



DEM - Departamento de Engenharia Mecânica

Relatório Final da Iniciação Científica

Identificação de Parâmetros pelo Método dos Mínimos Quadrados Não Linear

Aluna: Roberta de Queiroz Lima

Orientador: Rubens Sampaio



Rio de Janeiro, Agosto de 2009

Sumário

Lista de Figuras	ii
1 Introdução	3
2 Conceitos de Álgebra Linear Utilizados no Método dos Mínimos Quadrados	5
2.1 Convenção de Notação em Álgebra Linear	5
2.2 Espaços Vetoriais	6
2.3 Subespaços Vetoriais	6
2.4 Independência Linear	7
2.4.1 Dependência e Independência Linear	8
2.5 Bases	9
2.6 Subespaços Fundamentais de uma Matriz	11
2.6.1 Espaço Imagem e Espaço Nulo	11
2.7 Produto Interno	11
2.8 Matrizes Positivas Definidas	14
2.8.1 Matrizes Positivas Definidas	14
2.8.2 Matrizes Positivas Semi-Definidas	16
2.8.3 Matrizes de Gram	16
3 Método dos Mínimos Quadrados	19
3.1 Exemplos de Problemas de Minimização	19
3.2 Minimização de Funções Quadráticas	20
3.3 Solução de um Problema Geométrico	22
3.4 Mínimos Quadrados Aplicados à Solução de Sistema de Equações Lineares	24
3.5 Mínimos Quadrados Aplicados ao Ajuste de Curvas	26
3.5.1 Ajuste de Polinômios	28
4 Método dos Mínimos Quadrados Não Linear	30
4.1 Método de Gauss-Newton	32
4.2 Método de Levenberg-Marquardt	33
5 Parâmetros do Metabolismo Aeróbio	36
5.1 Introdução ao Metabolismo Aeróbio	36
5.2 Fases do Metabolismo Aeróbio	37
5.3 Identificação dos Parâmetros do Metabolismo Aeróbio	39

6	Levenberg-Marquardt Aplicado à Identificação dos Parâmetros da curva de $\dot{V}O_2(t)$	43
6.1	Programa Computacional Desenvolvido	43
6.1.1	Equações de Levenberg-Marquardt Aplicadas à curva de $\dot{V}O_2(t)$	43
6.1.2	Diagrama Modular do Programa Computacional	45
6.2	Resultados	47
7	Conclusões	56
	Referências Bibliográficas	57

Lista de Figuras

2.1	Norma Euclidiana para o \mathbb{R}^2 e o \mathbb{R}^3	12
3.1	Ajuste de $y(t) = \alpha + \beta t$ pelo método dos mínimos quadrados	26
5.1	Esquema mostrando as três fases exponenciais de $\dot{V}O_2(t)$ para exercícios de intensidade pesada	38
5.2	Esquema mostrando as curvas de $\dot{V}O_2(t)$ para exercícios de intensidade moderada, pesada e severa. A parte hachurada de cinza mostra a fase III para os exercícios de intensidade pesada e severa.	39
5.3	Esquema mostrando as combinações válidas dos instantes de mudança de fase TDP e TDs para os m instantes de tempo avaliados.	41
6.1	Diagrama Modular do Programa Computacional	47
6.2	Paciente 1 - Curva de $\dot{V}O_2(t)$ ajustada por Levenberg-Marquardt	48
6.3	Paciente 2 - Curva de $\dot{V}O_2(t)$ ajustada por Levenberg-Marquardt	49
6.4	Paciente 3 - Curva de $\dot{V}O_2(t)$ ajustada por Levenberg-Marquardt	49
6.5	Paciente 4 - Curva de $\dot{V}O_2(t)$ ajustada por Levenberg-Marquardt	50
6.6	Combinações entre TDP e TDs avaliadas pelo programa <i>Matlab</i>	50
6.7	Paciente 1 - Erro do Ajuste em função das combinações entre TDP e TDs	51
6.8	Paciente 1 - Comparação entre as curvas de $\dot{V}O_2(t)$ geradas por x_0 e pelos parâmetros calculados por Levenberg-Marquardt	52
6.9	Paciente 2 - Comparação entre as curvas de $\dot{V}O_2(t)$ geradas por x_0 e pelos parâmetros calculados por Levenberg-Marquardt	52
6.10	Paciente 3 - Comparação entre as curvas de $\dot{V}O_2(t)$ geradas por x_0 e pelos parâmetros calculados por Levenberg-Marquardt	53
6.11	Paciente 4 - Comparação entre as curvas de $\dot{V}O_2(t)$ geradas por x_0 e pelos parâmetros calculados por Levenberg-Marquardt	53
6.12	Paciente 1 - $F(x)$ em função do número de iterações	54
6.13	Paciente 2 - $F(x)$ em função do número de iterações	54
6.14	Paciente 3 - $F(x)$ em função do número de iterações	55
6.15	Paciente 4 - $F(x)$ em função do número de iterações	55

Resumo

Foi desenvolvido um programa computacional em *Matlab* que faz a identificação de parâmetros pelo método dos mínimos quadrados não linear. São identificados os parâmetros da curva de variação no tempo da taxa de absorção de oxigênio ($\dot{V}O_2(t)$) de uma pessoa durante a realização de exercício físico. O exercício possui carga constante e sua intensidade é classificada como pesada ou severa. Os valores de $\dot{V}O_2$ são medidos através do processo de ventilação pulmonar.

Na literatura, a variação da taxa de $\dot{V}O_2$ para exercícios de intensidade pesada-severa é modelada por três funções exponenciais seqüenciais. Cada uma dessas funções é identificada como uma fase fisiológica distinta: fases I, II e III. Elas representam a resposta do organismo ao aumento do metabolismo energético. E refletem o ajuste sistêmico que ocorre entre os sistemas respiratório, cardíaco, vascular e muscular, responsáveis pela captação, transporte e utilização de O_2 .

O ajuste matemático das três fases exige a identificação de nove parâmetros, incluindo os instantes de tempo em que há a mudança das fases I para II e II para III. O método escolhido para fazer-se o ajuste pelo mínimos quadrados não linear foi **Levenberg-Marquardt**.

PALAVRAS - CHAVE: Mínimos Quadrados Não Linear - Levenberg-Marquardt.

Abstract

The main purpose of this project is develop a computer program in *Matlab* to make identification of parameters by the nonlinear least squares method. Will be identify the parameters of the curve of oxygen absorption rate ($\dot{V}O_2(t)$) of a person during the performance of physical exercise. The exercise load is constant and its intensity is classified as heavy or severe. The values of $\dot{V}O_2$ are measured through the process of means ventilation.

In literature, the variation of $\dot{V}O_2$ for physical exercise of heavy or severe intensity is modeled by three sequential exponential functions. Each of these functions is identified as a distinct physiological stage: stages I, II and III. They represent the body's response to increased energy metabolism. And reflect the systemic adjustment that occurs between the respiratory system, cardiac, vascular and muscle responsible for the capture, transport and use of O_2 .

The mathematical fitting of the three phases requires the identification of nine parameters, including the moments of time where occurs the change of phases I to II and II to III. The nonlinear least squares method chosen for the fit of the three phases was **Levenberg-Marquardt**.

KEYWORDS: Nonlinear Least Squares - Levenberg-Marquardt.

Capítulo 1

Introdução

A aplicação de modelos matemáticos para identificação de parâmetros em sistemas é uma ferramenta antiga, mas somente com a evolução e popularização de métodos numéricos computacionais, a partir da década de 70, seu uso se tornou mais abrangente e deixou de ser restrito ao meio acadêmico e às grandes corporações. Com a utilização de um modelo matemático, pode-se simular o comportamento de sistemas de natureza física, biológica, química, econômica, etc. Isto possibilita identificar como alterações em variáveis específicas de um sistema afetam o seu funcionamento.

Este trabalho objetiva identificar os parâmetros que caracterizam a curva de variação no tempo da taxa de absorção de oxigênio ($\dot{V}O_2(t)$) de uma pessoa durante a realização de exercício físico. A capacidade de ajuste do metabolismo à necessidade de síntese de ATP gerada pela atividade física é um dos fatores determinantes da maior ou menor tolerância ao exercício físico de uma pessoa. Quanto mais rápido ocorrer o ajuste, ou seja, quanto mais rápida for a resposta do organismo ao aumento do esforço físico, menor será o distúrbio celular. Esse ajuste é progressivo e depende da intensidade, duração, modo de exercício e nível de preparo físico da pessoa. Tem como objetivo alcançar o estado estável e retomar a homeostasia celular.

Vários estudos recentes demonstram que um elevado nível de aptidão aeróbia é necessário para uma alta resistência à prática de esportes. Porém, o inverso, ou seja, o baixo nível de aptidão aeróbia de um paciente pode estar associado a várias doenças degenerativas. Isso acontece porque a taxa de $\dot{V}O_2$ reflete o ajuste sistêmico do transporte de oxigênio e do metabolismo muscular, assim como também da integração entre os sistemas respiratório, cardíaco, vascular e muscular, responsáveis pela captação, transporte e utilização de O_2 .

O treinamento físico sistematizado gera adaptações fisiológicas e bioquímicas que capacitam um indivíduo a ter melhores resultados na prática de exercícios físicos. A natureza e magnitude dessas adaptações dependem do tipo de atividade realizada, volume, intensidade, carga genética e nível de aptidão inicial da pessoa.

Na literatura, a variação da taxa de $\dot{V}O_2$ para exercícios de intensidade pesada-severa é modelada por três funções exponenciais sequenciais. Cada uma dessas funções é identificada como uma fase fisiológica distinta: fases I, II e III. Elas representam a resposta do organismo ao aumento do metabolismo energético. Ao todo, nove parâmetros caracterizam a curva de $\dot{V}O_2$, incluindo os

dois instantes de tempo em que há a mudança da fase I para fase II, e da fase II para fase III.

Sabendo que pacientes saudáveis apresentam valores para esses nove parâmetros diferentes dos valores apresentados por pacientes que possuem doenças degenerativas, a grande motivação desse trabalho foi desenvolver um programa computacional que através do cálculo dos valores desses parâmetros para pacientes saudáveis, possa no futuro, ser utilizado como ferramenta de auxílio para o diagnóstico de várias doenças. Ou seja, os parâmetros identificados poderão ajudar determinar como o funcionamento o sistema biológico humano é afetado por alterações em variáveis específicas da curva de $\dot{V}O_2$.

A identificação de parâmetros não se resume a determinar parâmetros para um certo modelo se um sistema. O processo deve englobar também o cálculo dos erros associados aos parâmetros. Neste trabalho é mostrado como os parâmetros que caracterizam um determinado modelo matemático podem ser calculados a partir do método dos mínimos quadrados linear e não linear, utilizando conceitos da álgebra linear e de métodos numéricos.

Para isso, a primeira parte do trabalho apresenta uma revisão dos conceitos básicos de álgebra linear. Esses conceitos são utilizados nos próximos capítulos como ferramentas matemáticas muito importantes para a obtenção das soluções apresentadas pelo método dos mínimos quadrados nos problemas de otimização.

O capítulo 3 do trabalho é dedicado ao método dos mínimos quadrados linear. Ele mostra a formulação matemática aplicada na solução de problemas analíticos e geométricos, em ajuste de curvas lineares e na solução de sistemas lineares. O capítulo 4, introduz o método dos mínimos quadrados não linear. Explica a necessidade da utilização de métodos iterativos na identificação de parâmetros de curvas não lineares, e mostra os algoritmos propostos pelos métodos de **Gauss-Newton** e **Levenberg-Marquardt** para a solução de minimização dos problemas não lineares.

O programa computacional desenvolvido em *Matlab* para identificação dos nove parâmetros que caracterizam a curva de $\dot{V}O_2$ utiliza o método de **Levenberg-Marquardt**, e é composto por oito arquivos. Seu funcionamento é explicado no capítulo 6 do trabalho.

Capítulo 2

Conceitos de Álgebra Linear Utilizados no Método dos Mínimos Quadrados

Neste capítulo serão revistos alguns conceitos básicos de álgebra linear. Esses conceitos serão utilizados para embasar a teoria do método dos mínimos quadrados linear e não linear que será apresentada nos próximos capítulos.

2.1 Convenção de Notação em Álgebra Linear

Uma matriz é um conjunto de números reais representados na forma:

$$A_{m \times n} = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{bmatrix} \quad (2.1)$$

Para uma matriz de dimensão $m \times n$, m faz referência ao número de linhas de A , e n ao número de colunas. Uma matriz é dita quadrada se $m = n$, ou seja, possui o mesmo número de linhas e colunas. Um vetor em forma de coluna é uma matriz com dimensão $m \times 1$, enquanto um vetor em forma de linha é uma matriz com dimensão $1 \times n$. Uma matriz 1×1 possui apenas uma entrada, e por isso é simultaneamente um vetor em forma de linha e coluna.

Supondo a representação em coluna como padrão neste trabalho, um vetor x que pertença ao \mathbb{R}^n será formado pelos m elementos x_1, x_2, \dots, x_n e terá a forma:

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} \quad (2.2)$$

O número real localizado na i -ésima linha e na j -ésima coluna da matriz A é chamado como a entrada (i, j) de A e representado por a_{ij} . Vale ressaltar que o primeiro índice refere-se à linha e o segundo à coluna. Duas matrizes são consideradas iguais se tiverem a mesma dimensão e todas as suas entradas idênticas, ou seja: $a_{ij} = b_{ij}$.

2.2 Espaços Vetoriais

Um dos conceitos básicos em álgebra linear é o de espaço vetorial. A noção de vetores como objetos com tamanho, direção e sentido, juntamente com as operações de adição e multiplicação por números reais forma a idéia básica de um espaço vetorial.

Espaço vetorial é uma entidade que satisfaz as propriedades básicas do espaço n – dimensional Euclidiano \mathbb{R}^n . É definido como sendo o conjunto de todos os vetores reais de dimensão n .

As duas operações básicas de um espaço vetorial V são:

1. Adição: a soma de um par de vetores $x, y \in V$ resulta em um vetor $x + y \in V$.
2. Multiplicação por escalar: multiplicando um vetor x por um escalar $c \in \mathbb{R}$, obtém-se um vetor $cx \in V$.

Essas duas operações satisfazem os seguintes axiomas, para os vetores $x, y, z \in V$ e os escalares $c, d \in \mathbb{R}$:

- Adição Comutativa: $x + y = y + x$.
- Adição Associativa: $(x + y) + z = x + (y + z)$.
- Elemento Neutro da Adição: $x + 0 = x = 0 + x$.
- Elemento Inverso da Adição: Para cada $x \in V$ existe um elemento $-x \in V$ tal que $x + (-x) = 0 = (-x) + x$.
- Distributiva: $c(x + y) = (cx) + (cy)$, e $(c + d)x = (cx) + (dx)$.
- Multiplicação Associativa por escalar: $c(dx) = (cd)x$.
- Elemento Neutro da Multiplicação por escalar: para o escalar $1 \in \mathbb{R}$, tem-se $1x = x1 = x$.

2.3 Subespaços Vetoriais

Um subespaço W de um espaço vetorial V é um subconjunto $W \subset V$, de forma que W também é um espaço vetorial. Como os elementos de W também pertencem a V , as operações de adição e de multiplicação por um escalar também são operações básicas de W . Em particular, é necessário que W contenha o elemento zero de V para que o axioma de elemento neutro da adição ($x + 0 = x = 0 + x$) possa ser atendido.

Um subconjunto $W \subset V$ de um espaço vetorial só será um subespaço vetorial se e somente se:

1. Para todo $x, y \in W$ a soma $x + y \in W$.
2. Para todo $x \in W$ e para todo $c \in \mathbb{R}$, o produto escalar $cx \in W$.

Exemplos de subespaços do \mathbb{R}^3 são:

- o próprio \mathbb{R}^3 ;

- planos contendo a origem;
- retas contendo a origem e
- a origem.

2.4 Independência Linear

Sejam v_1, \dots, v_k finitos elementos do espaço vetorial V , uma soma que tenha a forma:

$$c_1 v_1 + c_2 v_2 + \dots + c_k v_k = \sum_{i=1}^k c_i v_i \quad (2.3)$$

na qual os coeficientes c_1, c_2, \dots, c_k são números reais, é definida como uma *combinação linear* dos vetores v_1, \dots, v_k . Esses vetores geram o subespaço $W \subset V$. W é formado por todas as possíveis combinações lineares de v_1, \dots, v_k .

Dessa forma, dado um subespaço vetorial W , diz-se que v_1, \dots, v_k são geradores de W se todo vetor w de W pode ser escrito como combinação linear de v_1, \dots, v_k .

1º Exemplo:

$$3v_1 + v_2 - 2v_3, \quad 8v_1 - \frac{1}{3}v_3, \quad 5v_1 + 1v_2 + 7v_3$$

são três exemplos de combinações lineares dos vetores $v_1, \dots, v_k \subset V$.

2º Exemplo:

Seja $V = \mathcal{F}(\mathbb{R})$ o espaço de todas as funções escalares $f(x)$, e c_1, c_2 e c_3 constantes arbitrárias. A combinação linear das funções $f_1(x) = 1$, $f_2(x) = x$, $f_3(x) = x^2$:

$$f(x) = c_1 f_1(x) + c_2 f_2(x) + c_3 f_3(x) = c_1 + c_2 x + c_3 x^2 \quad (2.4)$$

forma o subespaço de todos os polinômios de segunda ordem.

3º Exemplo:

O conjunto de soluções de uma equação diferencial linear homogênea é um espaço vetorial. A equação (2.5) possui duas soluções independentes $f_1(x) = e^x$ e $f_2(x) = e^{-3x}$. Qualquer outra solução pode ser gerada a partir da combinação linear entre essas duas soluções.

$$u'' + 2u' + 3u = 0 \quad (2.5)$$

$$f(x) = c_1 f_1(x) + c_2 f_2(x) = c_1 e^x + c_2 e^{-3x} \quad (2.6)$$

2.4.1 Dependência e Independência Linear

Os vetores $v_1, \dots, v_k \in V$ são chamados de linearmente dependentes caso exista um conjunto de escalares c_1, \dots, c_k , sendo não todos nulos, que:

$$c_1 v_1 + c_2 v_2 + \dots + c_k v_k = 0 \quad (2.7)$$

A condição de que c_i 's não sejam simultaneamente iguais a zero é essencial, pois se $c_1 = \dots = c_k = 0$, a combinação linear (2.7) automaticamente assume o valor zero. Vetores que não sejam linearmente dependentes são chamados de *linearmente independentes*. Um conjunto S de vetores é dito linearmente independente se nenhum dos seus elementos for combinação linear dos outros.

1º Exemplo:

Os vetores abaixo são linearmente dependentes:

$$v_1 = \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix} \quad v_2 = \begin{pmatrix} 0 \\ 3 \\ 1 \end{pmatrix} \quad v_3 = \begin{pmatrix} -1 \\ 4 \\ 3 \end{pmatrix} \quad (2.8)$$

$$v_1 - 2v_2 + v_3 = 0 \quad (2.9)$$

Porém, os dois primeiros vetores v_1 e v_2 são linearmente independentes. Para comprovar isso, supõem-se:

$$c_1 v_1 + c_2 v_2 = \begin{pmatrix} c_1 \\ 2c_1 + 3c_2 \\ -c_1 + c_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad (2.10)$$

Dessa forma é necessário que os coeficientes c_1 e c_2 satisfaçam o sistema linear:

$$\begin{aligned} c_1 &= 0 \\ 2c_1 + 3c_2 &= 0 \\ -c_1 + c_2 &= 0 \end{aligned} \quad (2.11)$$

Esse sistema possui apenas a solução trivial $c_1 = c_2 = 0$, provando dessa forma que v_1 e v_2 são linearmente independentes.

A partir desse exemplo, pode-se fazer as seguintes observações:

Sejam $v_1, \dots, v_k \in \mathbb{R}^n$, e seja $A = (v_1 \dots v_k)$ a matriz $n \times k$ correspondente:

1. Os vetores $v_1, \dots, v_k \in \mathbb{R}^n$ são linearmente dependentes se e somente se existir uma solução $c \neq 0$ para o sistema linear homogêneo $Ac = 0$.
2. Os vetores são linearmente independentes se e somente se a solução para o sistema linear homogêneo $Ac = 0$ for a solução trivial $c = 0$.

