

**MODELAGEM NÃO LINEAR DA EXTRAÇÃO
DE ZINCO EM UM SOLO TRATADO COM
LODO DE ESGOTO**

EDCARLOS MIRANDA DE SOUZA

2008

EDCARLOS MIRANDA DE SOUZA

**MODELAGEM NÃO LINEAR DA EXTRAÇÃO DE ZINCO EM UM
SOLO TRATADO COM LODO DE ESGOTO**

Dissertação apresentada à Universidade Federal de Lavras como parte das exigências do Programa de Pós-Graduação em Estatística e Experimentação Agropecuária, para a obtenção do título de "Mestre".

Orientador
Prof. Dr. Joel Augusto Muniz

LAVRAS
MINAS GERAIS - BRASIL
2008

**Ficha Catalográfica Preparada pela Divisão de Processos Técnicos da
Biblioteca Central da UFLA**

Souza, Edcarlos Miranda de.

Modelagem não linear da extração de zinco em um solo tratado
com lodo de esgoto / Edcarlos Miranda de Souza. -- Lavras : UFLA, 2008.
95 p. : il.

Dissertação (Mestrado) – Universidade Federal de Lavras, 2008.

Orientador: Joel Augusto Muniz

Bibliografia.

1. Regressão não linear. 2. Elementos-traço. 3. Lodo de esgoto. 4.
“Bootstrap”. 5. Curvatura de Bates e Watts. I. Universidade Federal de
Lavras. II. Título.

CDD- 519.536

EDCARLOS MIRANDA DE SOUZA

**MODELAGEM NÃO LINEAR DA EXTRAÇÃO DE ZINCO EM UM
SOLO TRATADO COM LODO DE ESGOTO**

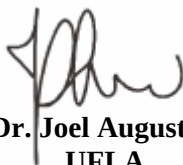
Dissertação apresentada à Universidade Federal de Lavras como parte das exigências do Programa de Pós-graduação em Estatística e Experimentação Agropecuária, para a obtenção do título de "Mestre".

APROVADA em 04 de Julho de 2008

Prof. Dr. Augusto Ramalho de Moraes UFLA

Prof. Dr. João Domingos Scalon UFLA

Dr. Giuliano Marchi EMBRAPA/CERRADOS



Prof. Dr. Joel Augusto Muniz
UFLA
(Orientador)

**LAVRAS
MINAS GERAIS – BRASIL**

À memória de:

meu avô, Oscar Felício de Souza; meu pai, José
Cordeiro de Souza e meu amigo Mário Antônio Lino
dos Santos, pelo amor, bondade e confiança.

OFEREÇO

A minha mãe, Cosma, pelo apoio,
confiança, paciência e amor que me dedica.

À Carla e aos meus filhos, Eduardo, Luis e Sabrina.

Aos meus irmãos, Ângelo, Luciano, Emanuel, Gleide e Eliana.

Aos meus tios Major, Damiana, Celina, Socorro e Raimundo Alves.

DEDICO

AGRADECIMENTOS

A Deus, por ter me concedido essa oportunidade de concluir este trabalho.

Ao professor Joel Augusto Muniz, pelo apoio, estímulo, confiança e orientação no desenvolvimento deste trabalho e por ser um verdadeiro amigo.

Ao Giuliano Marchi, pela co-orientação e sua esposa Edilene, por todo o companheirismo.

Aos professores João Domingos Scalon e Augusto Ramalho de Moraes, pelas valiosas sugestões.

À Universidade Federal do Acre, pela liberação e apoio financeiro.

Aos professores do Centro de Ciências Exatas e Tecnológicas da UFAC, pela confiança depositada. Especialmente, ao professor José Ivan, pela orientação durante a graduação e pelo incentivo à carreira acadêmica e ao professor Antônio Carlos, pelo tempo dedicado às discussões, desde o processo seletivo e ao Haroldo, meu colega de sala, por nossa bela amizade.

À Universidade Federal de Lavras e ao Programa de Pós-Graduação em Estatística e Experimentação Agropecuária, bem como a todos os funcionários do DEX-UFLA, especialmente a Josi, Joyce, Edila e Selminha.

A todos os professores da Pós-Graduação em Estatística e Experimentação Agropecuária da UFLA, especialmente aos professores Daniel Furtado, Lucas Monteiro e Marcos Vinício, pela assistência valiosa com relação a dúvidas pertinentes durante todo o curso.

A Silva, pelo apoio, compreensão, carinho, incentivo e pelas prazerosas e proveitosas discussões que tivemos em vários momentos.

