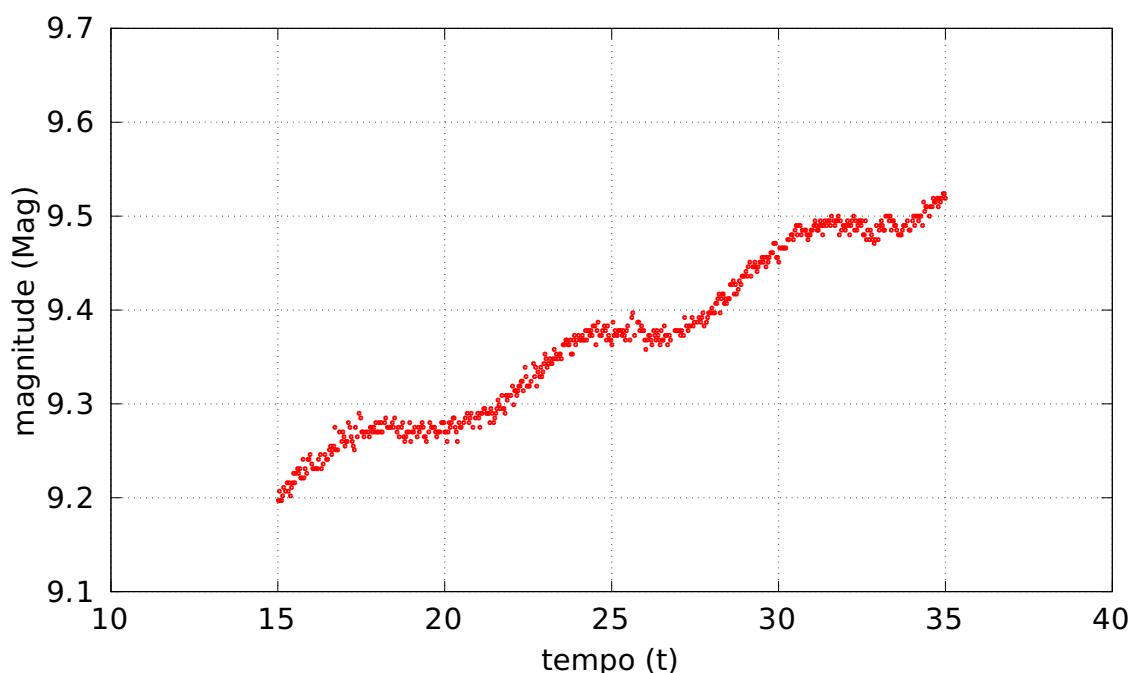


Figura 3.6: Gráfico de dispersão dos dados: Experimento da curva de luz.

Para esse caso, pelo próprio comportamento do gráfico de dispersão, o ideal é fazer o ajuste dos dados por meio de combinações de funções trigonométricas. Mas considerando o ponto de vista da investigação, foi realizado testes de ajustes por meio de polinômios com graus variando de 3 a 10. As figuras 3.7, 3.8, 3.9, 3.10, 3.11 e 3.12 apresentam gráficos dos polinômios de ajuste para os dados do experimento, onde se tem os referidos polinômios obtidos resolvendo-se utilizando as equações normais e os polinômios obtidos usando Transformação de Householder para resolver os sistemas lineares sobredeterminados necessários

à resolução do problema de mínimos quadrados. Após as figuras seguem comentários.