- O vetor b pode ser escrito como combinação linear dos vetores v_1, \dots, v_k se e somente se o sistema linear $Ac = b$ possuir no mínimo uma solução.
- Um conjunto de k vetores $\in \mathbb{R}^n$ é linearmente independente se e somente se a sua matriz correspondente $n \times k$ tem ordem $k \leq n$.

2.5 Bases

Considere um espaço vetorial V e uma coleção finita de vetores $v_1, v_2, \dots, v_n \in V$. Diz-se que $\beta = v_1, v_2, \dots, v_n$ é uma base de V se:

- os vetores de β geram V , ou seja, todo vetor $v \in V$ pode ser escrito como combinação linear da forma: $v = c_1v_1 + c_2v_2 + \dots + c_kv_k$;
- os vetores de β são linearmente independentes.

As seguintes propriedades sobre bases se verificam:

- Uma base de \mathbb{R}^2 sempre tem dois vetores.
- Uma base de \mathbb{R}^3 sempre tem três vetores.
- Uma base de um plano de \mathbb{R}^3 (contendo a origem) sempre tem dois vetores.
- Uma base de uma reta de \mathbb{R}^3 ou \mathbb{R}^2 (contendo a origem) sempre tem um vetor.
- Dois vetores linearmente independentes de \mathbb{R}^2 formam uma base de \mathbb{R}^2 .
- Três vetores linearmente independentes de \mathbb{R}^3 formam uma base de \mathbb{R}^3 .
- Dois vetores linearmente independentes de um plano π de \mathbb{R}^3 contendo a origem formam uma base de π .

1º Exemplo:

Os vetores:

$$\beta = \{(1\ 1\ 1)^T, (1\ 2\ 2)^T, (1\ 3\ 3)^T, (1\ 2\ 1)^T, (2\ 1\ 1)^T\} \quad (2.12)$$

geram o \mathbb{R}^3 , porém não formam uma base pois não são linearmente independentes. é possível obter uma base de \mathbb{R}^3 a partir da coleção de β , eliminando alguns vetores, como por exemplo:

$$\beta' = \{(1\ 1\ 1)^T, (1\ 2\ 2)^T, (1\ 2\ 1)^T\} \quad (2.13)$$

2º Exemplo:

Uma base padrão muito utilizada para o n -dimensional \mathbb{R}^n consiste dos n vetores:

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \quad \dots \quad e_n = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \quad (2.14)$$

Esses vetores também podem ser escritos como colunas da matriz identidade $n \times n$. Eles geram o \mathbb{R}^n , pois qualquer vetor $x \in \mathbb{R}^n$ pode ser escrito como combinação linear de e_1, e_2, \dots, e_n .

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = x_1 e_1 + x_2 e_2 + \dots + x_n e_n \quad (2.15)$$

3º Exemplo:

Uma notação padrão para a base do tri-dimensional \mathbb{R}^3 é:

$$i = e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad j = e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad k = e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \quad (2.16)$$

Dessa forma, supondo V como um espaço vetorial de dimensão n , tem-se:

1. Todo conjunto de vetores $\in V$ com mais de n elementos é linearmente dependente.
2. Nenhum conjunto de vetores com menos de n elementos gera V .
3. Um conjunto de n vetores $\in \mathbb{R}^n$ é uma base de V se e somente se geram V .
4. Um conjunto de n vetores é uma base de V se e somente se forem linearmente independentes.

Um conjunto de vetores $v_1, \dots, v_k \in V$ formará uma base de V se e somente se todo vetor $x \in V$ puder ser escrito unicamente como uma combinação linear de v_1, \dots, v_k .

$$x = c_1 v_1 + \dots + c_k v_k = \sum_{i=1}^n c_i v_i \quad (2.17)$$

Os coeficientes (c_1, \dots, c_k) são chamados de as *coordenadas* do vetor x com respeito a base de v_1, \dots, v_k . Na base padrão do \mathbb{R}^n mostrada em (2.14), esses coeficientes são chamados de *coordenadas Cartesianas*.

2.6 Subespaços Fundamentais de uma Matriz

Retornando ao estudo dos sistemas lineares de equações, representados na forma:

$$Ax = b \quad (2.18)$$

na qual, A é uma matriz $m \times n$, m é o número de equações e n é o número de parâmetros desconhecidos, ou seja, as entradas de $x \in \mathbb{R}^n$.

2.6.1 Espaço Imagem e Espaço Nulo

Existem dois importantes espaços vetoriais associados com qualquer matriz. Esses espaços tem grande importância na solução dos sistemas lineares e estão definidos abaixo.

O **espaço imagem** (*range space*) de A é o subespaço $\text{rng } A \subset \mathbb{R}^m$ gerado pelas colunas de A . O **espaço nulo** (*null space*) ou núcleo (*kernel*) de A é o subespaço $\text{ker } A \subset \mathbb{R}^n$ formado por todos os vetores que multiplicados por A resultam no vetor nulo. Dessa forma:

$$\text{ker}(A) = \{z \in \mathbb{R}^n \mid Az = 0\} \subset \mathbb{R}^n \quad (2.19)$$

Um vetor $b \in \mathbb{R}^m$ pertence ao $\text{rng } A$ se e somente se puder ser escrito como combinação linear das colunas de $A = (v_1, v_2, \dots, v_n)$.

$$b = x_1 v_1 + \dots + x_n v_n \quad (2.20)$$

O lado direito de (2.20) é igual ao produto Ax da matriz A com o vetor $x = (x_1, x_2, \dots, x_n)^T$. Consequentemente, tem-se que $b = Ax$ para algum $x \in \mathbb{R}^n$. Dessa forma:

$$\text{rng}(A) = \{Ax \mid x \in \mathbb{R}^n\} \subset \mathbb{R}^m \quad (2.21)$$

Portanto, um vetor b está contido no espaço imagem de A se e somente se o sistema linear $Ax = b$ possuir uma solução.

A partir dessa afirmação, verifica-se que se o sistema linear $Ax = b$ possuirá uma solução x^* se e somente se o b estiver contido no espaço imagem de A . Se isso ocorrer, então x será solução do sistema linear se e somente se:

$$x = x^* + z \quad (2.22)$$

onde $z \in \text{ker } A$ é um elemento qualquer do espaço nulo de A .

2.7 Produto Interno

O produto interno escalar no espaço vetorial real V ocorre entre os vetores $v, w \in \mathbb{R}^n$ e produz um número real \mathbb{R} . O produto escalar possui três axiomas para todos $u, v, w \in V$ e $c, d \in \mathbb{R}$.

1. Distributiva

$$\begin{aligned} \langle cu + dv; w \rangle &= c \langle u; w \rangle + d \langle v; w \rangle \\ \langle u; cv + dw \rangle &= c \langle u; v \rangle + d \langle u; w \rangle \end{aligned} \quad (2.23)$$

2. Simetria

$$\langle v; w \rangle = \langle w; v \rangle \quad (2.24)$$

3. Positividade

$$\begin{aligned} \langle v; v \rangle &\geq 0 \quad \text{para } v \neq 0 \\ \langle 0; 0 \rangle &= 0 \end{aligned} \quad (2.25)$$

O produto interno padrão do espaço euclidiano é o produto escalar.

$$\langle v; w \rangle = v \cdot w = v_1 w_1 + v_2 w_2 + \dots + v_n w_n = \sum_{i=1}^n v_i w_i \quad (2.26)$$

O produto escalar ocorre entre os vetores $v = (v_1, v_2, \dots, v_n)^T$, $w = (w_1, w_2, \dots, w_n)^T$ escritos no *Espaço Euclidiano* \mathbb{R}^n . Uma importante observação é que o produto (2.26) pode ser escrito como um produto de matrizes, entre um vetor em linha $v \in \mathbb{R}^n$ e um vetor em coluna $w \in \mathbb{R}^n$.

$$\langle v; w \rangle = v^T \cdot w = (v_1 \ v_2 \ \dots \ v_n) \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix} \quad (2.27)$$

O produto interno de um vetor com ele próprio é:

$$\langle v; v \rangle = v_1^2 + v_2^2 + \dots + v_n^2 \quad (2.28)$$

ou seja, é a soma dos quadrados de suas entradas e por isso igual ao quadrado de seu comprimento. Assim sendo, a *norma euclidiana* ou *comprimento* de um vetor de um vetor v , representada por $\|v\|$, é a raiz quadrada do produto interno $v \cdot v$, como mostrado abaixo:

$$\|v\| = \sqrt{v \cdot v} = \sqrt{v_1^2 + v_2^2 + \dots + v_n^2} \quad (2.29)$$

Verifica-se que em (2.29) cada um dos elementos assume valor positivo. Dessa forma, o comprimento de um vetor é também positivo, $\|v\| \geq 0$, e somente tem valor zero para o vetor 0. A norma de vetores do \mathbb{R}^2 e o \mathbb{R}^3 pode ser representada como na figura (2.1).

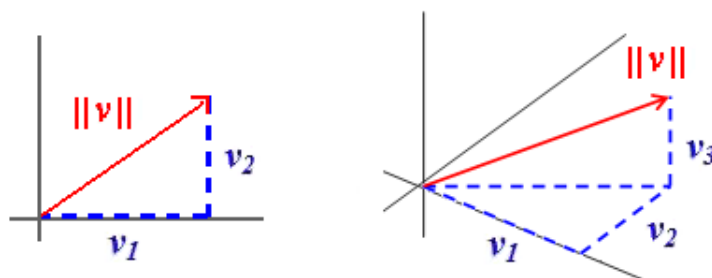


Figura 2.1: Norma Euclidiana para o \mathbb{R}^2 e o \mathbb{R}^3

1º Exemplo:

Um exemplo muito interessante é o produto interno entre funções. Dado um intervalo fechado $[a, b] \subset \mathbb{R}$ e o espaço vetorial $C^0 = C^0[a, b]$ formado por todas as funções contínuas $f: [a, b] \rightarrow \mathbb{R}$, a integral

$$\langle f; g \rangle = \int_a^b f(x)g(x) dx \quad (2.30)$$

define o produto interno no espaço vetorial C^0 . Pode-se também definir a norma de uma função f a partir da equação (2.29) para o intervalo $[a, b]$.

$$\|f\| = \sqrt{\int_a^b f(x)^2 dx} \quad (2.31)$$

2º Exemplo:

Outro exemplo interessante é o produto interno ponderado entre dois vetores $v, w \in \mathbb{R}^n$. Cada uma das parcelas $v_i w_i$ é multiplicada por uma constante c_i . Entre dois vetores do \mathbb{R}^2 , por exemplo, atribuindo os valores de 2 e 5 para c_1 e c_2 respectivamente, tem-se:

$$\begin{aligned} \langle v; w \rangle &= c_1 v_1 w_1 + c_2 v_2 w_2 \\ \langle v; w \rangle &= 2 v_1 w_1 + 5 v_2 w_2 \end{aligned} \quad (2.32)$$

Nesta seção uma importante observação a ser feita é que existem normas vetoriais que não são provenientes do produto interno. Como exemplo tem-se a *norma-1* $\|v\|_1$, e a norma infinita $\|v\|_\infty$ de um vetor $v = (v_1 v_2 \dots v_n)^T$. Essas normas são definidas respectivamente como:

$$\|v\|_1 = |v_1| + \dots + |v_n| \quad (2.33)$$

$$\|v\|_\infty = \sup \{|v_1|, \dots, |v_n|\} \quad (2.34)$$

A norma euclidiana, a norma-1 e a norma infinita são apenas três representações da forma mais geral p-norma.

$$\|v\|_p = \sqrt[p]{\sum_{i=1}^n |v_i|^p} \quad (2.35)$$

O valor de $\|v\|_p$ define uma norma para qualquer $1 \leq p < \infty$. A norma infinita é o caso limite da norma-p, e ocorre quando $p \rightarrow \infty$. A norma euclidiana mostrada na equação (2.29) equivale a norma-2 (quando $p = 2$).

2.8 Matrizes Positivas Definidas

2.8.1 Matrizes Positivas Definidas

Uma classe de matrizes extremamente importante é a classe das matrizes positivas definidas. Essas matrizes têm um papel fundamental na solução de problemas das mais variadas áreas, como, por exemplo, em problemas de mecânica, circuitos elétricos e equações diferenciais.

Seja $|\langle v; w \rangle|$ o produto interno entre os vetores $x = (x_1 x_2 \dots x_n)^T$ e $y = (y_1 y_2 \dots y_n)^T$, $\in \mathbb{R}^n$, escrevendo os vetores na base padrão vetorial, tem-se:

$$\begin{aligned} x &= x_1 e_1 + \dots + x_n e_n = \sum_{i=1}^n x_i e_i \\ y &= y_1 e_1 + \dots + y_n e_n = \sum_{j=1}^n y_j e_j \end{aligned} \quad (2.36)$$

Analisando os três axiomas do produto interno citados anteriormente, pode-se utilizar o axioma da distributiva para expandir:

$$\langle x; y \rangle = \left\langle \sum_{i=1}^n x_i e_i; \sum_{j=1}^n y_j e_j \right\rangle = \sum_{i,j=1}^n x_i y_j \langle e_i; e_j \rangle \quad (2.37)$$

Pode-se assim escrever:

$$\langle x; y \rangle = \sum_{i,j=1}^n k_{ij} x_i y_j = x^T K y \quad (2.38)$$

onde K é uma matriz $n \times n$ do produto interno da base vetorial, com as entradas:

$$k_{ij} = \langle e_i; e_j \rangle, \quad i, j = 1, \dots, n \quad (2.39)$$

Os outros dois axiomas do produto interno impõem algumas condições na matriz do produto interno K . O axioma de simetria garante que:

$$k_{ij} = \langle e_i; e_j \rangle = \langle e_j; e_i \rangle = k_{ji}, \quad i, j = 1, \dots, n \quad (2.40)$$

Consequentemente a matriz K do produto interno é simétrica:

$$K = K^T \quad (2.41)$$

A simetria de K assegura que:

$$\langle x; y \rangle = x^T K y = (x^T K y)^T = y^T K^T x = y^T K x = \langle y; x \rangle \quad (2.42)$$

O último axioma do produto interno é a positividade, por isso:

$$\|x\|^2 = \langle x; x \rangle = x^T K x = \sum_{i,j=1}^n k_{ij} x_i x_j \geq 0, \quad \text{para todo } x \in \mathbb{R}^n \quad (2.43)$$

A partir dessa equação define-se que uma matriz K $n \times n$ é positiva definida se for simétrica, $K = K^T$, e satisfizer a condição:

$$x^T K x > 0, \text{ para todo } 0 \neq x \in \mathbb{R}^n \quad (2.44)$$

Porém, deve-se ter cuidado porque uma matriz ser positiva definida não tem obrigatoriamente todas as suas entradas com valores positivos. Existem matrizes positivas definidas que têm entradas com valores negativos. O contrário também acontece, matrizes simétricas e com todas as entradas positivas nem sempre são positivas definidas.

Os conceitos apresentados nesse capítulo resultaram na seguinte caracterização do produto interno de um espaço vetorial com dimensão finita: todo produto interno do \mathbb{R}^n pode ser representado por

$$\langle x; y \rangle = x^T K y, \quad \text{para todo } x, y \in \mathbb{R}^n \quad (2.45)$$

Dada uma matriz K simétrica, a expressão

$$q(x) = x^T K x = \sum_{i,j=1}^n k_{ij} x_i x_j \quad (2.46)$$

é conhecida como uma forma quadrada no \mathbb{R}^n . Essa forma quadrada é chamada de positiva definida se:

$$q(x) > 0, \quad \text{para todo } 0 \neq x \in \mathbb{R}^n \quad (2.47)$$

Dessa forma, uma forma quadrada é positiva definida se e somente se sua matriz coeficiente for positiva definida.

Exemplo:

Dada a forma quadrática:

$$q(x, y, z) = x^2 + 4xy + 6y^2 - 2xz + 9z^2 \quad (2.48)$$

que possui três variáveis, a matriz coeficiente correspondente é:

$$K = \begin{pmatrix} 1 & 2 & -1 \\ 2 & 6 & 0 \\ -1 & 0 & 9 \end{pmatrix} \quad (2.49)$$

pois:

$$q(x, y, z) = (x \ y \ z) \begin{pmatrix} 1 & 2 & -1 \\ 2 & 6 & 0 \\ -1 & 0 & 9 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \quad (2.50)$$

2.8.2 Matrizes Positivas Semi-Definidas

Uma forma quadrática associada com uma matriz coeficiente simétrica é definida como positiva semi-definida se:

$$q(x) = x^T K x \geq 0, \quad \text{para todo } x \in \mathbb{R}^n \quad (2.51)$$

As matrizes positivas semi-definidas podem possuir direções nulas, o que significa vetores $z \neq 0$ para os quais $q(z) = z^T K z = 0$. Qualquer vetor pertencente ao espaço nulo de K ($z \in \ker(K)$) define uma direção nula. As matrizes positivas definidas possuem $\ker(K) = \{0\}$.

Outras definições importantes são as de matrizes negativas definidas e a de matrizes negativas semi-definidas.

Uma forma quadrática $q(x) = x^T K x$ associada a uma matriz simétrica K é chamada de negativa semi-definida se $q(x) \leq 0$ para todo x , e é chamada de negativa definida se $q(x) < 0$ para todo $x \neq 0$. Uma forma quadrática é indefinida se for simultaneamente positiva e negativa definida, ou seja, caso existam alguns vetores x_+ para os quais $q(x_+) > 0$ e outros x_- para os quais $q(x_-) < 0$.

2.8.3 Matrizes de Gram

Sejam V o espaço vetorial gerado por v_1, \dots, v_n , pode-se definir uma matriz $n \times n$ cujas entradas são produtos internos entre v_1, \dots, v_n chamada *Matriz de Gram*.

$$K = \begin{pmatrix} \langle v_1 ; v_1 \rangle & \langle v_1 ; v_2 \rangle & \cdots & \langle v_1 ; v_n \rangle \\ \langle v_2 ; v_1 \rangle & \langle v_2 ; v_2 \rangle & \cdots & \langle v_2 ; v_n \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle v_n ; v_1 \rangle & \langle v_n ; v_2 \rangle & \cdots & \langle v_n ; v_n \rangle \end{pmatrix} \quad (2.52)$$

O axioma de simetria do produto interno (2.24) garante que a matriz de Gram é simétrica.

$$k_{ij} = \langle v_i ; v_j \rangle = \langle v_j ; v_i \rangle = k_{ji} \quad \Leftrightarrow \quad K = K^T \quad (2.53)$$

Outra característica da matriz de Gram é que ela é positiva semi-definida. Será positiva definida se e somente se os elementos $v_1, \dots, v_n \in V$ forem linearmente independentes. A demonstração é feita através da forma quadrática associada a uma matriz K .

$$q(x) = x^T K x = \sum_{i,j=1}^n k_{ij} x_i x_j \quad (2.54)$$

Substituindo os valores de (2.53), tem-se:

$$q(x) = \sum_{i,j=1}^n \langle v_i ; v_j \rangle x_i x_j \quad (2.55)$$

$$q(x) = \left\langle \sum_{i=1}^n x_i v_i ; \sum_{j=1}^n x_j v_j \right\rangle = \langle v ; v \rangle = \|v\|^2 \geq 0 \quad (2.56)$$

onde:

$$v = x_1 v_1 + \cdots + x_n v_n \quad (2.57)$$

Sendo v_1, \dots, v_n os geradores de V , comprova-se que K é positiva semi-definida e consequentemente $q(x) = \|v\|^2 \geq 0$ enquanto $v \neq 0$. Se os vetores v_1, \dots, v_n forem linearmente independentes, $v = 0$ se e somente se $x_1 = \cdots = x_n = 0$. Nesse caso, $q(x)$ e K são positivas definidas.