Ao meu amigo Altemir (Bibi), por todo o companheirismo durante essa jornada e pela sincera amizade que temos.

A todos os meus amigos da graduação e especialização, especialmente Itamar, Jairo, Osvaldo, André, Aledias, Edileudo, José Roberto (Empresário), Wilson, Tony, Rodiney e Alessandro, pelas “horas alegres” de estudo.

A todos os meus amigos das vilas Santa Cecília, Albert Sampaio e Dom Moacir e a todos os membros da minha família, sobrinhos, primos, tios, etc.

Ao meu grande amigo Mário Luiz, pelo apoio, incentivo, torcida e principalmente, pela nossa bela e sincera amizade.

A todos os professores do cursinho pré-vestibular Ideal, pelo apoio, amizade e compreensão em diversas situações.

A “dona” Maria, que me motivou a continuar na carreira docente com sua alegria, respeito e amizade, em muitos momentos.

A todos os colegas da pós-graduação em Estatística e Experimentação Agropecuária da UFLA. Especialmente, ao Fábio e à Patrícia (Campo Belo).

À galera da Bahia: Wal, Vitória, Cleíton, Laine e Elma, pela amizade.

Ao trio parada dura (swat), Altemir (Bibi), Paulo (Palito) e Ricardo (Bebê) pela convivência harmoniosa e amizade sincera durante essa jornada.

Aos meus grandes amigos do grupo de estudo, Patrícia (Patinha), Augusto, Ana Paula, Tânia e Richardson, pela grande amizade, companheirismo, troca de conhecimentos, e inúmeros momentos de alegria.

Ao meu grande amigo Eustáquio, a minha linda amiga Francisca (minha “Franzinha”) e a Dalvinha, pelo carinho, estímulo, sugestões e, principalmente, pela grande e sincera amizade construída durante todo este tempo.

A minha mãe, minha maior paixão, mais uma vez quero agradecer por tudo.

SUMÁRIO

RESUMO.....	i
ABSTRACT.....	ii
1 INTRODUÇÃO.....	01
2 REFERENCIAL TEÓRICO.....	03
2.1 Ácidos orgânicos e extractabilidade de elementos-traço em solos tratados com lodo de esgoto.....	03
2.2 Modelos de regressão linear.....	05
2.2.1 Os estimadores de mínimos quadrados e de máxima verossimilhança.....	08
2.3 Modelos de regressão não linear.....	12
2.3.1 Modelos de regressão intrinsecamente lineares.....	13
2.3.2 O problema das transformações.....	15
2.3.3 Modelos de regressão intrinsecamente não lineares.....	15
2.4 Estimativas dos mínimos quadrados não lineares.....	18
2.4.1 Aproximação linear.....	19
2.4.2 O método de Gauss-Newton.....	20
2.4.3 Convergência.....	23
2.4.4 Valores iniciais.....	24
2.5 Adequação do modelo.....	25
2.5.1 O estimador de σ^2	26
2.5.2 Coeficiente de determinação – R^2	26
2.6 Considerações importantes sobre $\hat{\theta}$	27
2.7 Regiões de inferência aproximada para os parâmetros.....	28
2.8 Medidas de não linearidade.....	29

2.8.1 O local da solução.....	30
2.8.2 Linhas de parâmetro.....	33
2.8.3 Medidas de curvatura de Bates e Watts.....	35
2.8.4 Curvaturas.....	39
2.8.5 Calculando curvaturas.....	41
2.8.6 Curvaturas RMS.....	43
2.9 Estudos de simulação.....	45
3 MATERIAL E MÉTODOS.....	48
3.1 Material.....	48
3.2 Métodos.....	50
3.2.1 Estimativas iniciais.....	50
3.2.2 Estimativas.....	53
3.2.3 Bandas de confiança para resposta esperada.....	53
3.2.4 Região de confiança conjunta para os parâmetros.....	55
3.2.5 Intervalos de confiança, correlação entre os parâmetros e R^2	56
3.2.6 Estudos de simulação: “bootstrap” paramétrico.....	57
3.2.6.1 Intervalos de confiança “bootstrap”.....	59
3.2.7 Estudo da curvatura de Bates e Watts.....	60
4 RESULTADOS E DISCUSSÃO.....	61
4.1 Valores iniciais.....	61
4.2 Estimativas e resultados inferenciais.....	61
4.3 Resultados gráficos.....	63
4.4 Simulação “bootstrap”.....	70
4.5 Medidas de curvatura de Bates e Watts.....	76
5 CONCLUSÕES.....	79
REFERÊNCIAS BIBLIOGRÁFICAS.....	81
ANEXOS.....	84