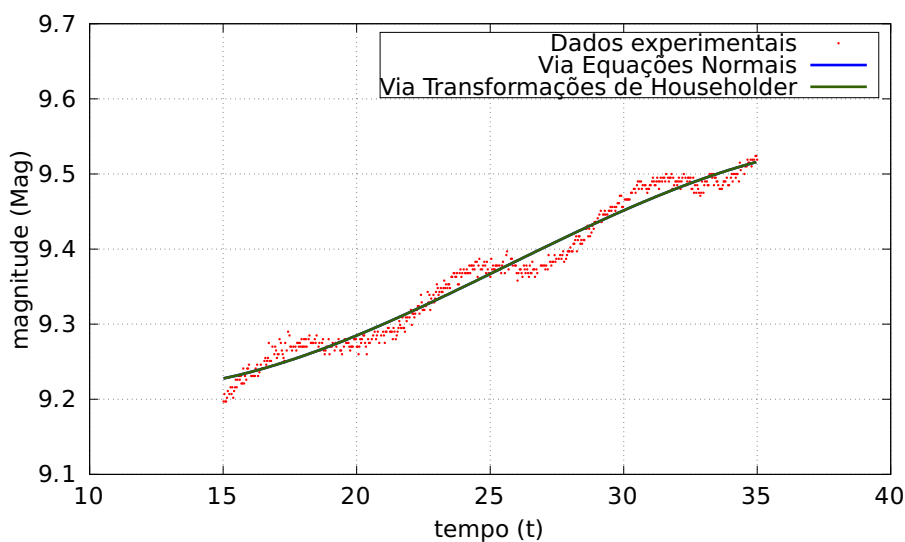


Figura 3.7: Gráfico de dispersão dos dados: Ajuste via polinômios de grau 3.

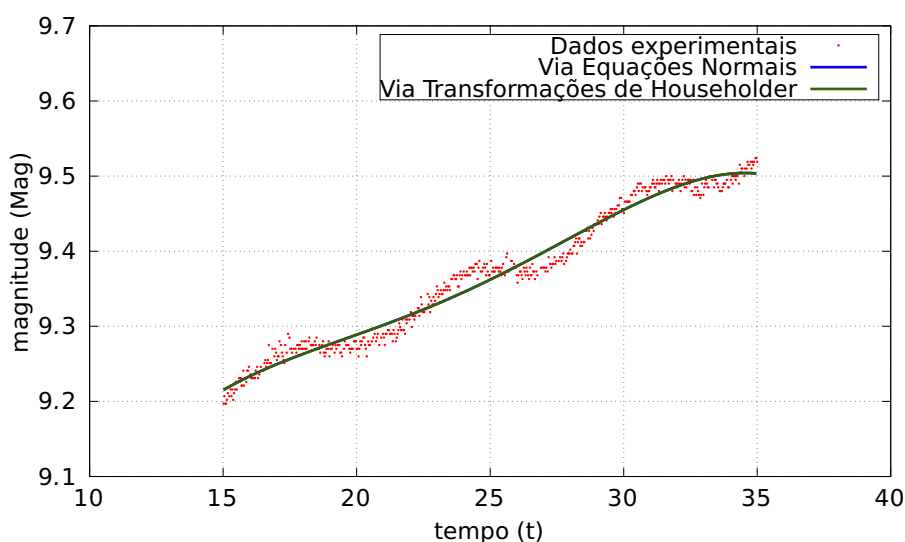


Figura 3.8: Gráfico de dispersão dos dados: Ajuste via polinômios de grau 4.

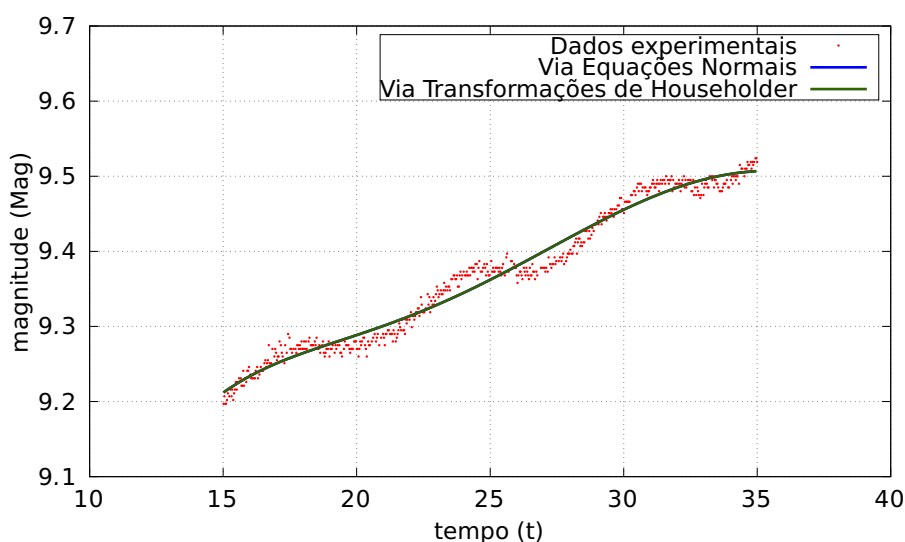


Figura 3.9: Gráfico de dispersão dos dados: Ajuste via polinômios de grau 5.