1º Exemplo:

Considerando os vetores $v_1 = \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}$ e $v_2 = \begin{pmatrix} 3 \\ 0 \\ 6 \end{pmatrix}$ pertencentes ao \mathbb{R}^3 , para o produto escalar padrão do espaço Euclidiano, a matriz Gram associada é:

$$K = \begin{pmatrix} v_1 \cdot v_1 & v_1 \cdot v_2 \\ v_2 \cdot v_1 & v_2 \cdot v_2 \end{pmatrix} = \begin{pmatrix} 6 & -3 \\ -3 & 45 \end{pmatrix} \quad (2.58)$$

A forma quadrática associada a matriz K é positiva definida para todos $(x_1, x_2) \neq 0$.

$$q(x_1, x_2) = 6x_1^2 - 6x_1x_2 + 45x_2^2 > 0 \quad (2.59)$$

Para o produto interno ponderado $\langle x; y \rangle = 3x_1y_1 + 2x_2y_2 + 5x_3y_3$ a matriz Gram positiva definida associada é:

$$K = \begin{pmatrix} \langle v_1; v_1 \rangle & \langle v_1; v_2 \rangle \\ \langle v_2; v_1 \rangle & \langle v_2; v_2 \rangle \end{pmatrix} = \begin{pmatrix} 16 & -21 \\ -21 & 207 \end{pmatrix} \quad (2.60)$$

No caso do produto escalar, a construção da matriz Gram pode ser feita de maneira fácil, como mostrado a seguir.

Dado os vetores $v_1, \dots, v_n \in \mathbb{R}^m$, pode-se formar a matriz $m \times n$ $A = (v_1 \ v_2 \ \cdots \ v_n)$. Devido a relação (2.27) entre o produto escalar e a multiplicação entre as linhas e colunas dos vetores, a entrada (i, j) da matriz K é:

$$k_{ij} = v_i \cdot v_j = v_i^T v_j \quad (2.61)$$

ou seja, é a multiplicação entre a i -ésima e a j -ésima linha e a coluna de A . Dessa forma, a matriz Gram é definida como sendo o produto entre a transposta de A e a própria matriz A :

$$K = A^T A \quad (2.62)$$

Para o último exemplo mostrado tem-se:

$$A = \begin{pmatrix} 1 & 3 \\ 2 & 0 \\ -1 & 6 \end{pmatrix} \quad (2.63)$$

$$K = A^T A = \begin{pmatrix} 1 & 2 & -1 \\ 3 & 0 & 6 \end{pmatrix} \begin{pmatrix} 1 & 3 \\ 2 & 0 \\ -1 & 6 \end{pmatrix} = \begin{pmatrix} 6 & -3 \\ -3 & 45 \end{pmatrix} \quad (2.64)$$

Sabendo-se que os geradores de V , ou seja, os vetores v_1, \dots, v_n , formam uma matriz Gram positiva definida se e somente se forem linearmente independentes, verifica-se que a matriz Gram $K = A^T A$ será positiva definida se e somente se as colunas de A foram linearmente independentes. Dessa forma, dada uma matriz A , sabe-se que:

1. A matriz Gram $m \times n$ associada $K = A^T A$ é positiva definida.
2. A matriz A tem colunas linearmente independentes.
3. $\text{rank}\{A\} = n \leq m$.
4. $\text{ker}\{A\} = \{0\}$.

Capítulo 3

Método dos Mínimos Quadrados

Este capítulo introduz e explica a metodologia de solução de problemas de minimização. A solução desses problemas tem um enorme gama de aplicações direta, incluindo o ajuste de curvas pelo método dos mínimos quadrados, problemas de interpolação e solução de sistemas de equações lineares.

3.1 Exemplos de Problemas de Minimização

O problema de minimização será introduzido a partir de dois importantes exemplos: um analítico e outro geométrico.

Exemplo de Solução de um Sistema de Equações

Suponha que deseja-se resolver o seguinte sistema de equações:

$$f_1(x) = 0, \quad f_2(x) = 0, \quad \dots \quad f_m(x) = 0 \quad (3.1)$$

com $x = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$. Esse sistema pode ser convertido em um problema de minimização. Para isso basta definir:

$$p(x) = [f_1(x)]^2 + [f_2(x)]^2 + \dots + [f_m(x)]^2 = \|f(x)\|^2 \quad (3.2)$$

onde $\|\cdot\|$ representa a norma euclidiana no \mathbb{R}^m . Claramente, $p(x) \geq 0$ para todo x , e $p(x^*) = 0$ se cada uma das suas parcelas for zero, e por isso, x^* é uma solução de (3.1). Além disso, o mínimo valor possível para $p(x)$ é zero, e como consequência, é alcançado se e somente se $x = x^*$ for solução do sistema.

O caso mais importante de solução de sistemas é quando temos o sistema linear:

$$Ax = b \quad (3.3)$$

constituído de n incógnitas, e m equações. Neste caso, as soluções podem ser obtidas minimizando-se a função:

$$p(x) = \|Ax - b\|^2 \quad (3.4)$$

A conversão de $Ax = b$ em um problema de minimização aparentemente não representa um grande avanço, pois $Ax = b$ pode ser resolvido por eliminação gaussiana. Porém, ao contrário do que parece, é um artifício simples gerador de grandes alterações na forma de resolver o problema.

Suponha que o sistema $Ax = b$ não possui uma solução, ou seja, o vetor b não pertence ao espaço imagem da matriz A ($b \notin \text{rng}(A)$). Isso é muito comum quando se tem $m > n$, ou seja, um número maior de incógnitas do que variáveis, ou em problemas de ajuste de curvas, nos quais os pontos medidos estão sujeitos à erros experimentais e conseqüentemente não assumem os valores exatos esperados pelo modelo matemático que descreve o fenômeno.

Apesar desses problemas não possuírem uma solução, ainda deseja-se descobrir qual é o vetor x^* que está o mais próximo possível da solução do sistema. Nesse caso, a conversão do problema de solução de sistemas em um problema de minimização é muito vantajosa, pois uma das opções para descobrir o melhor x^* é através da minimização do módulo do vetor residual $\|r\| = \|Ax - b\|$.

No caso do sistema possuir uma solução exata, ou seja $Ax^* = b$, x^* é o vetor que resulta no resíduo $\|r\|$ mínimo absoluto, ou seja $p(x^*) = 0$. O cálculo de x^* para minimização do valor de (3.4) feito pelo método dos mínimos quadrados.

Exemplo Geométrico

O próximo problema de minimização a ser mostrado provém da geometria. Dado um ponto $b \in \mathbb{R}^m$ e o subconjunto $V \subset \mathbb{R}^m$, deseja-se achar o ponto $v^* \in V$ que é mais próximo possível de b . Em outras palavras, deseja-se minimizar a distância $d(b, v) = \|v - b\|$ dentro de todas as possibilidades de $v \in V$.

O caso mais simples desse tipo de problema ocorre quando V é um subespaço do \mathbb{R}^m . Nesse caso, o problema pode ser convertido em um problema de mínimos quadrados. Sejam v_1, \dots, v_n uma base de V , qualquer elemento $v \in V$ pode ser escrito como uma combinação linear dos vetores v_1, \dots, v_n . A partir da equação mostrada em (3.4) de multiplicação de matrizes, pode-se escrever um elemento v , na forma:

$$v = x_1 v_1 + \dots + x_n v_n = Ax \tag{3.5}$$

onde $A = (v_1 \ v_2 \ \dots \ v_n)$ é a matriz $m \times n$ formada pelos vetores geradores da base. Conseqüentemente, o ponto mais próximo de b existente em V pode ser achado pela minimização:

$$\|v - b\|^2 = \|Ax - b\|^2 \tag{3.6}$$

dentro de todas as possibilidades $x \in \mathbb{R}^n$. Observa-se que a equação (3.6) é a mesma mostrada em (3.4) no exemplo anterior. Conseqüentemente, sendo x^* a solução do método dos mínimos quadrados para o sistema $Ax = b$, então $v^* = Ax^*$ será o ponto mais próximo de b que pertence a V .

Os dois exemplos de problemas de minimização mostrados nesta seção tiveram como proposta de solução o uso do método dos mínimos quadrados, que será explicado nas próximas seções.

3.2 Minimização de Funções Quadráticas

O mais simples problema de minimização é o problema de minimização de funções quadráticas. Vale ressaltar que funções lineares $p(x) = bx + c$ não possuem um valor mínimo. Nesta seção,

será mostrado a resolução dos problemas de minimização de funções quadráticas com n variáveis através de técnicas de álgebra linear.

As ideias apresentadas na solução do problema de minimização funções quadráticas serão o ponto de partida para a demonstração das soluções dos dois exemplos introduzidos no início deste capítulo.

O exemplo mais simples é o caso da minimização de funções quadráticas escalares:

$$p(x) = ax^2 + 2bx + c \quad (3.7)$$

Para esta equação, sabe-se que quando $a > 0$ o gráfico assume a forma de uma parábola com concavidade voltada para cima. Nesse caso, existe um único valor de mínimo para a função. Porém, quando $a < 0$ o gráfico possui concavidade voltada para baixo e conseqüentemente não há um valor de mínimo.

O ponto de mínimo existente quando $a > 0$ é calculado igualando a expressão da derivada de $p(x)$ em relação a variável x a zero.

$$p'(x) = 2ax + 2b = 0 \quad (3.8)$$

O valor de mínimo de $p(x)$ ocorrerá em x^* .

$$x^* = -\frac{b}{a}, \quad \text{onde} \quad p(x^*) = c - \frac{b^2}{a} \quad (3.9)$$

Um problema mais complicado que a minimização de funções escalares é a minimização de funções de várias variáveis. Dada a função quadrática:

$$p(x) = p(x_1, \dots, x_n) = \sum_{i,j}^n k_{ij} x_i x_j - 2 \sum_i^n f_i x_i + c \quad (3.10)$$

com n variáveis $x = (x_1, \dots, x_n)^T \in \mathbb{R}^n$, e coeficientes $k_{i,j}$, f_i reais, pode-se assumir que os coeficientes dos termos quadráticos são simétricos, ou seja, $k_{i,j} = k_{j,i}$. Reescrevendo a equação (3.10) na notação matricial, tem-se:

$$p(x) = x^T Kx - 2x^T f + c \quad (3.11)$$

Nessa equação, $K = (K_{i,j})$ é uma matriz $n \times n$, f é um vetor de constantes e c uma constante escalar. No caso de minimização de funções quadráticas escalares (3.7) impôs-se a condição de que o coeficiente a era positivo para obter-se um único valor de mínimo da função. Na minimização de funções de várias variáveis é necessário fazer-se uma imposição correspondente a essa feita anteriormente. Assume-se que a matriz K é positiva definida.

Teorema: Sendo $K > 0$ uma matriz positiva definida, a forma quadrática (3.11) terá um único valor de mínimo. Esse valor de mínimo ocorrerá em x^* , que é a solução do sistema linear:

$$Kx = f \quad (3.12)$$

$$x^* = K^{-1}f \quad (3.13)$$

Dessa forma, o valor mínimo de $p(x)$ pode ser expresso pelas seguintes equações:

$$P(x^*) = p(K^{-1}f) = c - f^T K^{-1}f = c - f^T x^* = c - (x^*)^T Kx^* \quad (3.14)$$

A prova desse teorema pode ser feita assumindo-se que $x^* = K^{-1}f$ é uma solução única para (3.12). Assim, para qualquer $x \in \mathbb{R}^n$ pode-se escrever:

$$\begin{aligned} p(x) &= x^T Kx + 2x^T f + c = x^T Kx + 2x^T Kx^* + c \\ &= (x - x^*)^T K(x - x^*) + [c - (x^*)^T Kx^*] \end{aligned} \quad (3.15)$$

Utilizou-se a simetria de $K = K^T$ para identificar que $x^T Kx^* = (x^*)^T Kx$. Fazendo uma simples análise de (3.15), verifica-se que o segundo termo na final da equação não depende da variável x , e que o primeiro termo está na forma $y^T Ky$, onde $y = x - x^*$. Assumindo que K é uma matriz positiva definida, ou seja, $y^T Ky \geq 0$ e terá valor nulo se e somente se $y = x - x^* = 0$, verifica-se que o mínimo de $p(x)$ ocorre quando $x = x^*$. E dessa forma, a expressão (3.13) foi comprovada.

Pode-se assim afirmar que sendo $K > 0$ uma matriz positiva definida, a forma quadrática $p(x) = x^T Kx + 2x^T f + c$ possui um único mínimo global que ocorre em x^* . Se a matriz $K \geq 0$ for positiva semi-definida, e $f \in \text{rng}K$, então toda solução para $Kx^* = f$ será um mínimo global de $p(x)$. No caso de matriz positiva semi-definida, o mínimo não é único pois $p(x^* + z) = p(x^*)$ para qualquer vetor $z \in \ker K$.

3.3 Solução de um Problema Geométrico

Nesta seção será mostrado a solução do problema geométrico de minimização que foi introduzido na primeira seção deste capítulo.

O problema consiste em dado um subespaço $V \subset \mathbb{R}^m$, e um elemento $b \in \mathbb{R}^m$, achar $v^* \in V$ dentro de todas as possibilidades de $v \in V$ que minimiza a distância $\|v - b\|$.

É importante ressaltar que $b \notin V$. Caso b pertencesse ao subespaço V , a menor distância seria automaticamente zero, pois o ponto escolhido seria o próprio b , de forma que $x^* = b$.

Dessa forma, deseja-se minimizar a distância:

$$\|v - b\|^2 = \|v\|^2 - 2\langle v; b \rangle + \|b\|^2 \quad (3.16)$$

dentro de todas as possibilidades de v pertencentes a $V \subset \mathbb{R}^m$. Assumindo que uma base de V v_1, \dots, v_n com dimensão n é conhecida, pode-se escrever todo vetor $v \in V$ como uma combinação linear dos vetores v_1, \dots, v_n .

$$v = x_1 v_1 + \dots + x_n v_n \quad (3.17)$$

Dessa forma:

$$\|v\|^2 = \langle v; v \rangle = \langle x_1 v_1 + \dots + x_n v_n; x_1 v_1 + \dots + x_n v_n \rangle = \sum_{i,j=1}^n x_i x_j \langle v_i; v_j \rangle \quad (3.18)$$

e conseqüentemente:

$$\|v\|^2 = \sum_{i,j=1}^n k_{ij} x_i x_j = x^T K x \quad (3.19)$$

onde K é a matriz *Gram* simétrica $n \times n$. Cada uma de suas entradas (i, j) é o produto interno entre os vetores que definem a base do subespaço V .

$$k_{ij} = \langle v_i ; v_j \rangle \quad (3.20)$$

De forma similar, tem-se que:

$$\langle v ; b \rangle = \langle x_1 v_1 + \dots + x_n v_n ; b \rangle = \sum_{i=1}^n x_i \langle v_i ; b \rangle \quad (3.21)$$

$$\langle v ; b \rangle = \sum_{i=1}^n x_i f_i = x^T f \quad (3.22)$$

onde f é um vetor cuja i -ésima entrada é o produto interno:

$$f_i = \langle v_i ; b \rangle \quad (3.23)$$

Dessa forma, a equação (3.16) pode ser escrita como uma função quadrática, com a forma:

$$p(x) = x^T K x + 2x^T f + c = \sum_{i,j=1}^n k_{ij} x_i x_j - 2 \sum_{i=1}^n f_i x_i + c \quad (3.24)$$

com K e f definidos em (3.20) e (3.23) respectivamente, e tendo $c = \|b\|^2$.

Assumindo que os vetores v_1, \dots, v_n da base de V são linearmente independentes, a matriz Gram $K = A^T A$ é positiva definida, e conseqüentemente a forma quadrática $p(x) = x^T K x + 2x^T f + c$ possui um único mínimo global que ocorre em x^* .

Teorema: Seja $V \subset \mathbb{R}^m$ um subespaço vetorial formado pelos vetores v_1, \dots, v_n . Dado um ponto b , o ponto $v^* = x^* v_1 + \dots + x_n^* v_n \in V$ mais próximo de b é obtido pela solução $x^* = K^{-1} f$ do sistema linear:

$$K x = f \quad (3.25)$$

onde K e f são definidos em (3.20) e (3.23) respectivamente. A distância do ponto ao subespaço será:

$$\|v^* - b\| = \sqrt{\|b\|^2 - f^T x^*} \quad (3.26)$$

Utilizando o produto interno padrão do espaço Euclidiano e a norma do \mathbb{R}^n para determinar-se a distância $v^* - b$, as entradas da matriz Gram K e do vetor f podem ser representadas pelo produto escalar:

$$k_{ij} = v_i \cdot v_j = v_i^T v_j \quad f_i = v_i \cdot b = v_i^T b \quad (3.27)$$

Sendo $A = (v_1 \ v_2 \ \dots \ v_n)$ a matriz $n \times m$ formada pelos vetores da base de V , tem-se que:

$$K = A^T C A \qquad f = A^T C b \qquad c = \|b\|^2 \qquad (3.28)$$

Sabendo-se que:

$$v = x_1 v_1 + \dots + x_n v_n = Ax \qquad (3.29)$$

verifica-se que:

$$\begin{aligned} \|v - b\|^2 &= \|Ax - b\|^2 = (Ax - b)^T (Ax - b) = (x^T A^T - b^T)(Ax - b) \\ &= x^T A^T A x - 2x^T A^T b + b^T b = x^T K x + 2x^T f + c \end{aligned} \qquad (3.30)$$

3.4 Mínimos Quadrados Aplicados à Solução de Sistema de Equações Lineares

Define-se que a solução do sistema de equações lineares

$$Ax = b \qquad (3.31)$$

pelo método dos mínimos quadrados é o vetor $x^* \in \mathbb{R}^n$ que minimiza a norma euclidiana $\|Ax - b\|$.

Se o sistema $Ax = b$ possuir uma solução, poderá ser resolvido por eliminação gaussiana. Porém, se não possuir uma solução exata, ou seja, o vetor b não pertence ao espaço imagem da matriz A ($b \notin \text{rng}(A)$), o método dos mínimos quadrados permite calcular um vetor x^* que é o mais próximo possível da solução do sistema.

Como objetiva-se determinar uma solução única para o sistema, é necessário que $\ker A = \{0\}$ ou de forma equivalente, que as colunas da matriz A sejam linearmente independentes. Identificando V como sendo o espaço imagem formado pelas colunas da matriz A , ou seja, $V = \text{rng} A \subset \mathbb{R}^m$, tem-se que as colunas de A formam uma base para V .

Dessa forma, como qualquer elemento do espaço imagem de A pode ser escrito como $v = Ax$, minimizar $\|Ax - b\|$ é o mesmo que minimizar a distância $\|v - b\|$ entre o ponto b e o subespaço V . A solução x^* obtida pelo método dos mínimos quadrados fornecerá o valor de $v^* = Ax^*$ que pertence ao subespaço $V = \text{rng} A$.

Feita a correlação entre o problema geométrico de minimização da distância de b a V e o problema da determinação de uma solução para $Ax = b$, utiliza-se as relações explicitadas em (3.28) para escreve-se o sistema (3.25) e a mínima distância do ponto ao subespaço.

Teorema: A solução pelo método dos mínimos quadrados para o sistema $Ax = b$, assumindo que $\ker A = \{0\}$ e que $K = A^T A$ e $f = A^T b$ é a solução do sistema:

$$Kx = f \qquad \text{ou} \qquad (A^T A)x = A^T b \qquad (3.32)$$

Dessa forma:

$$x^* = (A^T A)^{-1} A^T b \quad (3.33)$$

O erro associado ao método dos mínimos quadrados é:

$$\|Ax^* - b\|^2 = \|b\|^2 - f^T x^* = \|b\|^2 - b^T A (A^T A)^{-1} A^T b \quad (3.34)$$

É importante observar que a equação (3.32) é obtida a partir da multiplicação de A^T em ambos os lados do sistema original $Ax = b$. Se a matriz A for quadrada e invertível, então $(A^T A)^{-1} = A^{-1} (A^T)^{-1}$ e dessa forma, a equação (3.33) é reduzida a $x^* = A^{-1} b$. Além disso, os dois termos da expressão (3.34) se cancelam, fazendo com que o erro associado ao método dos mínimos quadrados seja zero.