RESUMO

SOUZA, Edcarlos Miranda. **Modelagem não linear da extração de zinco em um solo tratado com lodo de esgoto**. 2008. 95 p. Dissertação (Mestrado em Estatística e Experimentação Agropecuária) - Universidade Federal de Lavras, Lavras, MG.¹

Os modelos de regressão não lineares são de grande importância para cientistas e pesquisadores das diversas áreas do conhecimento, pela natureza do ajuste e a interpretação prática dos parâmetros. Na ciência do solo, a predição da disponibilidade de elementos-traço em solos tratados com lodo de esgoto por extratores em uma única extração é dificilmente obtida. Por essa razão, Marchi (2005) propôs um método que simulou a solução da rizosfera, por meio de extrações sucessivas com misturas de ácidos orgânicos sintéticos para acessar a disponibilidade potencial desses elementos a longo prazo. Os dados obtidos pelas extrações sucessivas foram ajustados por meio de um modelo não linear, baseado nos mecanismos cinéticos de primeira ordem, observando-se a dissolução de Cd, Zn, Cr, Ni e Cu. Foi realizado um estudo de forma rigorosa sobre as propriedades estatísticas dos estimadores, bem como os ajustes, os intervalos de confiança, as bandas de confiança para resposta esperada e a região de confiança conjunta. Realizou-se um esboço sobre a não linearidade do modelo, tendo como referência a medida de curvatura de Bates e Watts, com o objetivo de verificar as propriedades assintóticas dos estimadores, apesar do tamanho reduzido da amostra. Neste trabalho, considerou-se apenas o elemento-traço Zn e utilizou-se a metodologia clássica da amostragem e da verossimilhança, tendo como técnica alternativa a metodologia “bootstrap”. Este estudo proporciona um roteiro detalhado e rotina elaborada no software R versão 2.7.0, que poderá orientar trabalhos futuros sobre extrações sucessivas, tanto de elementos-traço quanto de nutrientes de plantas. Os resultados inferenciais assintóticos, com base na aproximação linear por série de Taylor, foram validados pela metodologia “bootstrap” e pela medida de não linearidade de Bates e Watts, para todos os extratores, exceto para o extrator NH_4OAc , que apresentou problemas quanto à medida de curvatura de efeito parâmetro.

Palavras chaves: regressão não linear, elementos-traço, lodo de esgoto, “bootstrap”, curvatura de Bates e Watts.

¹ **Comitê orientador:** Joel Augusto Muniz – UFLA (Orientador) e Giuliano Marchi – EMBRAPA - Cerrados (Co-orientador).

ABSTRACT

SOUZA, Edcarlos Miranda. **Nonlinear modelling of zinc extraction from sewage sludge-treated soil**. 2008. 95 p. Dissertation (Master of Statistics and Agricultural Experimentation) – Federal University of Lavras, Lavras, MG.¹

Nonlinear regression models are of great importance for scientists and researchers. These models are important as they adjust easily to research data and present practical interpretation of parameters. In soil science studies, assess to trace elements availability from sewage sludge-treated soil in a simple chemical extraction is hardly achieved, reason why Marchi (2005) proposed a method which tried to simulate the rizosphere solution with the use of organic acid mixtures. Trace elements, such as Cd, Zn, Cr, Ni, and Cu, were successively extracted from soil in order to assess the potential availability in the long term. A nonlinear model based in first order kinetics mechanisms was adjusted to data obtained with the successive extractions. The present study has been carried out in the strict sense on estimators statistic properties, as well as adjustments, confidence intervals and confidence bands for expected responses, and joint confidence region. A sketch on the model nonlinearity, having the Bates and Watts curvature measure as a reference, has been carried out in order to check the estimator asymptotic properties, despite the sample reduced size. In the present work only zinc has been considered. Classical methodology for sampling and likelihood was used, having as an alternative the bootstrap technique. The present study provides a detailed script and elaborated routine in 2.7.0 software R version which will guide future works on successive extractions of trace elements and plant nutrients. The asymptotic inferential results based on the linear approach by Taylor series were validated through the bootstrap methodology, and by Bates and Watts nonlinearity measure for all zinc extractors, except for zinc extracted by NH_4OAc . The latter presented problems concerning the curvature measure of parameter effect.

Key words: nonlinear regression, trace element, sewage sludge, bootstrap, Bates and Watts curvature.

¹ **Guidance committee:** Joel Augusto Muniz – UFLA (Adviser) e Giuliano Marchi – EMBRAPA - Cerrados (Adviser).