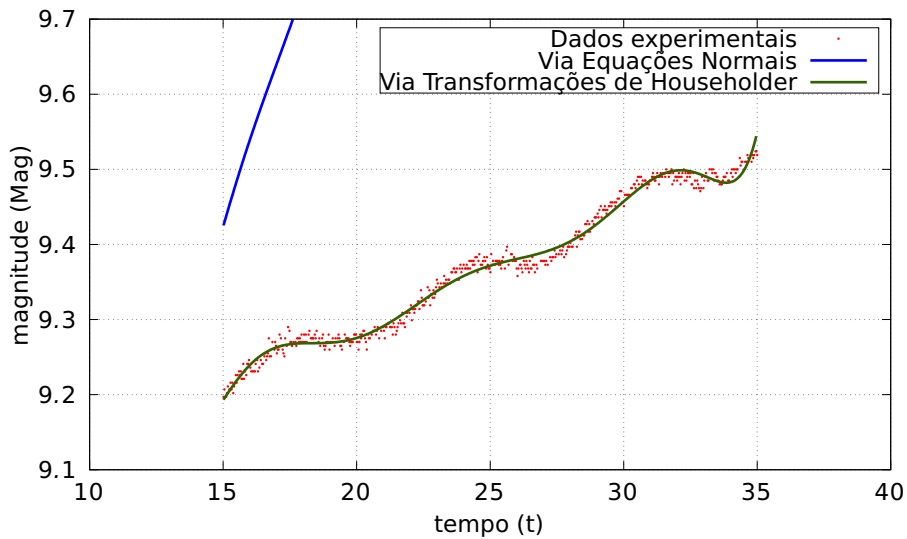


Figura 3.10: Gráfico de dispersão dos dados: Ajuste via polinômios de grau 8.

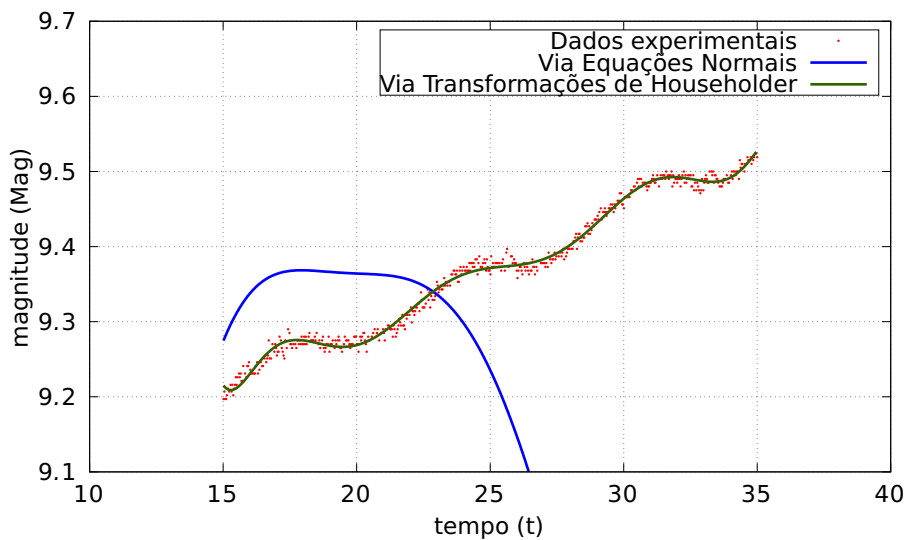


Figura 3.11: Gráfico de dispersão dos dados: Ajuste via polinômios de grau 9.