1º Exemplo:

Dado o sistema linear de 5 equações e três incógnitas abaixo:

$$\begin{array}{rccccrcr} x_1 & + & 2x_2 & & & = & 1 \\ 3x_1 & - & x_2 & + & x_3 & = & 0 \\ -x_1 & + & 2x_2 & + & x_3 & = & -1 \\ x_1 & - & x_2 & - & 2x_3 & = & 2 \\ 2x_1 & + & x_2 & - & x_3 & = & 2 \end{array} \quad (3.35)$$

define-se que a matriz A e o vetor b , necessários para se escrever o sistema na forma matricial são:

$$A = \begin{pmatrix} 1 & 2 & 0 \\ 3 & 1 & 1 \\ -1 & 2 & 1 \\ 1 & 1 & 2 \\ 2 & 1 & 1 \end{pmatrix} \quad b = \begin{pmatrix} 1 \\ 0 \\ -1 \\ 2 \\ 2 \end{pmatrix}. \quad (3.36)$$

A expressão (3.33) permite calcular solução pelo método dos mínimos quadrados. Para utilizá-la, calcula-se:

$$K = A^T A = \begin{pmatrix} 16 & -2 & -2 \\ -2 & 11 & 2 \\ -2 & 2 & 7 \end{pmatrix} \quad b = A^T b = \begin{pmatrix} 8 \\ 0 \\ 7 \end{pmatrix} \quad (3.37)$$

A solução do sistema $Kx = f$ é:

$$x^* = K^{-1} b \approx \begin{pmatrix} 0,4119 \\ 0,2482 \\ 0,9532 \end{pmatrix} \quad (3.38)$$

O erro associado à solução do método dos mínimos quadrados é:

$$\|b - Ax^*\| = \|(-0,0917, 0,0342, 0,1310, 0,0701, 0,0252)^T\| = 0,1799 \quad (3.39)$$

3.5 Mínimos Quadrados Aplicados ao Ajuste de Curvas

Uma das aplicações mais importantes do método dos mínimos quadrados é o ajuste de curvas. Ele permite determinar os parâmetros de um modelo matemático através de valores dessa função em pontos discretos de um intervalo. Para reduzir-se a influência dos erros presentes nos valores da variável avaliada, costuma-se utilizar um número de pontos superior ao número de parâmetros desconhecidos do modelo matemático. Supondo que foi feito um experimento para medir uma variável dependente do tempo. A cada instante t_i foi medido um valor para y_i , resultando em m pontos experimentais.

$$(t_1, y_1), \quad (t_2, y_2), \quad \dots \quad (t_m, y_m). \quad (3.40)$$

Supondo que o modelo matemático dos pontos experimentais seja uma reta, como mostrado em (3.41), objetiva-se achar valores para α e β que resultem no melhor ajuste possível para os pontos experimentais (3.40).

$$y(t) = \alpha + \beta t \quad (3.41)$$

Como os pontos medidos estão sujeitos à erros experimentais e conseqüentemente não assumem os valores exatos esperados pelo modelo matemático que descreve o fenômeno, há um erro associado a cada valor de y_i medido no instante t_i , como mostra o gráfico da figura (3.1).

$$e_i = y_i - (\alpha + \beta t_i), \quad i = 1, \dots, m. \quad (3.42)$$

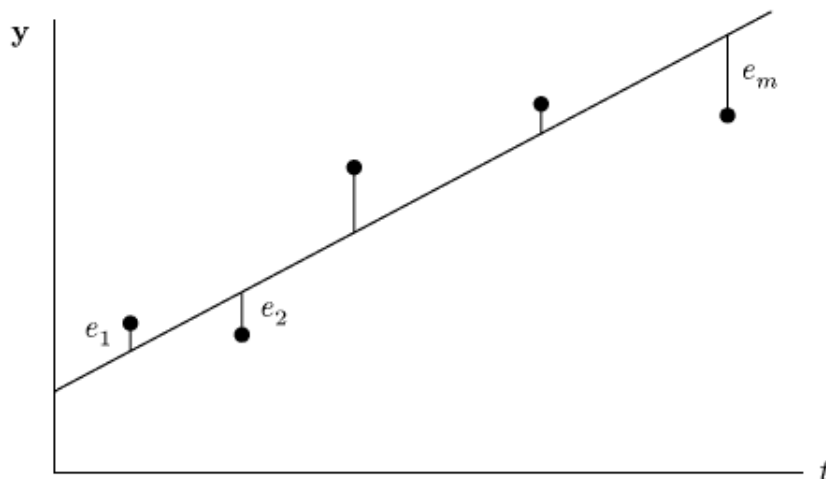


Figura 3.1: Ajuste de $y(t) = \alpha + \beta t$ pelo método dos mínimos quadrados

Escrevendo esse sistema na forma matricial, tem-se:

$$e = \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_m \end{pmatrix}, \quad y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix}, \quad A = \begin{pmatrix} 1 & t_1 \\ 1 & t_2 \\ \vdots & \vdots \\ 1 & t_m \end{pmatrix}, \quad x = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \quad (3.43)$$

Se os pontos experimentais medidos tivessem todos sobre uma reta, ou seja, fossem colineares, então $y_i = \alpha + \beta t_i$ e e_i para todo i . Dessa forma seria possível resolver o sistema $Ax = y$ e consequentemente $y \in \text{rng}A$. Porém como as medidas de y_i estão sujeitas à erros experimentais, o método dos mínimos quadrados tem o objetivo de minimizar a norma do vetor erro.

$$\text{Erro} = \|e\| = \sqrt{e_1^2 + \dots + e_m^2} \quad (3.44)$$

Dessa forma, objetiva-se determinar os coeficientes do vetor $x = (\alpha, \beta)^T$ que minimize a norma do vetor:

$$\|e\| = \sqrt{Ax - y} \quad (3.45)$$

A solução para esse problema já foi explicitada na seção anterior. De acordo com (3.32) pode-se escrever:

$$(A^T A)x = A^T b \quad (3.46)$$

com a solução:

$$x^* = (A^T A)^{-1} A^T y \quad (3.47)$$

A como a matriz Gram é $K = A^T A$, a matriz A deve possuir colunas linearmente independentes e consequentemente os valores de t_i devem ser distintos. Dessa forma, para que o método dos mínimos quadrados possa ser aplicado, é necessário que haja no mínimo dois instantes de tempo diferentes.

No caso das matrizes mostradas em (3.43), pode-se escrever:

$$A^T A = \begin{pmatrix} 1 & 1 & \dots & 1 \\ t_1 & t_2 & \dots & t_m \end{pmatrix} \begin{pmatrix} 1 & t_1 \\ 1 & t_2 \\ \vdots & \vdots \\ 1 & t_m \end{pmatrix} = \begin{pmatrix} m & \sum t_i \\ \sum t_i & \sum (t_i)^2 \end{pmatrix} = m \begin{pmatrix} 1 & \bar{t} \\ \bar{t} & \overline{(t)^2} \end{pmatrix} \quad (3.48)$$

$$A^T y = \begin{pmatrix} 1 & 1 & \dots & 1 \\ t_1 & t_2 & \dots & t_m \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix} = \begin{pmatrix} \sum y_i \\ \sum t_i y_i \end{pmatrix} = m \begin{pmatrix} \bar{y} \\ \overline{t y} \end{pmatrix} \quad (3.49)$$

Considera-se:

$$\bar{t} = \frac{1}{m} \sum_{i=1}^m t_i \quad \bar{y} = \frac{1}{m} \sum_{i=1}^m y_i \quad \overline{t^2} = \frac{1}{m} \sum_{i=1}^m t_i^2 \quad \overline{t y} = \frac{1}{m} \sum_{i=1}^m t_i y_i \quad (3.50)$$

Substituindo a equação (3.48) em (3.47), determina-se o sistema de equações:

$$\alpha + \bar{t}\beta = \bar{y} \quad \bar{t}\alpha + \overline{t^2}\beta = \overline{t y} \quad (3.51)$$

Dessa forma, as soluções para α e β serão:

$$\alpha = \bar{y} - \bar{t}\beta \quad (3.52)$$

$$\beta = \frac{\bar{t}\bar{y} - \bar{t}\bar{y}}{\bar{t}^2 - \bar{t}^2} = \frac{\sum(t_i - \bar{t})y_i}{\sum(t_i - \bar{t})^2} \quad (3.53)$$

1º Exemplo:

Determinar valores para α e β que resultem no melhor ajuste possível da curva $y(t) = \alpha + \beta t$ para os pontos experimentais mostrados na tabela abaixo.

t_i	0	1	3	6
y_i	2	3	7	12

Escrevendo na forma matricial, tem-se:

$$A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 3 \\ 1 & 6 \end{pmatrix}, \quad A^T = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 3 & 6 \end{pmatrix}, \quad y = \begin{pmatrix} 2 \\ 3 \\ 7 \\ 12 \end{pmatrix}. \quad (3.54)$$

Dessa forma:

$$A^T A = \begin{pmatrix} 4 & 10 \\ 10 & 46 \end{pmatrix}, \quad A^T y = \begin{pmatrix} 24 \\ 96 \end{pmatrix}. \quad (3.55)$$

Substituindo a equação (3.55) em (3.47), determina-se o sistema de equações:

$$\begin{cases} 4\alpha + 10\beta = 24 \\ 10\alpha + 46\beta = 96 \end{cases} \quad (3.56)$$

Resolvendo o sistema obtêm-se:

$$\alpha = \frac{12}{7} \quad \beta = \frac{12}{7} \quad \Rightarrow \quad y = \frac{12}{7} + \frac{12}{7}t \quad (3.57)$$

3.5.1 Ajuste de Polinômios

O método dos mínimos quadrados pode ser expandido para o ajuste de funções polinomiais de grau n , como mostrado em (3.58).

$$y(t) = \alpha_0 + \alpha_1 t + \dots + \alpha_n t^n \quad (3.58)$$

Esse problema pode ser escrito na forma matricial, para isso defini-se:

$$A = \begin{pmatrix} 1 & t_1 & t_1^2 & \dots & t_1^n \\ 1 & t_2 & t_2^2 & \dots & t_2^n \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 1 & t_m & t_m^2 & \dots & t_m^n \end{pmatrix}, \quad x = \begin{pmatrix} \alpha_0 \\ \alpha_1 \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix}. \quad (3.59)$$

De acordo com (3.46), escreve-se:

$$(A^T A)x = A^T b \quad (3.60)$$

com a solução:

$$x^* = (A^T A)^{-1} A^T y \quad (3.61)$$

O erro associado ao ajuste será:

$$\|e\|^2 = \sum_{i=1}^m [y_i - y(t_i)]^2 = \|y - Ax\|^2 \quad (3.62)$$

Capítulo 4

Método dos Mínimos Quadrados Não Linear

No capítulo anterior foi determinado uma solução para o ajuste de funções polinomiais de grau n através do método dos mínimos quadrados. O problema foi escrito na forma de um sistema linear $Ax = b$ (4.1), e assumindo que $\ker A = \{Ax - b\}$, a solução é expressa pela equação (4.2).

$$(A^T A)x = A^T b \quad (4.1)$$

$$x^* = (A^T A)^{-1} A^T b \quad (4.2)$$

Neste capítulo serão mostrados alguns métodos de solução do problema dos mínimos quadrados para ajuste de curvas que possuem modelos matemáticos não lineares. Nesses casos o problema não pode ser escrito na forma de um sistema linear $Ax = b$, e conseqüentemente a solução mostrada em (4.2) não pode ser aplicada.

Funções não lineares estão presentes em vários modelos matemáticos da engenharia mecânica. O ajuste de curvas e estimativa de seus parâmetros são importantes etapas do estudo experimental e estatístico da engenharia. Um exemplo típico em fadiga é calcular os parâmetros da curva de Coffin-Manson que melhor ajustam os dados dos testes de vários corpos de prova de um mesmo material. Este não é um problema trivial, mas é solúvel, por exemplo, pelo algoritmo de Levenberg-Marquardt (LM), que busca numericamente os parâmetros que ajustam pelo método dos mínimos quadrados uma dada função genérica (em geral não linear) a um dado conjunto de pontos.

O método dos mínimos quadrados busca um vetor x^* que minimize a expressão:

$$F(x) = \frac{1}{2} \sum_{i=1}^m (f_i(x))^2 = \frac{1}{2} \|f(x)\|^2 = \frac{1}{2} f(x)^T f(x) \quad (4.3)$$

sendo as funções $f_i: \mathbb{R}^n \mapsto \mathbb{R}^m$, $i = 1, \dots, m$ com $m \geq n$. No caso dos métodos serem aplicado ao ajuste de um modelo matemático y através de m pontos experimentais:

$$(t_1, y_1), \quad (t_2, y_2), \quad \dots \quad (t_m, y_m). \quad (4.4)$$

A expressão $F(x)$ representa o erro total $\|e(x)\|^2$ resultante do ajuste a ser feito.

$$F(x) = \|e(x)\|^2 = \sum_{i=1}^m e_i^2 \quad (4.5)$$

e conseqüentemente a função $f(x)$ será:

$$f_i = y_i - y(t_i). \quad (4.6)$$

O ajuste de funções não lineares pelo método dos mínimos quadrados é feito através do processo iterativo. A partir de um vetor inicial x_0 , o método produz uma série de vetores x_1, x_2, \dots que devem convergir para x^* . A maioria dos métodos utilizados para convergirem utilizam a condição de que:

$$F(x_{k+1}) < F(x_k) \quad (4.7)$$

É importante ressaltar que no caso de $F(x)$ possuir diversos mínimos, o resultado x^* irá depender do valor inicial de x_0 . Não se sabe qual dos mínimos será encontrado e dessa forma, o vetor x^* encontrado não é necessariamente o mais próximo de x_0 .

Em cada iteração, x deve ser substituído por uma nova estimativa $x + h$. Para determinar-se o vetor $h = (h_1, h_2, \dots, h_n)^T$, a função $F(x_i + h)$ é aproximada pela expansão de Taylor:

$$F(x + h) = F(x) + h^T g + \frac{1}{2} h^T H h + O(\|h\|^3) \quad (4.8)$$

onde, g é o gradiente:

$$g \equiv F'(x) = \begin{bmatrix} \frac{\partial F}{\partial x_1}(x) \\ \vdots \\ \frac{\partial F}{\partial x_n}(x) \end{bmatrix} \quad (4.9)$$

e H a hessiana:

$$H \equiv F''(x) = \left[\frac{\partial^2 F}{\partial x_i \partial x_j}(x) \right] \quad (4.10)$$

No caso de x^* ser um mínimo local e $\|h\|$ suficientemente pequeno, não poderá ser determinado um outro ponto $x^* + h$ que resulte um valor menor de F . Dessa forma, sendo x^* um mínimo local, tem-se que:

$$g^* \equiv F'(x^*) = 0 \quad (4.11)$$

Define-se que h é uma direção descendente em $F(x)$ se:

$$h^T F'(x) < 0 \quad (4.12)$$

Os métodos de solução dos mínimos quadrados apresentados neste capítulo necessitam das expressões utilizadas para as derivadas de F . Supondo que f tenha derivadas parciais contínuas, pode-se escrever a expansão de Taylor:

$$f(x + h) = f(x) + J(x)h + O(\|h\|^2) \quad (4.13)$$

com $J \in \mathbb{R}^{m \times n}$ sendo a matriz Jacobiana.

$$(J(x))_{ij} = \frac{\partial f_i}{\partial x_j}(x) \quad (4.14)$$

Como $F : \mathbb{R}^n \mapsto \mathbb{R}$:

$$\frac{\partial F}{\partial x_j}(x) = \sum_{i=1}^m f_i(x) \frac{\partial f_i}{\partial x_j}(x) \quad (4.15)$$

seu gradiente será:

$$F'(x) = J(x)^T f(x) \quad (4.16)$$

Também será necessário definir a matriz Hessiana de F . A partir da equação (4.15), verifica-se que:

$$\frac{\partial^2 F}{\partial x_j \partial x_k}(x) = \sum_{i=1}^m \left(\frac{\partial f_i}{\partial x_j}(x) \frac{\partial f_i}{\partial x_k}(x) + f_i(x) \frac{\partial^2 f_i}{\partial x_j \partial x_k}(x) \right), \quad (4.17)$$

dessa forma:

$$F''(x) = J(x)^T J(x) + \sum_{i=1}^m f_i(x) f_i''(x) \quad (4.18)$$

4.1 Método de Gauss-Newton

O método de Gauss-Newton é baseado em uma aproximação linear para os componentes de f nas proximidades de x . Para valores pequenos de $\|h\|$, a expansão de Taylor (4.13) mostra que:

$$f(x+h) \cong l(h) \equiv f(x) + J(x)h \quad (4.19)$$

Substituindo (4.13) na expressão de $F(x)$ mostrada em (4.3) verifica-se que:

$$\begin{aligned} F(x+h) \cong L(h) &\equiv \frac{1}{2} l(h)^T l(h) \\ &= \frac{1}{2} [(f + Jh)^T (f + Jh)] \\ &= \frac{1}{2} f^T f + h^T J^T f + \frac{1}{2} h^T J^T J h \\ &= F(x) + h^T J^T f + \frac{1}{2} h^T J^T J h \end{aligned} \quad (4.20)$$

com $f = f(x)$ e $J = J(x)$. Como a direção descendente do método de Gauss-Newton h_{gn} deve minimizar $L(h)$, calcula-se a o gradiente e a matriz Hessiana de L .

$$\begin{aligned} L'(h) &= J^T f + J^T J h \\ L''(h) &= J^T J \end{aligned} \quad (4.21)$$

Comparando as expressões (4.21) e (4.16) tem-se que $L'(0) = F'(x)$. Se as colunas da matriz J forem linearmente independentes, então $L''(h)$ é uma matriz positiva definida. Dessa forma, $L(h)$ possui um único mínimo que é achado pela expressão (4.23):

$$\begin{aligned} L'(h_{gn}) &= 0 \\ J^T f + J^T J h_{gn} &= 0 \end{aligned} \quad (4.22)$$

$$\boxed{(J^T J)h_{gn} = -J^T f} \quad (4.23)$$

O vetor h_{gn} é uma direção descendente desde que:

$$h_{gn}^T F'(x) = h_{gn}^T (J^T f) = -h_{gn}^T (J^T J)h_{gn} < 0 \quad (4.24)$$

No processo iterativo de solução pelo método de Gauss-Newton, o vetor x na k -ésima iteração é calculado por:

$$x_{k+1} = x_k + h_{gn} \quad (4.25)$$

4.2 Método de Levenberg-Marquardt

O método de Levenberg-Marquardt é muito semelhante ao método de Gauss-Newton. A diferença é que um fator de amortecimento μ é introduzido na equação (4.23) método de Gauss-Newton. Dessa forma, a direção h_{lm} fica definida como:

$$\boxed{(J^T J + \mu I)h_{lm} = -g} \quad (4.26)$$

com $g = J^T f$ para $\mu \geq 0$.

Nessa equação $f = f(x)$ e $J = J(x)$. O fator de amortecimento μ introduzido é responsável por:

1. Para $\mu > 0$ a matriz $J^T J + \mu I$ é positiva definida e isso assegura que h_{lm} é uma direção descendente.
2. Para valores grandes de μ tem-se:

$$h_{lm} \approx \frac{1}{\mu} J^T f = \frac{1}{\mu} F'(x) \quad (4.27)$$

isso resulta em um passo pequeno na direção descendente.

3. Para valores muito pequenos de μ , tem-se que $h_{lm} \cong h_{gn}$, ou seja, o h_{lm} do método de Levenberg-Marquardt fica muito próximo do h_{gn} do método de Gauss-Newton. Isso é vantajoso nas últimas iterações, quando x está próximo de x^* .