1 INTRODUÇÃO

Nas ciências aplicadas, especialmente na ciência do solo, utilizam-se modelos úteis na descrição dos dados de análises realizadas por pesquisadores em diversas situações. Em muitas dessas situações, os modelos utilizados possuem parâmetros não lineares e precisam ser avaliados mais profundamente por estatísticos, para a obtenção de resultados inferenciais válidos e confiáveis para a realização das predições.

Em pesquisa ambiental, há necessidade de dados relativos à extração de elementos-traço dos solos e à demanda por extratores capazes de fornecer informações sobre o teor desses elementos nos solos e sua disponibilidade para as plantas (Anderson, 1975). Um bom extrator de elementos-traço do solo deveria ser capaz de extraí-los das mesmas frações que as plantas os extraem, com a mesma intensidade. As plantas não extraem o total de elementos-traço do solo de uma só vez; elas o fazem ao longo do tempo, lentamente dissolvendo, solubilizando, complexando e absorvendo-os da solução do solo. As extrações sucessivas de um elemento-traço, usando solução que imite a rizosfera, poderiam ser capazes de acessar a quantidade potencialmente disponível daquele elemento a longo prazo (Marchi, 2005).

Um caminho para descrever e prever as relações entre eventos que ocorrem na natureza, como no caso de extrações sucessivas de elementos-traço, é a procura de modelos que relacionem as variáveis que descrevam a realidade de forma confiável (Pereira, 2004). A avaliação da relação entre uma variável dependente, como quantidade de metal extraído do solo por um extrator, com uma ou mais variáveis independentes, tais como número de extrações realizadas ao longo do tempo, é tarefa da análise estatística. Isso pode ser feito por meio

dos bem conhecidos modelos de regressão, lineares e não lineares (Mazucheli & Achcar, 2002).

Marchi (2005) utilizou um modelo não linear, conhecido na literatura como modelo de cinética de primeira ordem, com o objetivo de obter a disponibilidade potencial de elementos-traço em solos tratados com lodo de esgoto a longo prazo. O método de extração proposto pelo autor simulou a solução da rizosfera com soluções compostas dos principais íons mono e divalentes e diferentes composições de ácidos orgânicos, para extrair elementos-traço do solo sucessivamente e obter sua disponibilidade potencial a longo prazo. Os dados obtidos pelas extrações sucessivas foram estudados por meio de um modelo não linear de cinética de primeira ordem, observando-se a dissolução de Cd, Zn, Cr, Ni e Cu.

Neste estudo, foram avaliadas estatisticamente as principais características do modelo não linear apresentado por Marchi (2005) na extractabilidade de zinco de solos tratados com lodo de esgoto por ácidos orgânicos, por sete extratores. Foram analisadas as propriedades estatísticas dos estimadores do modelo, determinando intervalos de confiança, estimativas de erros padrões e bandas de confiança para resposta esperada e, também, realizou-se estudo da não linearidade do modelo, considerando a medida de curvatura de Bates & Watts (1980). Além disso, foi utilizada a técnica “bootstrap” paramétrico para a avaliação e o diagnóstico do modelo não linear.

2 REFERENCIAL TEÓRICO

2.1 Ácidos orgânicos e extractabilidade de elementos-traço em solos tratados com lodo de esgoto

Os solos são, por natureza, sistemas multicomponentes muito complexos, nos quais muitas espécies iônicas diferentes em solução e fases sólidas com propriedades de superfície diferentes estão presentes (Kretzchmar & Volgelin, 2001 apud Marchi, 2005). O papel mais importante dos solos é a produção de alimentos, que é chave para a sobrevivência humana. Assim, a manutenção das funções agrícolas e ecológicas dos solos é de responsabilidade da humanidade (Kabata-Pendias & Pendias, 2001).

Um dos objetivos em se estudar o solo é que a persistência de contaminantes neste ambiente é muito maior do que em qualquer outro compartimento da biosfera e a sua contaminação, especialmente por elementos-traço¹, parece ser virtualmente permanente. Várias estimativas têm indicado claramente que a remoção completa de contaminantes dos solos é praticamente impossível (Kabata-Pendias & Pendias, 2001).