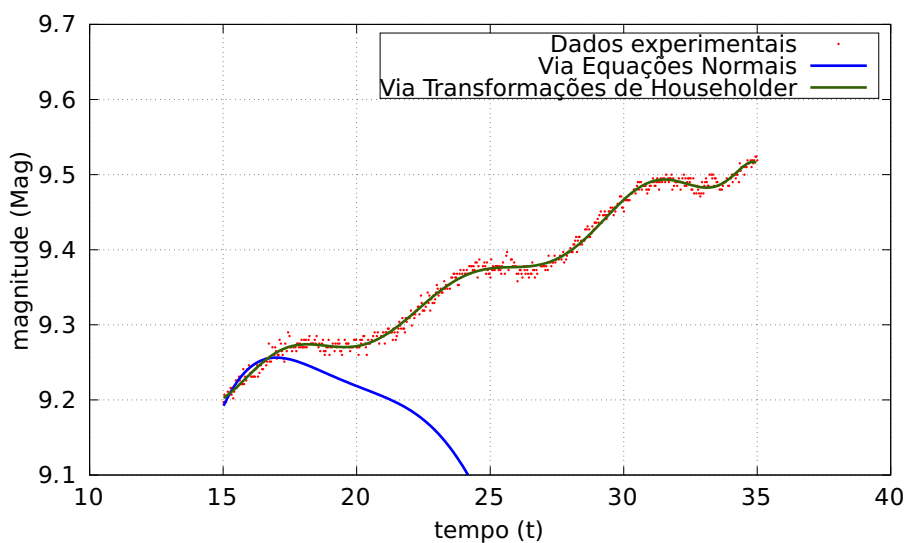


Figura 3.12: Gráfico de dispersão dos dados: Ajuste via polinômios de grau 10.

Pode-se notar que os ajustes usando as duas metodologias de obtenção do problema de mínimos quadrados não tem diferença visível nos gráficos no caso dos polinômios de graus 3,

4 e 5. Agora, nos casos onde os ajustes foram realizados por meio de polinômios de graus 8, 9 e 10 pode-se notar uma grande diferença entre as duas metodologias. Quando no ajuste foi empregado as transformações de Householder para resolver os sistemas lineares sobredeterminados não houve instabilidade devido a erros de arredondamento, sendo que os polinômios foram gradualmente se ajustando aos dados tabelados a medida que se aumentava o grau desses polinômios. Agora, os ajustes obtidos através da resolução das equações normais geraram polinômios sem nenhuma relação com os dados tabelados, divergindo em muito com os dados do gráfico de dispersão. Isso tudo se deve à tendência do mal condicionamento das equações normais, isto é, da sensibilidade de tais equações aos erros de arredondamento inerentes ao processo computacional. O mal condicionamento significa que a matriz dos coeficientes tende a ser singular, isto é, tende a ter determinante igual a zero. Assim sendo, todo sistema linear cujo determinante tende a se anular tende a ser muito sensível aos erros de arredondamento.

Embora na prática não se utiliza polinômios de graus elevados devido às tendências oscilatórias que em alguns casos não é aconselhável. A abordagem apresentada acima ilustra a aplicação do MMQ bem como os cuidados necessários aos métodos empregado no ajuste.

3.5 Decaimento radioativo

Os dados obtidos para o experimento do decaimento radioativo são de duas fontes distintas: um experimento real realizado em laboratório de física moderna e uma proposta de aula que simula o decaimento radioativo com moedas.

A proposta inicial para o experimento do decaimento radiativo é baseado em Carvalho [3], e apresenta uma proposta didática para uso em sala de aula. Considerando que a escola básica não tem acesso a materiais radioativos e a equipamentos de medição adequados para executar tal experimento físico real.

Para aplicar o MMQ no ajuste de dados do decaimento radioativo serão utilizados dados reais obtidos no laboratório de física moderna da UFU-Campus Ituiutaba no ano de 2015 durante a aula sobre decaimento radioativo da disciplina de Física Moderna do Curso de Mestrado Profissional em Ensino de Ciências e Matemática da UFU. Nesta aula foi realizada a medição da emissão de partículas radioativas do Ba 137, isótopo radioativo obtido através da eluição de Cs 137. Os dados referentes à medição da emissão de partículas foram coletados por um contador Giger Muller registrados em função do tempo, sendo a relação partículas por segundo.

É conhecido da física e da química que o decaimento radioativo segue uma lei física que pode ser descrita conforme uma equação exponencial. Essa equação apresenta o número de partícula em função de um valor inicial considerando o tempo de meia vida. Essa relação exponencial também é utilizada para a estimativa de população de bactérias e outros sistemas biológicos que aumentam ou diminuem seguindo o padrão de uma função exponencial.