O valor inicial de μ deve estar relacionado com os valores dos elementos de $A_0 = J(x_0)^T J(x_0)$, de forma que:

$$\mu_0 = \tau \cdot \max_i \{a_{ii}^{(0)}\} \quad (4.28)$$

O valor de τ é escolhido pelo usuário do método. Um valor sugerido é $\tau = 10^{-3}$. Durante as iterações o valor de μ é atualizado, e esse processo é controlado pela razão de ganho ζ mostrado em (4.29). Dependendo do valor de ζ em cada iteração, o valor de μ é aumentado ou reduzido.

$$\zeta = \frac{F(x) - F(x + h_{lm})}{L(0) - L(h_{lm})} \quad (4.29)$$

O denominador da expressão para ζ é calculado pela expressão (4.30).

$$\begin{aligned} L(0) - L(h_{lm}) &= -h_{lm}^T J^T f - \frac{1}{2} h_{lm}^T J^T J h_{lm} \\ &= -\frac{1}{2} h_{lm}^T (2J^T f + (J^T J + \mu I - \mu I) h_{lm}) \\ &= \frac{1}{2} h_{lm}^T (\mu h_{lm} - J^T f) \end{aligned} \quad (4.30)$$

Como $h_{lm}^T h_{lm}$ e $h_{lm}^T J^T f$ são positivos, então o denominador $L(0) - L(h_{lm})$ é positivo.

Observa-se que um valor grande de ζ indica que $L(h_{lm})$ é uma boa aproximação para $F(x + h_{lm})$. Nesse caso, deve-se reduzir o valor de μ para que o próximo passo do método de Levenberg-Marquardt se aproxime mais do passo do método de Gauss-Newton. No caso de ζ ter um valor pequeno, deve-se aumentar o valor de μ para que h_{lm} se aproxime mais da direção descendente.

Uma estratégia para o controle do valor de μ foi originalmente proposta por Marquardt (1963). Ela consiste em dobrar o valor de μ caso ζ seja menor do que 0,25, e dividir o valor de μ por três caso ζ seja maior do que 0,75, como é mostrado abaixo:

```

if  $\zeta < 0.25$ 
     $\mu = \mu * 2$ 
elseif  $\zeta > 0.75$ 
     $\mu = \mu / 3$ 

```

Essa estratégia de controle causa uma descontinuidade no valor de μ a cada iteração, e isso pode não ser vantajoso. Por isso, uma nova proposta para o controle do valor de μ foi elaborada por Nielsen (1999), como é mostrado abaixo:

```

if  $\zeta > 0$ 
     $\mu = \mu * \max \{ 1/3, 1 - (2\zeta - 1)^3 \}$ 
     $v = 2$ 
else
     $\mu = \mu * v$ 
     $v = 2 * v$ 

```

O fator v é inicializado em $v = 2$. Além do controle do valor de μ , outra questão relevante que deve ser tratada são os critérios de parada do processo iterativo. São identificadas três situações diferentes que podem caracterizar critérios de parada.

Sabe-se que se o mínimo local for encontrado, então tem-se $F'(x^*) = g(x^*) = 0$. Dessa forma, se a função g assumir um valor menor de um número muito pequeno (ϵ_1), pode-se considerar que o x^* foi encontrado. E conseqüentemente, o processo iterativo deve ser interrompido.

$$\|g\|_\infty \leq \epsilon_1 \quad (4.31)$$

A escolha do valor de ϵ_1 deve ser feita pelo usuário do método. Outra situação que determina que o processo iterativo pode ser interrompido é pequena variação da norma do vetor x entre cada uma das etapas do processo de iteração. Matematicamente, isso é expresso na forma:

$$\|x_{novo} - x\| \leq \epsilon_2(\|x\| + \epsilon_2) \quad (4.32)$$

O valor de ϵ_2 é escolhido pelo usuário do método. Finalmente, em todo processo iterativo deve existir um número máximo de iterações a fim de evitar que ocorra um loop infinito. Por isso, caso o número de iterações (k) exceda um limite máximo (k_{max}) determinado pelo usuário, deve-se interromper o processo de busca de x^* .

$$k \geq k_{max} \quad (4.33)$$

Um resumo do algoritmo de Levenberg-Marquardt é mostrado abaixo:

```

begin
   $k = 0$     $v = 2$     $x = x_0$ 
   $A = J(x)^T J(x)$ 
   $g = J(x)^T f(x)$ 
   $found = \|g\|_\infty \leq \epsilon_1$ 
   $\mu = \tau * \max\{a_{ii}\}$ 
  while (not found) and ( $k < k_{max}$ )
     $k = k + 1$ 
    Resolver  $(A + \mu I)h_{lm} = -g$ 
    if  $\|h_{lm}\| \leq \epsilon_2(\|x\| + \epsilon_2)$ 
       $found = true$ 
    else
       $x_{novo} = x + h_{lm}$ 
       $\zeta = (F(x) - F(x_{novo})) / (L(0) - L(h_{lm}))$ 
      if  $\zeta > 0$ 
         $x = x_{novo}$ 
         $A = J(x)^T J(x)$ 
         $g = J(x)^T f(x)$ 
         $found = \|g\|_\infty \leq \epsilon_1$ 
         $\mu = \mu * \max\{1/3, 1 - (2\zeta - 1)^3\}$ 
         $v = 2$ 
      else
         $\mu = \mu * v$ 
         $v = 2 * v$ 
    end

```

Capítulo 5

Parâmetros do Metabolismo Aeróbio

5.1 Introdução ao Metabolismo Aeróbio

A atividade física é um potente e importante estímulo para o aumento da atividade muscular, cardio-respiratória e vascular. No início do exercício ou de qualquer outra atividade diária (transição: repouso → atividade), em resposta ao aumento da intensidade de esforço e a necessidade de síntese de ATP nos músculos ativos, o metabolismo energético é acelerado, e a taxa de absorção de oxigênio nos pulmões ($\dot{V}O_2$) é aumentada para que haja manutenção do exercício.

A capacidade de ajuste do metabolismo à necessidade de síntese de ATP gerada pela atividade física é um dos fatores determinantes da maior ou menor tolerância ao exercício físico de uma pessoa. Quanto mais rápido ocorrer o ajuste, ou seja, quanto mais rápida for a resposta do organismo ao aumento do esforço físico, menor será o distúrbio celular. Esse ajuste é progressivo e depende da intensidade, duração, modo de exercício e nível de preparo físico da pessoa. Tem como objetivo alcançar o estado estável e retomar a homeostasia celular.

O estudo dos mecanismos fisiológicos responsáveis pela dinâmica da resposta da taxa de absorção de oxigênio nos pulmões $\dot{V}O_2$ durante o exercício e sua subsequente recuperação é conhecida como cinética do consumo de oxigênio.

Vários estudos recentes têm demonstrado que um elevado nível de aptidão aeróbia é necessário para uma alta resistência à prática de esportes e que o baixo nível de aptidão aeróbia está associado a um risco aumentado por várias doenças degenerativas. Para avaliar este nível de aptidão aeróbia, sugere-se que a taxa de $\dot{V}O_2$, medida por meio da ventilação pulmonar durante um exercício controlado de carga incremental ou carga constante, é uma variável que reflete o ajuste sistêmico do transporte de oxigênio e do metabolismo muscular, assim como também da integração entre os sistemas respiratório, cardíaco, vascular e muscular, responsáveis pela captação, transporte e utilização de O_2 .

O treinamento físico sistematizado tem como consequência no organismo adaptações fisiológicas e bioquímicas, levando à melhora no desempenho de tarefas específicas. A natureza e magnitude destas adaptações são dependentes do tipo de atividade realizada, volume, intensidade, carga genética e nível de aptidão inicial. Porém, essas adaptações esperadas só vão ocorrer caso o indivíduo realize atividade física regularmente, e em nível superior ao seu habitual.

5.2 Fases do Metabolismo Aeróbio

Os principais fatores responsáveis pela resposta do aumento da taxa de $\dot{V}O_2$ são o trabalho interno do metabolismo, a potência do trabalho externo submetido ao organismo e a taxa de eficiência.

Os sinais nervosos enviados pelo córtex para os músculos cardíaco-respiratórios e a modificação da relação [ATP]/[ADP] no músculo ativo são estímulos desencadeadores do sistema oxidativo na musculatura esquelética durante o exercício. As adaptações do $\dot{V}O_2$ às cargas de trabalho requerem um funcionamento otimizado do sistema cardiovascular, respiratório e metabólico periférico, com heterocronia de respostas entre eles. Os modelos matemáticos para expressar esses ajustes parecem ser dependentes da magnitude das modificações nesses sistemas, associadas à intensidade e às condições fisiológicas prévias (repouso x exercício; treinados x destreinados).

Classicamente, podem ser identificadas três fases fisiológicas distintas (Fases I, II e III) para o ajuste da curva de $\dot{V}O_2(t)$. Cada uma dessas fases pode ser modelada por uma função exponencial, como mostrado a seguir:

Fase I:

Os primeiros segundos do exercício são caracterizados por um atraso temporário na resposta da taxa de $\dot{V}O_2$ ocasionado pela dissociação entre a taxa de oxigênio absorvido no pulmão e a consumido na musculatura esquelética, especialmente nos músculos responsáveis pela contração muscular. A fase I, comumente chamada de fase ou componente cardio-dinâmica, é modelada por uma equação mono-exponencial da forma:

$$\dot{V}O_2(t) = \dot{V}O_{2BL} + A_c(1 - e^{-t/\tau_c}) \quad \text{se } t < TDp \quad (5.1)$$

onde:

$\dot{V}O_{2BL}$ = taxa de consumo de oxigênio no repouso (base line) [$ml/(Kg.min)$]

A_c = Constante Assintótica da Fase I [$ml/(Kg.min)^{-1}$];

τ_c = Constante de Tempo da Fase I [$1/s$];

Fase II:

Após os primeiros segundos de exercício, inicia-se a fase II. O instante de tempo do início dessa fase é denominado TDp (time delay). Durante a fase I, há utilização dos estoques de oxigênio pelos músculos esqueléticos, ocasionando uma redução no conteúdo de oxigênio do sangue venoso misto. A chegada desse sangue venoso nos pulmões, gera como resposta a fase II (comumente chamada de fase ou componente primária), que é modelada por uma equação mono-exponencial da forma:

$$\dot{V}O_2(t) = A'_c + A_p(1 - e^{-(t-TDp)/\tau_p}) \quad \text{se } TDp < t < TDs \quad (5.2)$$

Nessa equação o termo A'_c garante a continuidade do valor de $\dot{V}O_2$ no instante TDp. Seu valor é:

$$A'_c = \dot{V}O_2(b) + A_c(1 - e^{-TDp/\tau_c}) \quad (5.3)$$

os outros termos de (5.2) são:

A_p = Constante Assintótica da Fase II [$ml/(Kg.min)$];

τ_p = Constante de Tempo da Fase II [$1/s$];

Fase III:

Após o término da fase II, inicia-se a fase III (chamada de fase ou componente lento) no instante de tempo TDs (time delay). Essa fase também pode ser modelada como uma função mono-exponencial da forma:

$$\dot{V}O_2(t) = A'_p + A_s(1 - e^{-(t-TDs)/\tau_s}) \quad se : t > TDs \quad (5.4)$$

Nessa equação o termo A'_p garante a continuidade do valor de $\dot{V}O_{2BL}$ no instante TDs. Seu valor é:

$$A'_p = A'_c + A_p(1 - e^{-(TDs-TDp)/\tau_p}) \quad (5.5)$$

os outros termos de (5.4) são:

A_s = Constante Assintótica da Fase III [$ml/(Kg.min)$];

τ_s = Constante de Tempo da Fase III [$1/s$];

A partir dessas três fases descritas, obtém-se um modelo matemático tri-exponencial para a variação da taxa de consumo de oxigênio no tempo, como mostrado na figura (5.1).

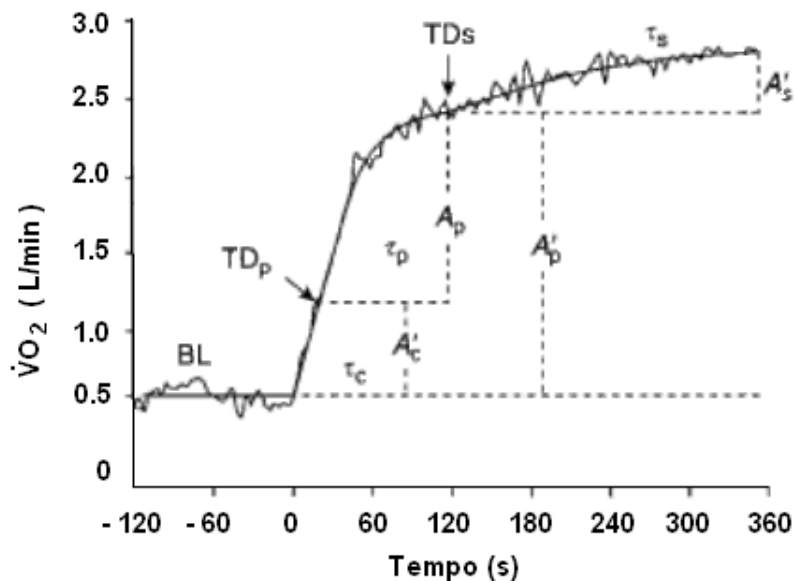


Figura 5.1: Esquema mostrando as três fases exponenciais de $\dot{V}O_2(t)$ para exercícios de intensidade pesada

A existência das três fases depende da intensidade do exercício. Para exercícios de intensidade moderada, a taxa $\dot{V}O_2(t)$ cresce de forma bi-exponencial e assume um valor constante, ou seja, entra em regime permanente, sem que haja a existência da fase III. Dessa forma, nessa intensidade de exercício, a curva de $\dot{V}O_2(t)$ é caracterizada pela existência de apenas as fases I e II. No caso de exercícios de intensidade pesada, a taxa de $\dot{V}O_2(t)$ cresce de forma tri-exponencial (fases I, II e III) e entra em regime permanente após um Δt .

No caso de exercícios de intensidade severa, assim como nos de intensidade pesada, a taxa de $\dot{V}O_2(t)$ cresce de forma tri-exponencial. Porém, a taxa não entra em regime permanente, pois o paciente atinge a fadiga antes de haver tempo suficiente para que $\dot{V}O_2(t)$ atinja um valor constante. A figura (5.2) mostra esquemas para as curvas da taxa $\dot{V}O_2(t)$ para as três intensidades de exercícios descritas acima.

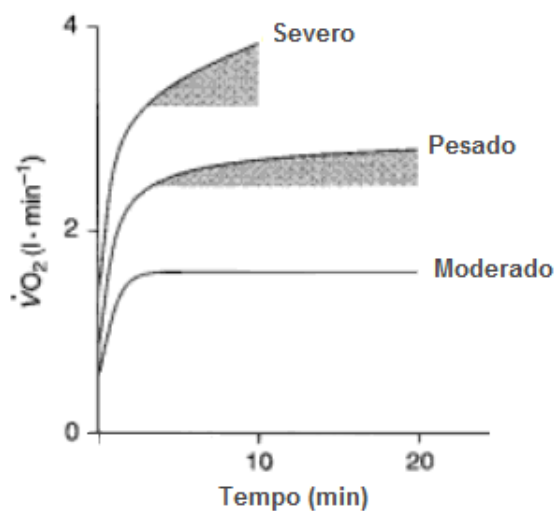


Figura 5.2: Esquema mostrando as curvas de $\dot{V}O_2(t)$ para exercícios de intensidade moderada, pesada e severa. A parte hachurada de cinza mostra a fase III para os exercícios de intensidade pesada e severa.

5.3 Identificação dos Parâmetros do Metabolismo Aeróbio

Fazer o ajuste dos pontos experimentais da taxa de consumo de oxigênio ($\dot{V}O_2(t)$) durante um determinado intervalo de tempo (Δt) para um paciente que realiza um exercício físico de intensidade pesada ou severa não é uma tarefa trivial. Utiliza-se o método dos mínimos quadrados, e o ajuste das curvas das três fases objetiva determinar os valores dos parâmetros $\dot{V}O_{2BL}$, A_c , τ_c , A_p , τ_p , A_s , τ_s , TDP e TDS para um determinado paciente.

Através dos valores obtidos pelo ajuste, um médico deverá ser capaz de avaliar a capacidade de ajuste do metabolismo do paciente à necessidade de síntese de ATP gerada pela atividade física. Estudos recentes têm demonstrado que baixo nível de aptidão aeróbia pode estar associado a várias doenças degenerativas. Por isso, uma avaliação médica da capacidade de ajuste do metabolismo pode ajudar muito no diagnóstico de doenças. Além disso, pode ajudar a determinar quais são

as conseqüências dessas doenças no funcionamento dos sistemas respiratório, cardíaco, vascular e muscular, responsáveis pela captação, transporte e utilização de O_2 .

A taxa de $\dot{V}O_2$ de um paciente é medida por meio da ventilação pulmonar durante um exercício controlado de carga incremental ou carga constante. Dessa forma, durante o exercício são medidos os valores de $\dot{V}O_{2i}$ e seu correspondentes instantes de tempo t_i . Dessa forma, ao fim do exercício tem-se uma seqüência de m pontos medidos experimentalmente, como mostrado abaixo:

$$(t_1, \dot{V}O_{21}), \quad (t_2, \dot{V}O_{22}), \quad \dots \quad (t_m, \dot{V}O_{2m}). \quad (5.6)$$

Para fazer-se o ajuste desses m pontos experimentais a curva de $\dot{V}O_2(t)$, é necessário determinar os valores das seguintes variáveis:

$$x = \begin{pmatrix} \dot{V}O_{2BL} \\ A_c \\ \tau_c \\ A_p \\ \tau_p \\ A_s \\ \tau_s \\ TDp \\ TDs \end{pmatrix} \quad (5.7)$$

Dessa forma, o ajuste da curva de $\dot{V}O_2(t) \in \mathbb{R}$ deve calcular um vetor de variáveis $x \in \mathbb{R}^9$. Deve-se ressaltar que as variáveis TDp e TDs são os instantes de tempo em que há a mudança das fases I para II e das fases II para III. Por isso, no ajuste da curva, as fases I, II e III têm domínio não pré-determinados.

Outro fator que dificulta o ajuste é o fato de que as curvas exponenciais que caracterizam as fases I, II e III são curvas não lineares, e por isso, o ajuste a ser feito em cada uma dessas fases deve utilizar o método dos mínimos quadrados não linear. Neste trabalho será apresentado o ajuste feito através do método de Levenberg-Marquardt.

Feita essa análise preliminar do problema do ajuste da curva de $\dot{V}O_2(t)$, verifica-se que o ajuste a ser feito na verdade engloba dois processos de otimização, que devem ser feitos de forma simultânea. O primeiro processo é determinação de TDp e TDs , de forma que as fases I, II e III possam ser ajustadas da melhor forma possível, e o segundo processo é o ajuste em si das fases I, II e III.

Verifica-se que para determinar os valores de $\dot{V}O_{2BL}$, A_c , τ_c , A_p , τ_p , A_s , τ_s é necessário determinar quais são os valores de TDp e TDs (instantes de mudança de fase). Porém, somente quando esse dois valores forem determinados, e assim o domínio das fases I, II e III definidos, será possível fazer o ajuste dessas três fases, determinando os valores dos outros sete parâmetros do vetor $x \in \mathbb{R}^9$.

A fim de resolver esse problema, o ajuste da curva de $\dot{V}O_2(t)$ nas três fases deve ser feito para todas as possíveis combinações de TDp e TDs . O processo de determinar todas essas combinações é feito da seguinte forma: fixa-se um valor de TDp e varia-se o valor de TDs . Posteriormente, varia-se o valor de TDp , fixando-o em um novo valor, e varia-se novamente o valor de TDp . Essa seqüência de etapas deve ser repetida sucessivamente até que todas as combinações de TDp e TDs

sejam avaliadas.

As possíveis combinações podem ser representadas em uma matriz quadrada. Nas colunas são representados os possíveis valores para TDP , e nas linhas os possíveis valores para TDs . Cada elemento dessa tabela representa uma combinação entre TDP e TDs . Vale ressaltar que como o instante que como obrigatoriamente $TDP < TDs$, os elementos a_{ij} dessa matriz com $j < i$ não representam uma combinação válida de TDP e TDs . A figura (5.3) mostra um esquema da matriz a ser percorrida durante o processo de ajuste da curva de $\dot{V}O_2(t)$.