É esperado que os dados forneçam uma curva exponencial negativa, do tipo

$$N(t) = N_0 e^{-\lambda(t)},$$

onde N_0 é quantidade inicial de partículas quando $t = 0$ e o parâmetro λ corresponde ao tempo de meia-vida do material radioativo, sendo este um valor específico por elemento químico.

Após realizar o ajuste o tempo de meia-vida para o Ba137 é possível determinar, através da determinação do coeficiente λ e o valor de N_0 . Sendo estes os valores do tempo de meia-vida para o elemento radioativo e o número inicial de partículas. O valor determinado

experimentalmente pode variar ligeiramente do valor de $N(0)$, valor quando $t = 0$, o que corresponde ao valor inicial do número de partículas emitidas pela fonte radioativa.

O gráfico na figura 3.13 apresenta o ajuste de dados, sendo que estes foram obtidos durante aula prática da disciplina física moderna na UFU em Ituiutaba em 2015. A dispersão dos dados inferiu que a função de ajuste mais adequada deve ter a forma

$$f(x) = ae^{bx},$$

um tipo conhecido dentro do MMQ como um caso não linear, sendo necessário antes de fazer o ajuste realizar uma linearização para que seja possível aplicar o MMQ. Tal linearização consiste nas seguintes operações analíticas

$$f(x) = ae^{bx} \Rightarrow \ln(f(x)) = \ln(a) + bx \Rightarrow F(x) = \alpha + \beta x \Rightarrow \begin{cases} F(x) = \ln(a) \\ \alpha = \ln(a) \\ \beta = b \end{cases}$$

sendo que depois de ser efetivado o ajuste da função linearizada deve-se retornar aos parâmetros originais, sendo que $a = e^\alpha$ e $b = \beta$.

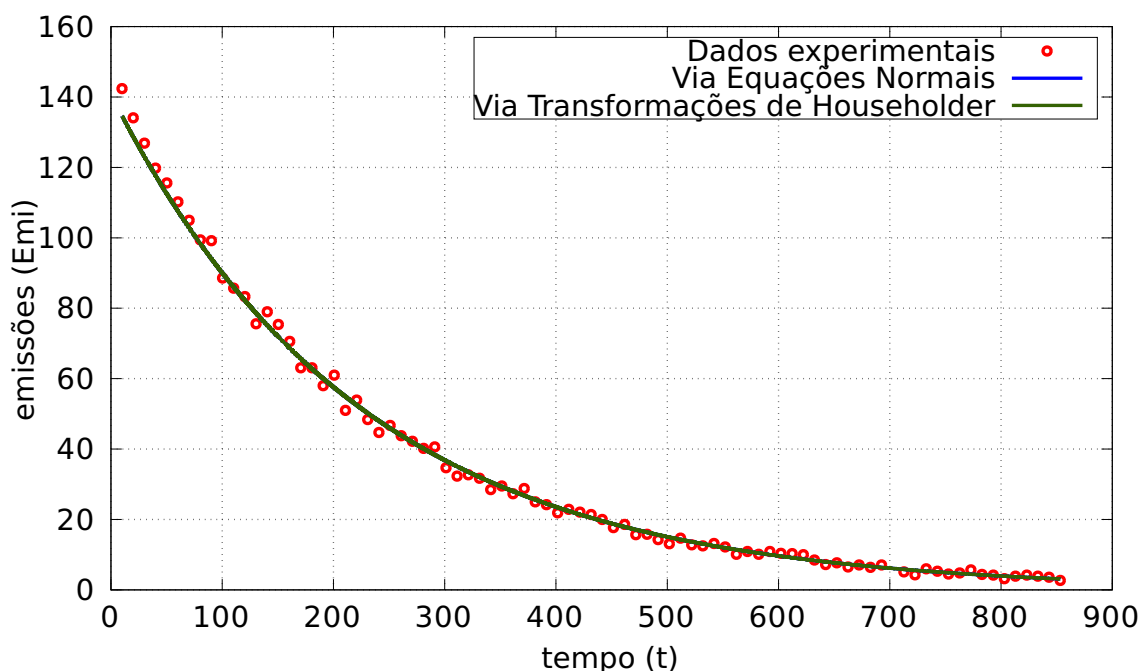


Figura 3.13: Gráfico de dispersão dos dados: Decaimento radioativo.

Do ponto de vista físico, a informação mais importante nesse ajuste é o módulo do parâmetro b , pois ele fornece o tempo de meia vida do elemento. Nesse experimento o valor desse parâmetro é dado pelo valor aproximado $4.4666 \cdot 10^{-3}$ segundos.

Esse experimento encerra a exemplificação no que tange à aplicação do MMQ em problemas de física, abordando vários aspectos teóricos e práticos do método. O próximo capítulo finaliza o texto apresentando as conclusões obtidas no estudo realizado.

CAPÍTULO 4

Conclusões

O estudo realizado apresenta conclusões em duas linhas, sendo uma no que tange os aspectos matemáticos e computacionais e a outra nos aspectos da aplicabilidade no estudo de problemas de física.

No que tange aos aspectos matemáticos foi observado a grande carga de teoria que existe nos fundamentos do MMQ. Das duas formas de apresentação do processo de resolução do referido método, a mais simples é a que desenvolve essa resolução do problema de mínimos quadrados através do uso do Cálculo Diferencial e Integral. De fato, o processo de montagem da função $F(\alpha_1, \alpha_2, \dots, \alpha_m)$ e o cálculo de suas derivadas parciais para determinar o ponto crítico dessa função seguido pela montagem do sistema linear, cujas equações constituem as equações normais, é o modo mais simples de entendimento do método. Isso talvez justifique o fato de que muitos livros de Cálculo Numérico, em nível de graduação, apresente a teoria do MMQ por esse metodologia. Já a apresentação teórica do MMQ via embasamentos teóricos da Álgebra Linear exige uma carga maior de conhecimento para seu entendimento pois exige que se tenha bem claro temas como espaços vetoriais, base e dimensão de espaços vetoriais, coordenadas de vetores em relação a uma dada base, produto interno e projeção ortogonal. Além disso, muitas vês dependendo de quem estuda o tema, nem sempre é fácil visualizar as projeções de vetores em subespaços para entender a resolução do processo de obtenção da solução do problema de mínimos quadrados. Esses fatores talvez justifique o motivo de muitos autores de livros de Cálculo Numérico evitarem a explicação do MMQ por essa via, optando de modo justificado, o outro meio de dedução do método, facilitando assim o processo de ensino e aprendizagem.

No que tange aos aspectos computacionais, foi observado a importância de se escolher o método numérico adequado para resolver o sistema linear sobredeterminado para garantir estabilidade numérica nas operações aritméticas realizadas pelo computador. Dessa forma, evita-se que os erros de arredondamento, inerentes às operações computacionais, afetem a obtenção dos parâmetros de ajuste. Nesse sentido, do ponto de vista da resolução de um problema físico real, a utilização de métodos numéricos inadequados pode conduzir a conclusões erradas acerca do fenômeno físico em estudo. E ainda dentro da questão computacional, a resolução de problema de mínimos quadrados através da resolução do sistema de equações normais tem um custo computacional bem menor do que a resolução do mesmo problema de mínimos quadrados através da aplicação das transformações de Householder sobre os sistemas lineares sobredeterminados, porém, bem mais seguro e mais preciso.

E no campo da física, o MMQ mostrou ser um técnica de aproximação de funções valioso para o estudo de fenômenos físicos. O grau de relevância nesse aspecto é alto, visto

que o desenvolvimento de modelos matemáticos para fenômenos físicos é uma tarefa árdua, demandando grande acurácia na análise de propriedades físicas de tal forma a obter modelos com um determinado nível de confiança, mesmo sabendo de muitas imponderações que podem ocorrer no processo.

Considerando agora aspectos educacionais, em uma proposta futura, em nível de ensino médio, desde que o tema seja adequadamente preparado, é possível o desenvolvimento de projetos de ensino de física, dentro de problemas onde o MMQ é aplicável, a implementação de experimentos de física, bem como a análise dos mesmos, através de metodologias adequadas aos conhecimentos dos educandos. Dessa forma, é possível estimular o significado do que é estudado em teoria dentro da visão prática em experimentos laboratoriais. Para tanto, em termos de ensino de nível médio, é necessário um treinamento prévio no uso de softwares livres, como o GNU Octave empregado nesse trabalho, bem como o treinamento no que tange à preparação de bancadas de experimentos com os recursos de medição adequados, bem como o uso dos mesmos.