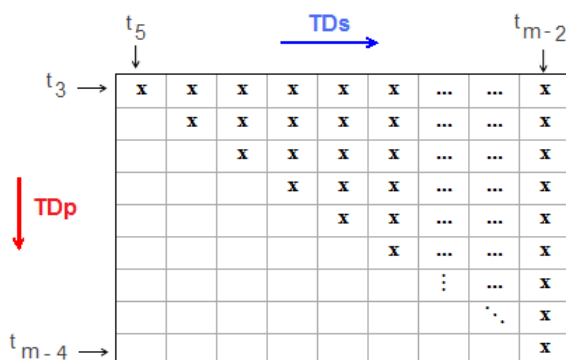


Figura 5.3: Esquema mostrando as combinações válidas dos instantes de mudança de fase TDP e TDs para os m instantes de tempo avaliados.

O domínio de possíveis valores para a variável TDs é um domínio discreto. Assume-se que essa variável só poderá assumir um valor igual a um dos m valores de instante de tempo que foram medidos durante realização do exercício físico. O mesmo ocorre para a variável TDP .

Para cada uma das combinações entre TDP e TDs , faz-se o ajuste das três fases. Inicialmente, faz-se o ajuste da fase I, determinado assim os valores de $Vo2_b$, A_c e tau_c . Depois, faz-se o ajuste da fase II, determinado Ap e τ_p , e por último da fase III, determinado os valores de As e τ_s .

O ajuste tem que obrigatoriamente ser nessa ordem, pois o calculo de $\dot{V}O_2(t)$ na fase II utiliza os valores de $Vo2_b$, A_c e τ_c através da variável A'_c , e o mesmo ocorre na fase III, que utiliza os valores de A'_c , τ_c e A'_p através da variável A'_p .

Além da condição de que $TDP < TDs$, para determinar as combinações válidas entre TDP e TDs , algumas outras condições devem ser impostas. Essas condições são baseadas em imposições do método dos mínimos quadrados não linear.

Como mostrado anteriormente, o método dos mínimos quadrados busca um vetor x^* que minimize a expressão:

$$F(x) = \frac{1}{2} \sum_{i=1}^m (f_i(x))^2 = \frac{1}{2} \|f(x)\|^2 = \frac{1}{2} f(x)^T f(x) \quad (5.8)$$

sendo as funções $f_i: \mathbb{R}^n \mapsto \mathbb{R}^m$, $i = 1, \dots, m$ com $m \geq n$. Isso indica que o número de incógnitas a serem determinadas pelo método, ou seja a dimensão n do vetor $x \in \mathbb{R}^n$, deve ser maior ou igual

ao número m de pontos avaliados. Dessa forma, podem ser impostas as seguintes condições a respeito dos valores válidos de TDP e TDS :

1. A equação característica da fase I possui três parâmetros a serem determinados: $\dot{V}O_{2BL}$, A_c e τ_c , fazendo com que nesse caso $n = 3$. Por isso, para que o método de Levenberg-Marquardt possa ser aplicado na fase I, é necessário que essa fase tenha duração mínima de três intervalos iniciais de tempo medidos durante o exercício físico. Assim, o valor mínimo para TDP corresponde ao terceiro instante de tempo (t_3) medido durante o exercício físico.
2. A equação característica da fase II possui dois parâmetros a serem determinados: A_p e τ_p , fazendo com que nesse caso $n = 2$. Por isso, para que o método de Levenberg-Marquardt possa ser aplicado na fase II, é necessário que essa fase tenha duração mínima de dois intervalos de tempo medidos durante o exercício físico. Porém, como a fase II ocorre após o término da fase I, e a fase I tem duração mínima de três intervalos de tempo, como visto no item anterior, o valor mínimo para TDS corresponde ao quinto instante de tempo (t_5) medido durante o exercício físico.
3. O valor máximo para TDP é determinado a partir do número m total de pontos medidos durante o exercício físico. Sabe-se que ao término da fase I, ocorrem ainda as fases II e III, que tem duração mínima cada uma delas de dois intervalos de tempo. Por isso, maior valor possível para o instante de término da fase I deve ser o instante correspondente ao último instante de tempo (t_m) menos quatro intervalos de tempo, ou seja (t_{m-4}).
4. O valor máximo para TDS é determinado a partir do número m total de pontos medidos durante o exercício físico. Sabe-se que ao término da fase II, ocorre ainda a fase III, que tem duração mínima de dois intervalos de tempo. Por isso, maior valor possível para o instante de término da fase II deve ser o instante correspondente ao último instante de tempo (t_m) menos dois intervalos de tempo, ou seja (t_{m-2}).

As possíveis combinações entre valores válidos de TDP e TDS são mostradas na figura (5.3).

Capítulo 6

Levenberg-Marquardt Aplicado à Identificação dos Parâmetros da curva de $\dot{V}O_2(t)$

6.1 Programa Computacional Desenvolvido

Neste trabalho a plataforma computacional escolhida para fazer a identificação dos parâmetros que caracterizam da curva de $\dot{V}O_2(t)$ foi o *Matlab*. O método utilizado pelo programa é **Levenberg-Marquardt**.

6.1.1 Equações de Levenberg-Marquardt Aplicadas à curva de $\dot{V}O_2(t)$

Abaixo são mostradas as equações que utilizadas pelo método de **Levenberg-Marquardt** aplicadas ao modelo matemático da curva de $\dot{V}O_2(t)$. O objetivo do método é obter o mínimo valor para a função $F(x)$.

$$F(x) = \frac{1}{2} \sum_{i=1}^m (f_i(x))^2 = \frac{1}{2} \|f(x)\|^2 = \frac{1}{2} f(x)^T f(x) \quad (6.1)$$

Porém como são identificadas três fases fisiológicas distintas (Fases I, II e III) e cada uma dessas fases pode ser modelada por uma função exponencial, a função $F(x)$ é definida como sendo o somatório ponderado entre as funções $f(x)$ de cada uma das fases, como mostrado a seguir:

$$F(x) = \frac{F(x)_{Fase I}}{n_{Fase I}} + \frac{F(x)_{Fase II}}{n_{Fase II}} + \frac{F(x)_{Fase III}}{n_{Fase III}} \quad (6.2)$$

onde:

$n_{Fase I}$ = Número de intervalos de tempo medidos entre o início o exercício e o fim da fase I.

$n_{Fase II}$ = Número de intervalos de tempo medidos entre o início o da fase II e o fim da fase II.

$n_{Fase III}$ = Número de intervalos de tempo medidos entre o início o da fase III e o fim do exercício.

É importante mencionar que o primeiro valor de $\dot{V}O_2$ medido experimentalmente não corresponde ao início do exercício $t = 0$. Dessa forma, número total de pontos medidos experimentalmente m corresponde a somatório do número de intervalos de tempo decorridos em cada uma das fases, como mostrado a seguir.

$$m = n_{Fase I} + n_{Fase II} + n_{Fase III} \quad (6.3)$$

As funções $F(x)$ para cada uma das três funções exponenciais que caracterizam a curva de $\dot{V}O_2(t)$ são definidas como:

$$F(x)_{Fase I} = \frac{1}{2} \sum_{i=1}^{n_{Fase I}} (f_{Ii}(x))^2 \quad (6.4)$$

$$F(x)_{Fase II} = \frac{1}{2} \sum_{i=1}^{n_{Fase II}} (f_{IIi}(x))^2 \quad (6.5)$$

$$F(x)_{Fase III} = \frac{1}{2} \sum_{i=1}^{n_{Fase III}} (f_{IIIi}(x))^2 \quad (6.6)$$

Em cada das fases, a função $f(x)$ resulta em um vetor, cuja i -ésima equivale a diferença entre o valor medido experimentalmente para $\dot{V}O_{2i}$, representado por $\dot{V}O_{2data i}$, e o valor de $\dot{V}O_{2i}$ correspondente, que é calculado pela equação exponencial.

Fase I:

$$f_{Ii} = \dot{V}O_{2data i} - (\dot{V}O_{2BL} + A_c(1 - e^{-t_i/\tau_c})) \quad t < TDp \quad (6.7)$$

Fase II:

$$f_{IIi} = \dot{V}O_{2data i} - (A'_p + A_p(1 - e^{-(t_i - TDp)/\tau_p})) \quad TDp < t < TDs \quad (6.8)$$

Fase III:

$$f_{IIIi} = \dot{V}O_{2data i} - (A'_s + A_s(1 - e^{-(t_i - TDs)/\tau_s})) \quad t > TDs \quad (6.9)$$

O método de Levenberg-Marquardt também exige o cálculo da matriz jacobiana de cada uma das três funções exponenciais. Em cada fase, a i -ésima linha da matriz jacobiana será:

Fase I:

$$J_{Ii} = \left(\frac{\partial f}{\partial \dot{V}O_{2BL}} \quad \frac{\partial f}{\partial A_c} \quad \frac{\partial f}{\partial \tau_c} \right) \quad (6.10)$$

$$J_{Ii} = \left(-1 \quad -(1 - e^{-t_i/\tau_c}) \quad \frac{A_c t_i}{\tau_c^2} e^{(-t_i/\tau_c)} \right) \quad (6.11)$$

Fase II:

$$J_{IIIi} = \begin{pmatrix} \frac{\partial f}{\partial A_p} & \frac{\partial f}{\partial \tau_p} \end{pmatrix} \quad (6.12)$$

$$J_{IIIi} = \begin{pmatrix} -(1 - e^{(-\frac{t_i - TDp}{\tau_c})}) & A_c \frac{(t_i - TDp)}{\tau_c^2} e^{(-\frac{t_i - TDp}{\tau_c})} \end{pmatrix} \quad (6.13)$$

Fase III:

$$J_{Fase IIIi} = \begin{pmatrix} \frac{\partial f}{\partial A_s} & \frac{\partial f}{\partial \tau_s} \end{pmatrix} \quad (6.14)$$

$$J_{Fase IIIi} = \begin{pmatrix} -(1 - e^{(-\frac{t_i - TDs}{\tau_c})}) & A_p \frac{(t_i - TDs)}{\tau_c^2} e^{(-\frac{t_i - TDs}{\tau_c})} \end{pmatrix} \quad (6.15)$$

Vale a pena observar que as fases II e III possuem matrizes jacobianas com derivadas parciais idênticas, se for feita a correlação entre A_p e A_s , τ_p e τ_s e por último, TDp e TDs .

6.1.2 Diagrama Modular do Programa Computacional

O programa computacional desenvolvido no software *Matlab* é constituído por 8 arquivos. Abaixo segue uma breve descrição das operações feitas em cada um desses arquivos:

1. Arquivo **ChamadaFit3expLM.m**: Esse arquivo inicia o todo o processo de identificação dos parâmetros, e por isso é o arquivo que deve ser compilado pelo *Matlab*. Uma planilha eletrônica contém os valores de $\dot{V}O_2$ medidos experimentalmente e os valores dos seus m instantes de tempo correspondentes. Por isso, a rotina desse arquivo *ChamadaFit3expLM.m* importa os valores de $\dot{V}O_2$ e t os armazenam em dois vetores de m posições cada, nomeados respectivamente de $\dot{V}O_{2\ data}$ e t_{data} . Posteriormente, a rotina chama a função *Fit3ExpLMpositivo*, que irá iniciar o processo de identificação dos parâmetros da curva de $\dot{V}O_2(t)$. Uma das saídas dessa função será o vetor $R \in \mathbb{R}^9$ com os valores dos nove parâmetros calculados.
2. Arquivo **Fit3ExpLMpositivo.m**: Esse arquivo define a função *Fit3ExpLMpositivo* e possui como entrada três vetores: $\dot{V}O_{2\ data}$, t_{data} e x_0 (um vetor de sete posições com os valores de chute inicial dado para as variáveis: $\dot{V}O_{2BL}$, A_c , τ_c , A_p , τ_p , A_s e τ_s). A saída da função é formada por dois vetores e uma matriz. O primeiro vetor (x) possui os valores dos nove parâmetros que caracterizam a curva de $\dot{V}O_2(t)$ (incluindo os instantes de mudança de fase TDp e TDs), e o segundo vetor (k) possui os valores do número de iterações realizadas em cada uma das fases ajustadas. A matriz de saída (*resnormTotal*) possui em cada uma de suas posições o valor do erro associado ao ajuste feito para uma combinação válida entre TDp e TDs . Nessa matriz, as colunas fazem referência ao valor de TDp e as linhas à TDs , de forma que a matriz possui os valores do erro do ajuste feito para todas as combinações válidas entre TDp e TDs . Para cada uma das combinações válidas, é feito o ajuste das três fases fisiológicas da curva de $\dot{V}O_2(t)$. Para a fase I, o ajuste é feito através da função *LevenbergMarquardtFase1*, e para as fases II e III através da função *LevenbergMarquardtFase23*. A escolha da combinação entre TDp e TDs que resulta no melhor ajuste da curva, é feita através da identificação do menor elemento da matriz *resnormTotal*.
3. Arquivo **LevenbergMarquardtFase1.m**: Define a função *LevenbergMarquardtFase1* responsável por fazer o ajuste da fase I. Suas entradas são: a função da equação exponencial

característica da fase I, um vetor com os valores de $\dot{V}O_{2\ data}$ medidos durante o intervalo considerado como fase I, um vetor com os valores $tdata$ correspondentes, e um vetor de três posições x_0 com os valores de chute inicial para as variáveis $\dot{V}O_{2BL}$, A_c e tau_c . Suas saídas são: um vetor R de três posições com os valores dos três parâmetros ajustados, o número de iterações, o erro associado ao ajuste, e um vetor com os valores desse erro ao longo do processo iterativo. A identificação dos parâmetros é feita através do algoritmo do método de Levenberg-Marquardt, mostrado no capítulo de Mínimos Quadrados Aplicado ao Ajuste de Funções Não Lineares. Para o cálculo da matriz jacobiana de $f_I(x)$ em cada uma das iterações do método, utiliza-se a função *myfJacobianaFaseI*.

4. Arquivo **LevenbergMarquardtFase1positivo.m**: Caso o parâmetro $\dot{V}O_{2BL}$ calculado para a fase I pela função *LevenbergMarquardtFaseI* assuma um valor negativo, refaz-se o ajuste fixando o valor de $\dot{V}O_{2BL}$ em $3.5 [ml/(Kg.min)^{-1}]$. Esse é um valor médio encontrado na literatura para a taxa de absorção de oxigênio no pulmão de uma pessoa enquanto ela está no repouso. Dessa forma, esse arquivo define a função *LevenbergMarquardtFaseIpositivo*, responsável por refazer o ajuste da fase I, assumindo que $\dot{V}O_{2BL} = 3.5$. Suas entradas e saídas são idênticas às da função *LevenbergMarquardtFaseI*. As únicas diferenças são em relação ao vetor x_0 que contém os valores de chute inicial somente para as variáveis A_c e tau_c , e ao vetor R , que possui somente duas posições (valores dos dois parâmetros ajustados).
5. Arquivo **LevenbergMarquardtFase23.m**: Define a função *LevenbergMarquardtFase23* responsável por fazer o ajuste das fases II e III. Suas entradas são: a função da equação exponencial característica da fase II ou III, um vetor com os valores de $\dot{V}O_{2\ data}$ medidos durante o intervalo considerado como fase II ou III, um vetor com os valores $tdata$ correspondentes, e um vetor de duas posições x_0 com os valores de chute inicial para as variáveis A_p e tau_p , na caso de ajuste da fase II, ou para as variáveis A_s e tau_p , na caso de ajuste da fase I. Suas saídas são: um vetor R de três posições com os valores dos parâmetros ajustados, o número de iterações feitas para obter-se o ajuste, o erro associado ao ajuste, e os valores desse erro ao longo do processo iterativo. Para fazer o ajuste a função utiliza o algoritmo do método de Levenberg-Marquardt, que foi mostrado no capítulo de Mínimos Quadrados Aplicado ao Ajuste de Funções Não Lineares. Para o cálculo da matriz jacobiana de $f_{II}(x)$ e $f_{III}(x)$ em cada uma das iterações do método, utiliza-se a função *myfJacobianaFaseI*.
6. Arquivo **myfJacobianaFase1.m**: Define a função *myfJacobianaFaseI* utilizada para calcular a matriz jacobiana durante o ajuste da fase I.
7. Arquivo **myfJacobianaFase1positivo.m**: Define a função *myfJacobianaFaseIpositivo* utilizada para calcular a matriz jacobiana durante o ajuste da fase I assumindo que o valor de $\dot{V}O_{2BL}$ está fixo em $3.5 [ml/(Kg.min)^{-1}]$.
8. Arquivo **myfJacobianaFase23.m**: Define a função *myfJacobianaFaseI* utilizada para calcular a matriz jacobiana durante o ajuste das fases II e III.

Na figura (6.1) é mostrado um diagrama simplificado da seqüência de operações que são executadas pelo programa.

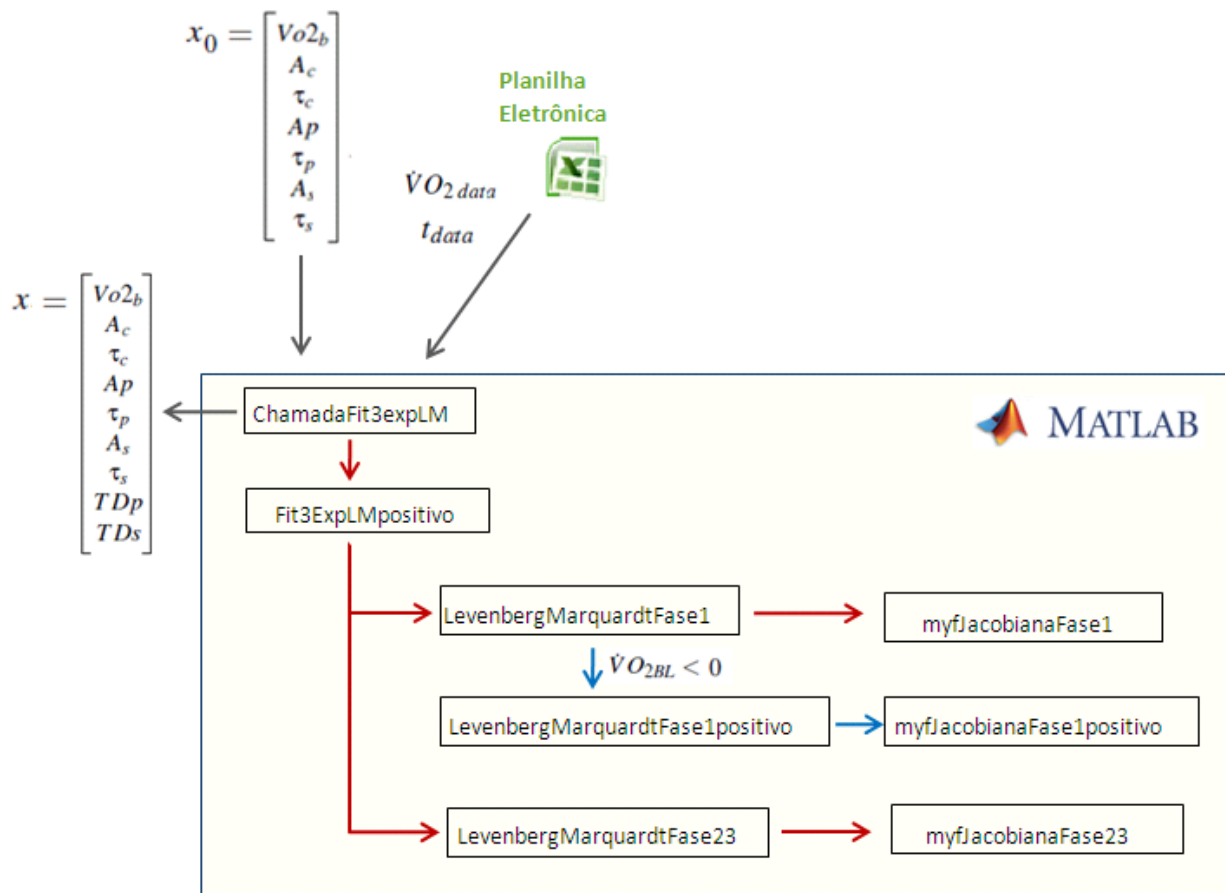


Figura 6.1: Diagrama Modular do Programa Computacional

6.2 Resultados

Esta seção do trabalho contém alguns dos resultados obtidos através do programa computacional desenvolvido em *Matlab*. Os resultados apresentados foram obtidos aplicando os seguintes valores (recomendados pela literatura) para as constantes utilizadas no processo iterativo do método de **Levenberg-Marquardt**:

$$\begin{aligned}
 k_{max} &= 200; \\
 \mu &= \tau * \max \{a_{ii}\}; \\
 \tau &= 10^{-3}; \\
 v &= 2; \\
 \varepsilon_1 &= 10^{-10}; \\
 \varepsilon_2 &= 10^{-10};
 \end{aligned}$$

O vetor $x_0 \in \mathbb{R}^7$ (chute inicial para os parâmetros que caracterizam a curva de $\dot{V}O_2(t)$) não foi definido como um parâmetro fixo para todos os pacientes. Algumas de suas posições (x_{02} , x_{03} e x_{06}) foram definidas como função de $\dot{V}O_{2\ data}$ e t_{data} . Verificou-se, que dessa forma o método de **Levenberg-Marquardt** resultava em ajustes com menor erro, quando comparado com os casos em que x_0 foi assumido constante. Por isso, para cada paciente, $x_0 \in \mathbb{R}^7$ é definido como:

$$x_0 = \begin{pmatrix} \dot{V}O_{2BL} \\ A_c \\ \tau_c \\ A_p \\ \tau_p \\ A_s \\ \tau_s \end{pmatrix} = \begin{pmatrix} 3,5 \\ 10(\dot{V}O_{2\ data\ 3} - \dot{V}O_{2\ data\ 1}) \\ t_{data\ 3} \\ 8 \\ 30 \\ 1,5(\dot{V}O_{2\ data\ m} - A_p - (0,3A_c) - Vo_{2BL}) \\ 2000 \end{pmatrix} \quad (6.16)$$

Os gráficos das figuras (6.2) , (6.3) e (6.4) mostram a curva de $\dot{V}O_2(t)$ ajustada para três pacientes distintos com os seus respectivos valores dos nove parâmetros ajustados. Os valores de $\dot{V}O_{2\ data}$ foram medidos em mililitros de oxigênio consumidos por quilograma de massa corporal por minuto, ou seja, $[ml/(Kg.min)]$, e os valores de t_{data} medidos em segundos.