Em síntese, esse estudo mostrou, em um nível mais elevado, que a aliança da Matemática com a Física, ambas com suporte da ciência da computação, representa uma união fundamental no processo de estudo e resolução de uma vasta gama de problemas dentro da física e, em um sentido mais amplo, das ciências naturais.

APÊNDICE A

Códigos Implementados em Octave

A.1 Ajuste de Curvas via Equações Normais

O código abaixo resolve o problema de mínimos quadrados através da montagem sistema linear das equações normais e da resolução desses sistemas utilizando rotinas prontas do Octave.

```
%=====
% UNIVERSIDADE FEDERAL DE UBERLÂNDIA
% FACULDADE DE MATEMÁTICA
% LICENCIATURA EM MATEMÁTICA EAD
% TRABALHO DE CONCLUSÃO DE CURSO
% PROGRAMA: RESOLUÇÃO DE PROBLEMA DE MÍNIMOS QUADRADOS USANDO
%           EQUAÇÕES NORMAIS
%=====
% LIMPEZA DA MEMÓRIA
%=====
clear
%=====
% CONFIGURAÇÃO DA QUANTIDADE DE CASAS PARA IMPRESSÃO EM TELA
%=====
%format long
%=====
% ENTRADA DE DADOS
%=====
% LEITURA DO ARQUIVO DE DADOS
%=====
MAT=load('dados.mat');
%=====
% NÚMERO DE PONTOS DA TABELA PARA AJUSTE
%=====
n=size(MAT,1);
%=====
% ATRIBUIÇÃO PARA X E Y
%=====
x=MAT(:,1);
```



```

f=MAT(:,2);
%=====
% DELETA MATRIZ COM ENTRADA DE DADOS AUXILIAR
%=====
clear MAT ;
%=====
% ENTRADA DO GRAU DO POLINÔMIO DE AJUSTE
%=====
disp('MMQ USANDO EQUAÇÕES NORMAIS');
N=input('Informe o grau do polinomio de ajuste: ');
%=====
% MONTAGEM DO SISTEMA LINEAR (MATRIZ DOS COEFICIENTES A E VETOR b)
%=====
for i=1:n
    for j=0:N
        A(i,j+1)=(x(i))^(j);
    end
    b(i)=f(i);
end
b=b';
MA0=A;
vb0=b;
%=====
% MONTAGEM DAS EQUAÇÕES NORMAIS
%=====
M=A';
A=M*A;
b=M*b;
clear M;
%=====
% RESOLUÇÃO DO SISTEMA LINEAR
%=====
c=inv(A)*b;
%=====
% CÁLCULO DO RESÍDUO
%=====
resmmqen=norm(vb0-MA0*c,2)
%=====
% IMPRESSÃO DOS COEFICIENTES DE AJUSTE
%=====
c=flip(c);
c'
%=====
% GERAÇÃO DE DADOS DO POLINÔMIO DE AJUSTE PARA PLOTAGEM VIA GNUPLT
%=====
xx=x(1):0.05:x(n);
yx=polyval(c,xx);