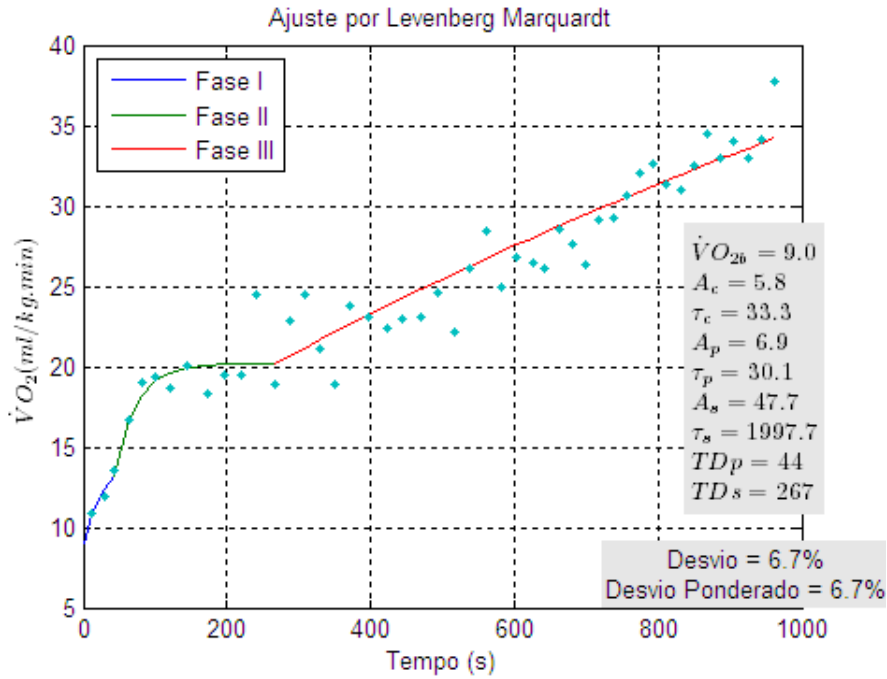


Figura 6.2: Paciente 1 - Curva de $\dot{V}O_2(t)$ ajustada por **Levenberg-Marquardt**

O gráfico da figura (6.5) mostra a curva de $\dot{V}O_2(t)$ para o qual os valores de $\dot{V}O_{2\ data}$ foram medidos em litros de oxigênio consumidos por minuto. Nesse caso, os parâmetros que caracterizam a curva de $\dot{V}O_2(t)$ não possuem as unidades definidas na capítulo 6. Os valores de $\dot{V}O_{2BL}$, A_c , A_p e A_s estão em $[l/(min)]$, e os de τ_c , τ_p e τ_s em $[1/s]$.

Para evitar que a rotina *Matlab* demorasse muito tempo para ser compilada, o número de combinações entre TDp e TDs avaliadas pelo programa não corresponde ao número total de combinações válidas (mostradas na figura (5.3)). Somente são avaliadas as combinações que envolvem os valores mais prováveis de TDp e de TDs .

Depois do programa ser compilado algumas vezes em sua fase de teste, observou-se que o instante TDp sempre correspondia ao terceiro instante de tempo do vetor t_{data} . Por isso, as combinações entre TDp e TDs avaliadas pela rotina *matlab* poderiam estar restritas àquelas que envolvessem valores de TDp inferiores à décima posição do vetor t_{data} .

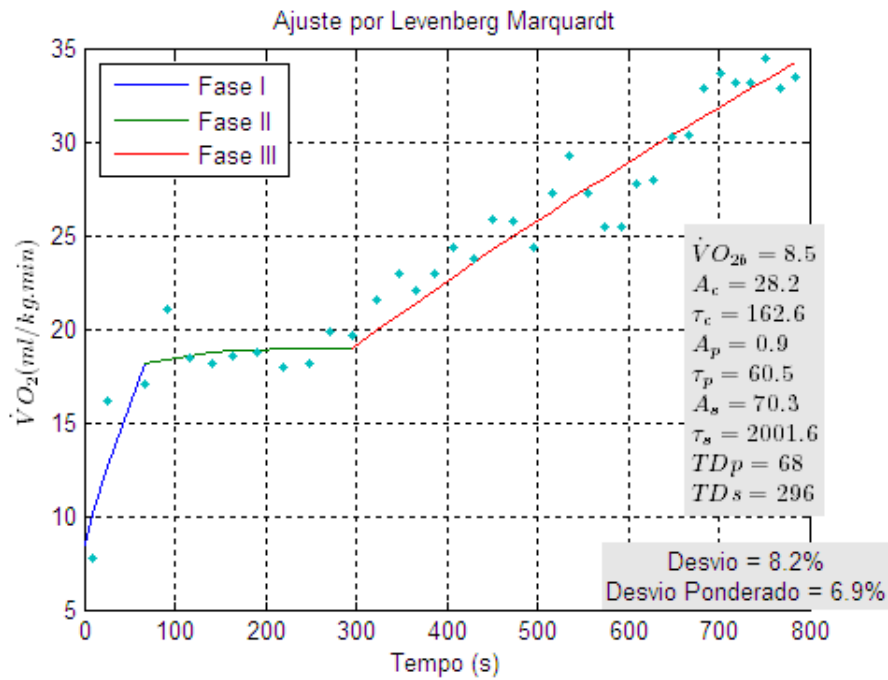


Figura 6.3: Paciente 2 - Curva de $\dot{V}O_2(t)$ ajustada por **Levenberg-Marquardt**

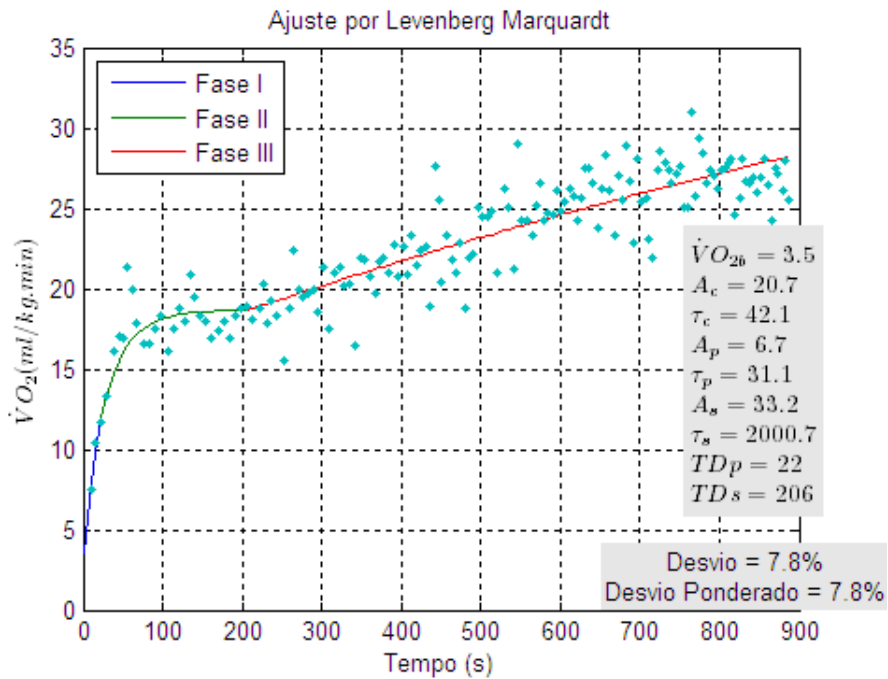


Figura 6.4: Paciente 3 - Curva de $\dot{V}O_2(t)$ ajustada por **Levenberg-Marquardt**

Também se observou que o instante TD_s sempre assumia valores superiores a 200 segundos. Por isso, outra restrição imposta ao programa é avaliar apenas as combinações entre TD_p e TD_s que envolvessem valores de TD_p superiores a 200 segundos.

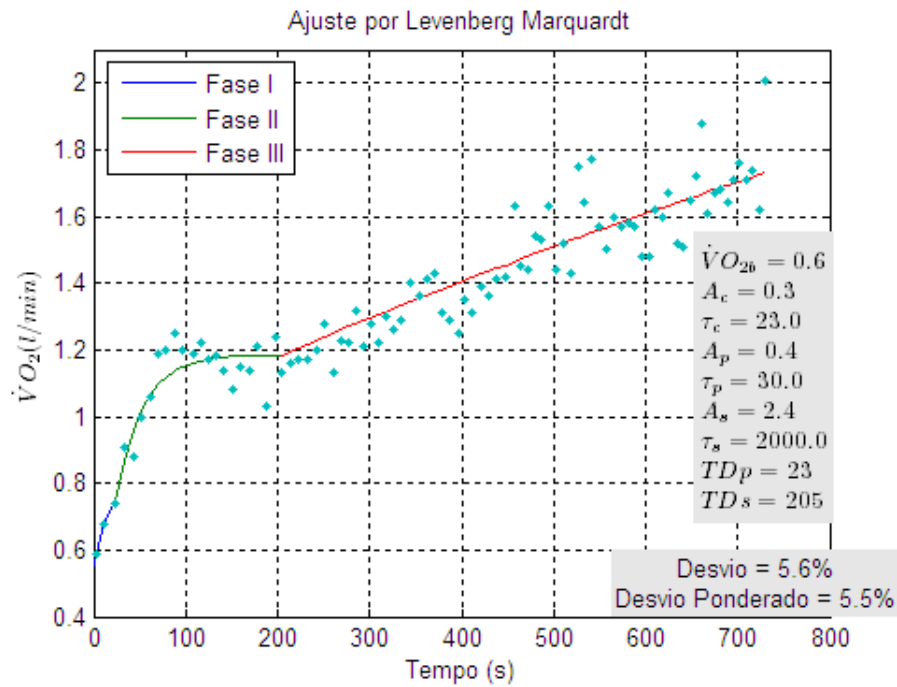


Figura 6.5: Paciente 4 - Curva de $\dot{V}O_2(t)$ ajustada por **Levenberg-Marquardt**

O esquema da figura (6.6) mostra as combinações entre TDp e TDs avaliadas pelo programa *Matlab* impondo restrições de que $TDp < t_{data10}$ e $TDs > 200$ segundos.

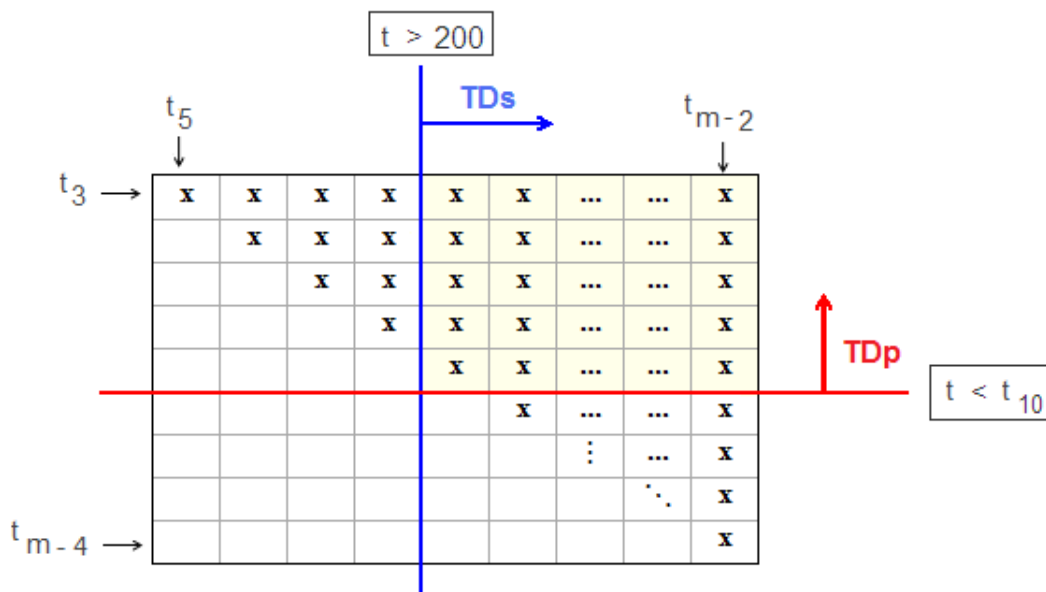


Figura 6.6: Combinações entre TDp e TDs avaliadas pelo programa *Matlab*

A figura (6.7) representa a matriz *resnormTotal* para o paciente 1 (6.2), com os valores do erro do ajuste para algumas das combinações válidas entre TDp e TDs . Nos eixos da base estão representados os instantes de tempo de mudança de fase, e no eixo vertical os erros de ajuste (calculados pela equação (6.2)).

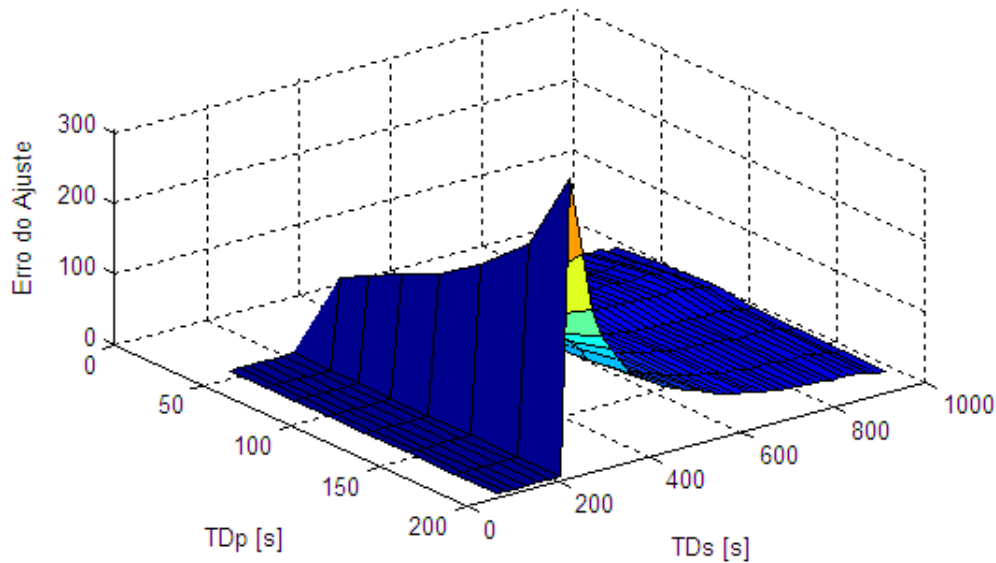


Figura 6.7: Paciente 1 - Erro do Ajuste em função das combinações entre TDp e TDs

As figuras (6.8), (6.9), (6.10) e (6.11) fazem referência aos quatro pacientes dos gráficos 6.2, 6.3, 6.4 e 6.5. Elas mostram a curva de $\dot{V}O_2(t)$ gerada com os valores iniciais dos parâmetros contidos em x_0 , e a curva final de $\dot{V}O_2(t)$ obtida com o método de **Levenberg-Marquardt**.

Deve-se ressaltar que esses gráficos poderiam ser construídos para cada uma das combinações entre TDp e TDs avaliadas pela rotina *Matlab*. Porém, as figuras (6.8), (6.9), (6.10) e (6.11) mostram os gráficos em relação a dupla de instantes de mudança de fase que resultaram em menor erro de ajuste.

Observa-se ainda nessas figuras que tanto no início da fases II quanto da III, o valor de $\dot{V}O_2$ calculado por x_0 coincide com o valor final da fase anterior que é calculado através dos parâmetros ajustados por **Levenberg-Marquardt**. Isso é devido a condição de continuidade entre as fases imposta pelos parâmetros A'_c e A'_p .

Outras quatro figuras relevantes a esse trabalho são os gráficos (6.12, 6.13, 6.14 e 6.15) do valor assumido pelo erro de ajuste ($F(x)$) ao longo do processo iterativo de cada uma das fases da curva de $\dot{V}O_2(t)$. Esse erro é calculado através das equações (6.4), (6.5) e (6.6). Verifica-se que para todas as fases, o número de iterações realizadas é menor do que 200 (valor definido para k_{max}), e que em poucos passes o erro assume um valor constante.

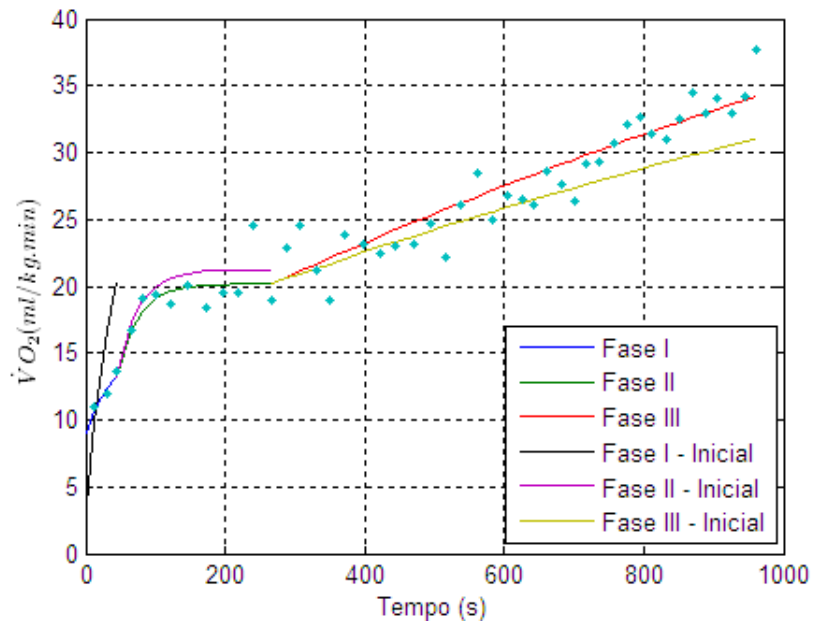


Figura 6.8: Paciente 1 - Comparação entre as curvas de $\dot{V}O_2(t)$ geradas por x_0 e pelos parâmetros calculados por **Levenberg-Marquardt**

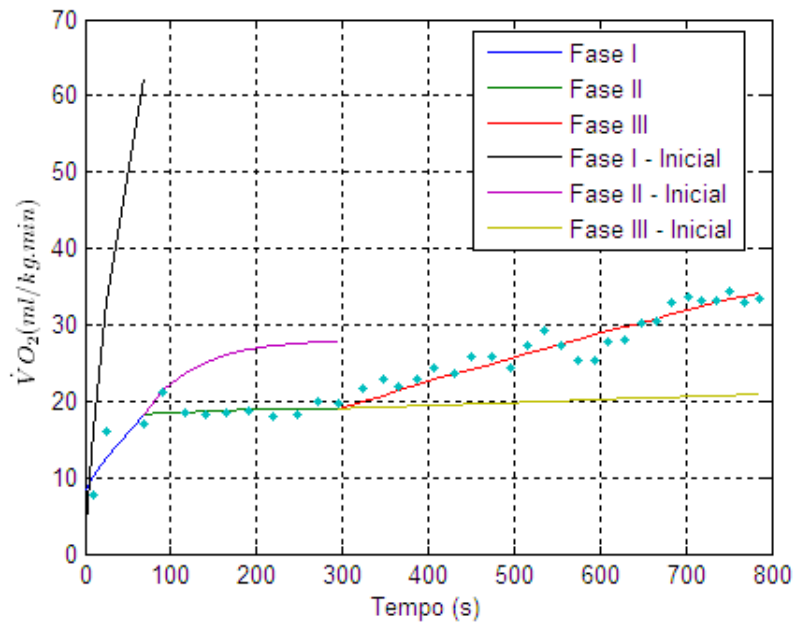


Figura 6.9: Paciente 2 - Comparação entre as curvas de $\dot{V}O_2(t)$ geradas por x_0 e pelos parâmetros calculados por **Levenberg-Marquardt**

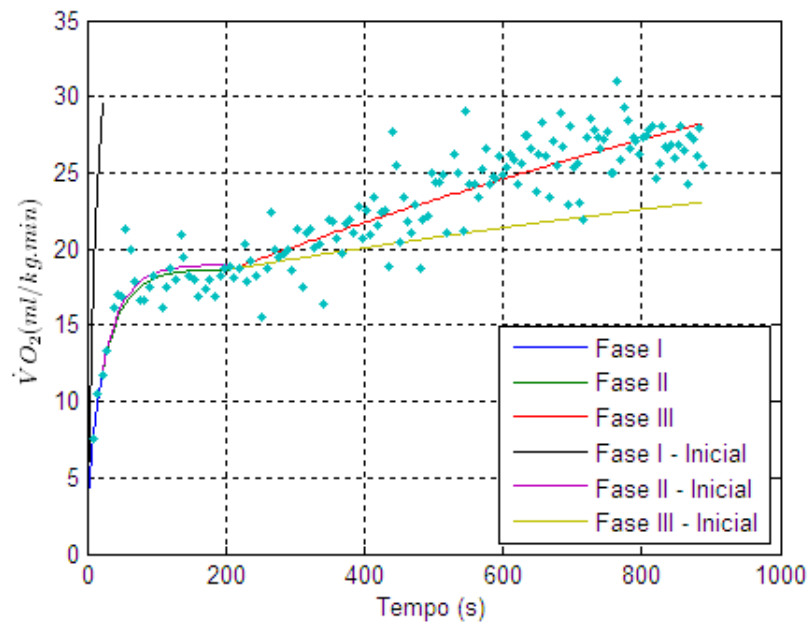


Figura 6.10: Paciente 3 - Comparação entre as curvas de $\dot{V}O_2(t)$ geradas por x_0 e pelos parâmetros calculados por **Levenberg-Marquardt**

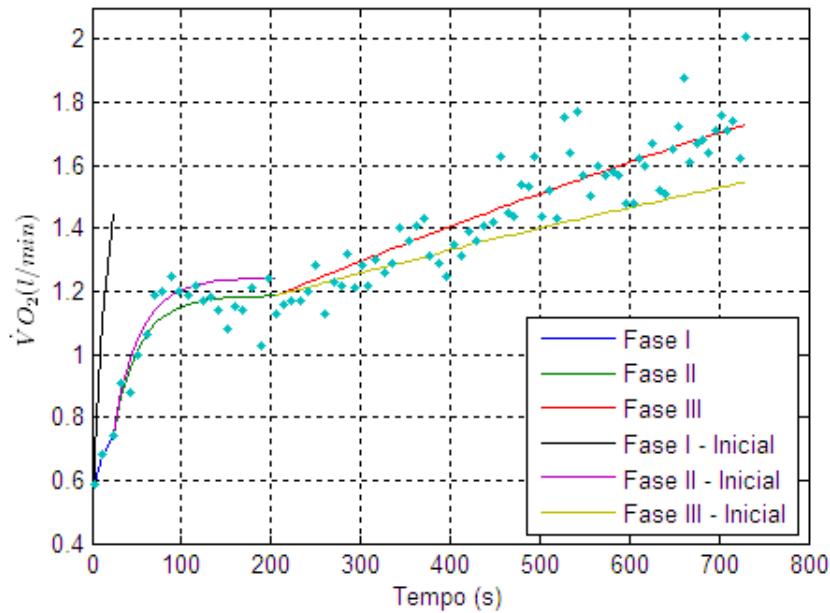


Figura 6.11: Paciente 4 - Comparação entre as curvas de $\dot{V}O_2(t)$ geradas por x_0 e pelos parâmetros calculados por **Levenberg-Marquardt**

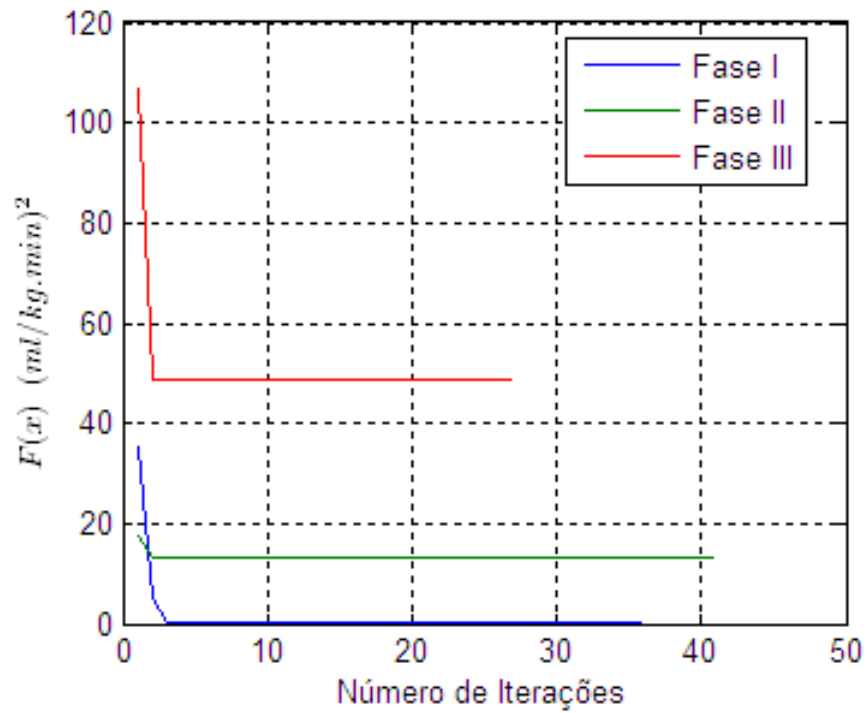


Figura 6.12: Paciente 1 - $F(x)$ em função do número de iterações

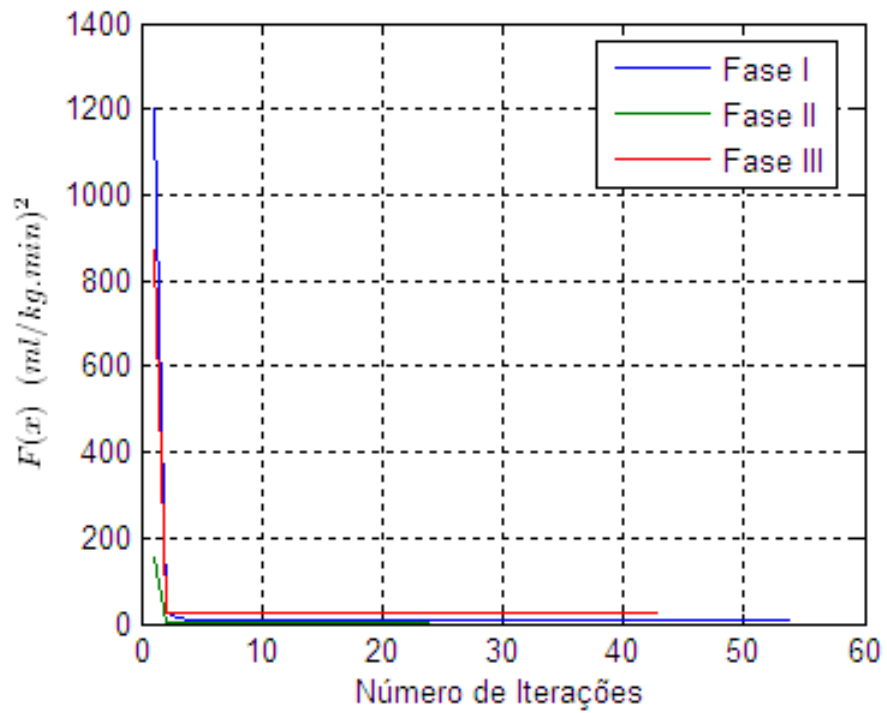


Figura 6.13: Paciente 2 - $F(x)$ em função do número de iterações

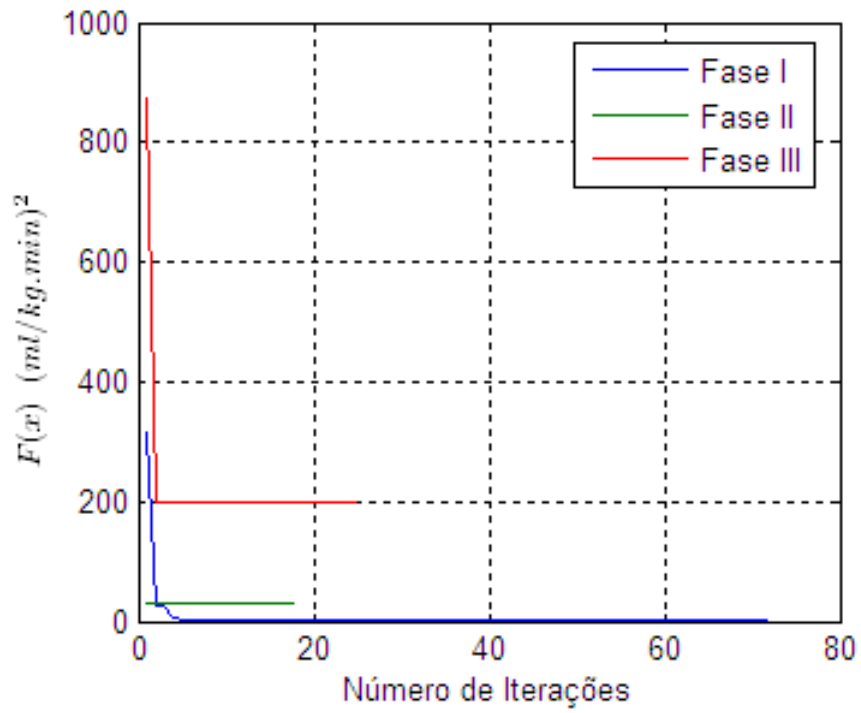


Figura 6.14: Paciente 3 - $F(x)$ em função do número de iterações

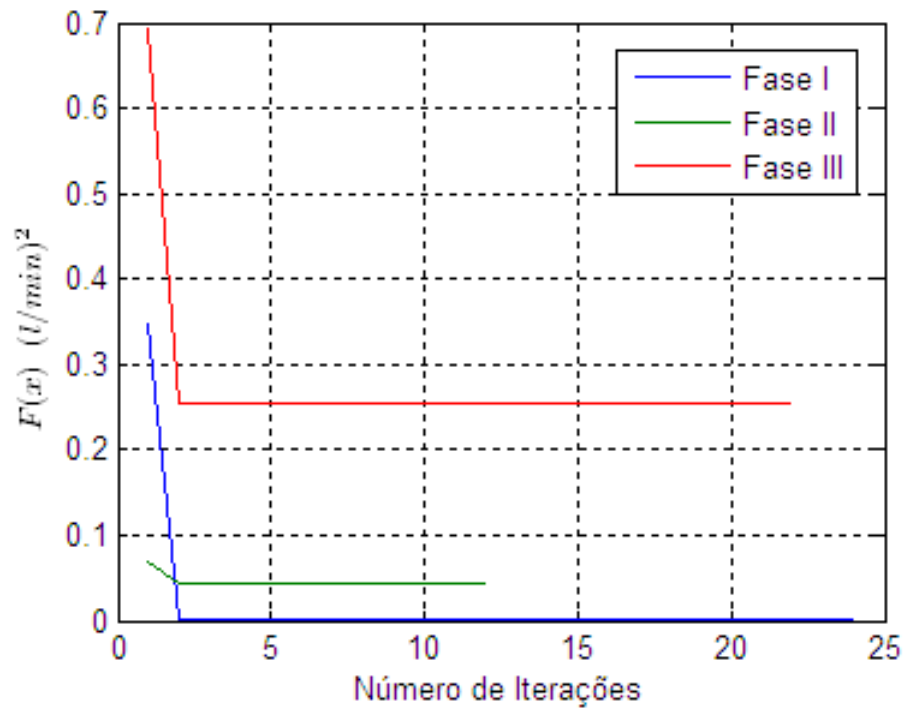


Figura 6.15: Paciente 4 - $F(x)$ em função do número de iterações

Capítulo 7

Conclusões

Neste trabalho mostrou-se como os parâmetros que caracterizam um determinado modelo matemático podem ser calculados a partir do método dos mínimos quadrados linear e não linear. Para isso, foram utilizados conceitos da álgebra linear e de métodos numéricos.

A proposta de se criar um programa automático de cálculo dos parâmetros que caracterizam a curva de $\dot{V}O_2(t)$ foi alcançada. Foi desenvolvida uma rotina *Matlab* capaz de identificar parâmetros pelo método dos mínimos quadrados não linear de **Levenberg-Marquardt**.

Estudos recentes mostram que pacientes saudáveis apresentam valores para os parâmetros da curva de $\dot{V}O_2(t)$ diferentes dos valores apresentados por pacientes que possuem doenças degenerativas. Dessa forma, a grande motivação desse trabalho foi desenvolver um programa computacional que através do cálculo dos valores desses parâmetros para pacientes saudáveis, possa no futuro, ser utilizado como ferramenta de auxílio para o diagnóstico de várias doenças.

No programa a curva de $\dot{V}O_2(t)$ é modelada por três funções exponenciais seqüenciais. Cada uma dessas funções é classificada como uma fase fisiológica distinta: fases I, II e III. Elas representam a resposta do organismo ao aumento do metabolismo energético. E refletem o ajuste sistêmico que ocorre entre os sistemas respiratório, cardíaco, vascular e muscular, responsáveis pela captação, transporte e utilização de O_2 .

O ajuste matemático das três fases exigiu a identificação de nove parâmetros (incluindo os instantes de tempo em que há a mudança das fases I para II e II para III). Esse ajuste foi feito de forma iterativa e produziu resultados muito satisfatórios. O programa desenvolvido disponibiliza os resultados em forma gráfica. Neste trabalho, são mostrados gráficos com as curvas ajustadas para quatro pacientes distintos.

Além do cálculo dos parâmetros, através da rotina *Matlab* verificou-se também que o método de **Levenberg-Marquardt** após um número muito pequeno de iterações obtém um valor constante de erro associado ao ajuste. Isso mostra que o método apresenta uma solução estável ao problema de identificação de parâmetros.

Referências Bibliográficas

- [1] HUFFEL S., VANDEWALLE J. **The Total Least Squares Problem: Computational Aspects and Analysis**. Philadelphia: Society for Industrial and Applied Mathematics, 1991. 300p.
- [2] LAWSON J., HANSON R. **Solving Least Squares Problems**. New Jersey: Prentice-Hall, 1974. 340p.
- [3] JONES, A., POOLE D. **Oxygen Uptake Kinetics in Sport, Exercise and Medicine**. Reino Unido: Routledge, 2005. 403p.
- [4] PENNY J., LINDFIELD G. **Numerical Methods Using MATLAB**. Reino Unido: Ellis Horwood, 1995. 328p.
- [5] JUANG J. **Applied System Identification**. Estados Unidos: Prentice-Hall, 1994. 394p.
- [6] TAYLOR J. **An Introduction to Error Analysis: The Study of Uncertainties in Physical Measurements**. 2.ed. Califórnia: University Science Books, 1996. 327p.
- [7] HENSON, L., POOLE, D., WHIPP B. Fitness as a determinant of oxygen uptake response to constant-load exercise. **European journal of applied physiology and occupational physiology**, v. 59, n. 1-2, p. 21-28, set. 1989.
- [8] WILLIAMS C., CARTER H., JONES A., DOUST J. H. Oxygen uptake kinetics during treadmill running in boys and men. **Journal of Applied Physiology**, v. 90, n. 5, p. 1700-1706, maio 2001.
- [9] TANABE Y., TAKAHASHI M., HOSAKA Y., ITO M., ITO E., SUZUKI K. Prolonged Recovery of Cardiac Output After Maximal Exercise in Patients With Chronic Heart Failure. **Journal of the American College of Cardiology**, v. 35, n. 5, p. 1228-1236, abril 2000.
- [10] BEARDEN S., MOFFATT R. $\dot{V}O_2$ and heart rate kinetics in cycling: transitions from an elevated baseline. **Journal of the American College of Cardiology**, v. 90, n. 6, p. 2081-2087, jun. 2001.
- [11] SILVA, A., OLIVEIRA, F. Consumo de oxigênio durante o exercício físico: aspectos temporais e ajustes de curvas. **Revista Brasileira de Cineantropometria & Desempenho Humano**, v. 6, n. 2, p. 73-82, nov. 2004.
- [12] SILVA A., GAGLIARDI J., LOTUFO R., KISS M. Ausência do componente lento no VO_2 em protocolo de rampa. **Revista de Educação Física**, v. 10, n. 3, p. 181-188, set./dez. 2004.

- [13] AZEVEDO, P. **Cinética do Consumo de Oxigênio no Domínio Severo: Comparação entre Homens e Mulheres Saudáveis e Sedentários**. São Carlos, 2007. 80p. Dissertação de mestrado (Programa de Pós-Graduação em Ciências Fisiológicas) - Universidade Federal de São Carlos.
- [14] MOTULSKY H., BROWN R. Detecting outliers when fitting data with nonlinear regression: a new method based on robust nonlinear regression and the false discovery rate. **BMC Bioinformatics**, v. 7, n. 123, p. 1-20, mar. 2006.
- [15] STRANG G. **Linear Algebra and its Applications**. 3.ed. Estados Unidos: Harcourt College Publishers, 1988. 505p.
- [16] RUGGIERO M., LOPES V. **Cálculo Numérico: Aspectos Teóricos e Computacionais**. São Paulo: Makron Books, 1997. 406p.
- [17] OLVER P. J. **Applied Mathematics**. Minneapolis: School of Mathematics, University of Minnesota, 2003. 1059p.
- [18] MARQUARDT D. W. An Algorithm for Least-Squares Estimation of Nonlinear Parameters. **Journal of the Society for Industrial and Applied Mathematics**, v. 11, n. 2, p. 431 - 441, jun. 1963. Published by: Society for Industrial and Applied Mathematics.
- [19] MADSEN K., NIELSEN H., TINGLEFF O. **Methods for Non-Linear Least Squares Problems**. 2.ed. Dinamarca: Technical University of Denmark , 2004. 58p.