```

```
M=[xx' yx'];  
save 'polmmqen.dat' M
```

A.2 Ajuste de Curvas via Transformação de Householder

O código abaixo foi implementado com a finalidade de resolver o problemas de mínimos quadrados empregando no processo de resolução dos sistemas lineares sobredeterminados as Transformações de Householder com a finalidade de garantir estabilidade numérica nas operações computacionais.

```
%=====
% UNIVERSIDADE FEDERAL DE UBERLÂNDIA
% FACULDADE DE MATEMÁTICA
% LICENCIATURA EM MATEMÁTICA EAD
% TRABALHO DE CONCLUSÃO DE CURSO
% PROGRAMA: RESOLUÇÃO DE PROBLEMA DE MÍNIMOS QUADRADOS USANDO
%           TRANSFORMAÇÃO DE HOUSEHOLDER
%=====
% LIMPEZA DA MEMÓRIA
%=====
clear
%=====
% CONFIGURAÇÃO DA QUANTIDADE DE CASAS PARA IMPRESSÃO EM TELA
%=====
%format long
%=====
% ENTRADA DE DADOS
%=====
% LEITURA DO ARQUIVO DE DADOS
%=====
MAT=load('dados.mat');
%=====
% NÚMERO DE PONTOS DA TABELA PARA AJUSTE
%=====
n=size(MAT,1);
%=====
% ATRIBUIÇÃO PARA X E Y
%=====
x=MAT(:,1);
f=MAT(:,2);
%=====
% DELETA MATRIZ COM ENTRADA DE DADOS AUXILIAR
%=====
clear MAT ;
%=====
% ENTRADA DO GRAU DO POLINÔMIO DE AJUSTE
%=====
disp('MMQ VIA TH');
```

```
N=input('Informe o grau do polinomio de ajuste: ');
%=====
% MONTAGEM DO SISTEMA LINEAR (MATRIZ DOS COEFICIENTES A E VETOR b)
%=====
for i=1:n
    for j=0:N
        A(i,j+1)=(x(i))^(j);
    end
    b(i)=log(f(i));
end
b=b';
MA0=A;
vb0=b;
%=====
% APLICAÇÃO DAS TRANSFORMAÇÕES DE HOUSEHOLDER SOBRE O SISTEMA LINEAR
%=====
for k=1:N+1
    xaux=A(k:n,k);
    alfa=norm(xaux,2);
    beta=alfa*(alfa-xaux(1));
    v=xaux;
    v(1)=xaux(1)-alfa;
    r=1/beta;
    H=eye(n-k+1)-r*v*v';
    A(k:n,k:N+1)=H*A(k:n,k:N+1);
    b(k:n)=H*b(k:n);
end
%=====
% RESOLUÇÃO DO SISTEMA LINEAR
%=====
c(N+1)=b(N+1)/A(N+1,N+1);
for i=N:-1:1
    soma=0;
    for j=i+1:N+1
        soma=soma+A(i,j)*c(j);
    end
    c(i)=(b(i)-soma)/A(i,i);
end
%=====
% CÁLCULO DO RESÍDUO
%=====
resmmqth=norm(vb0-MA0*c',2)
%=====
% IMPRESSÃO DOS COEFICIENTES DE AJUSTE
%=====
c=flip(c)
%=====
```

```
% GERAÇÃO DE DADOS DO POLINÔMIO DE AJUSTE PARA PLOTAGEM VIA GNUPLOT
%=====
##xx=x(1):0.05:x(n);
##yx=polyval(c,xx);
##M=[xx' yx'];
##save 'polmmqth.dat' M
%=====
% GERAÇÃO DE DADOS DA FUNÇÃO F(X)=A*EXP(B*X)
%=====
c=flip(c);
c1=exp(c(1));
c2=c(2);
xx=x(1):0.05:x(n);
yx=c1*exp(c2*xx);
M=[xx' yx'];
save 'polmmqth.dat' M
```

Referências Bibliográficas

- [1] H. Anton e C. Rorres: *Álgebra Linear com Aplicações*. Bookman, 2010.
- [2] T. C. Caetano: *O experimento "Curva de Luz" do Laboratório Remoto de Física: Uma Proposta de Atividade Investigativa Contextualizada Epistemologicamente*.
- [3] R. P. Carvalho e S. M. V. de Oliveira: *Aplicações da Energia Nuclear na Saúde*. SBPC, 2017.
- [4] J. W. Eaton e outros: *GNU Octave*. <https://www.gnu.org/software/octave/index>, 1988.
- [5] N. B. Franco: *Cálculo Numérico*. Pearson, 2006.
- [6] G. H. Golub e C. F. V. Loan: *Matrix Computations*. Johns Hopkins, 1996.
- [7] S. J. Leon: *Álgebra Linear com Aplicações*. LTC, 2008.
- [8] Projeto Debian e Outros: *Debian Linux*. <https://www.debian.org/>, 1993.
- [9] V. L. da Rocha Lopes e M. A. G. Ruggiero: *Cálculo Numérico - Aspectos Teóricos e Computacionais*. Pearson, 2000.
- [10] Universidade Federal de Itajubá - Instituto de Física e Química: *Laboratório Remoto de Física (LABREMOTO)*. <https://labremoto.unifei.edu.br/src/welcome.php>, 2021.