Independentemente de suas origens e das razões para o aumento da concentração nos solos, os elementos-traço são responsáveis, em alguns casos, por contaminar a cadeia alimentar devido a migração deles para a solução do solo e conseqüentemente acumulação nas plantas (Bermond & Ghestem, 2001). A contaminação do solo por elementos-traço é um tema importante e, em vista disso, alguns pesquisadores procuram alternativas para o descarte de lodo de esgoto de forma que não haja contaminação do meio ambiente. Alguns deles

¹ Um elemento-traço é qualquer elemento tendo uma medida de concentração menor do que $100 \mu\text{g/g}$ (International Union of Pure and Applied Chemistry - IUPAC, 1979).

acreditam que a incineração é a única maneira de dispor de um lodo de esgoto que contenha altos teores de elementos-traço, mas esta alternativa tem custo elevado e também é uma técnica que polui o ar. Apesar disso, o uso do lodo de esgoto contendo elementos-traço sem prejudicar o meio ambiente é possível. O lodo de esgoto pode ser aplicado na agricultura, sem risco de contaminação do solo, quando se conhecem os teores totais de elementos-traço e de outros nutrientes, como o nitrogênio. Esta prática é preferível ao escoamento de esgoto bruto nos rios, o que causa poluição desastrosa. A legislação brasileira há pouco tempo incentiva a preservação dos rios com a obrigação das cidades em construir estações de tratamento de esgoto. O próximo passo, é o destino deste lodo de esgoto gerado por estas estações (Marchi, 2005).

O destino final para o lodo de esgoto produzido em estações de tratamento urbano-industriais se torna um problema e a quantidade de lodo de esgoto produzida pela atividade humana vem crescendo. Esse lodo pode ser uma fonte significativa de nutrientes de plantas e de matéria orgânica aos solos. Entretanto, as propriedades benéficas dos lodos de esgoto são limitadas pelo teor de substâncias potencialmente prejudiciais, tais como os elementos-traço e de excesso de nutrientes, como o nitrogênio (Alloway, 1995 apud Marchi, 2005).

Em pesquisas sobre contaminação ambiental por lodo de esgoto há a necessidade de dados relativos aos teores de elementos-traço nos solos e grande demanda por extratores capazes de fornecer informações sobre a quantidade disponível desses elementos para as plantas (Anderson, 1975). Os lodos de esgoto de origem municipal e industrial são freqüentemente aplicados aos solos para se dar um destino a esse material e para fornecer nutrientes às plantas. Entretanto, essa prática deve seguir rígidas normas, pois muitos lodos apresentam concentrações elevadas de elementos-traço (Evans et al., 1995). A extração desses elementos dos solos, para fins de avaliação, é uma ferramenta importante para se conhecer sua disponibilidade potencial para as plantas. Dessa

forma, o uso de modelos estatísticos na predição deste tipo de extração tem seu uso bastante difundido.

Os modelos que surgem nas diversas áreas são construídos, em geral, de forma empírica. Tais modelos são equações matemáticas que podem ser ajustadas a resultados experimentais. Esses modelos podem ser utilizados para avaliar ou prever fenômenos observados e auxiliar no desenvolvimento de práticas adequadas de manejo de solo (Camargo et al., 2002 apud Pereira et al., 2005).

Bermond & Ghestem (2001) apresentaram um modelo não linear de cinética de primeira ordem, capaz de relacionar a quantidade de metais extraídos cumulativamente após sucessivas extrações de elementos-traço realizadas por um determinado extrator. Estes extratores podem ser químicos como também naturais, como por exemplo a planta. O modelo determinístico apresentado pelos autores e utilizado por Marchi (2005) é da forma:

$$C_t = C_l (1 - e^{-kt}) \quad (1)$$

em que C_t representa os metais removidos cumulativamente (mg kg^{-1}) na extração t ; C_l é a quantidade de metal extraído até o equilíbrio (mg kg^{-1}); t é o número de extrações (ε); k é uma constante positiva representando a taxa de extração (ε^{-1}); e é o número de Euler, base dos logaritmos neperianos.

2.2 Modelos de regressão linear

A análise de regressão consiste na realização de análise estatística com o objetivo de verificar a existência de relação funcional significativa entre variáveis. Em outras palavras, consiste na obtenção de uma equação que explica

a variação da variável dependente pela variação dos níveis da(s) variável(is) independente(s) (Charnet et al., 1999). A equação estimada com base numa amostra apenas estabelece uma relação funcional entre a variável dependente e a(s) variável(eis) independente(s), para representar o fenómeno em estudo. Portanto, a simples obtenção da equação estimada não responde ao pesquisador se a variação da variável independente influencia significativamente a variação da(s) variável(eis) dependente(s).

É necessário realizar um teste estatístico para estimativas dos coeficientes da equação de regressão estimada. Um teste que pode ser realizado para verificar tal fato é o teste F da análise de variância, proposto por Fisher e Snedecor (Bates & Watts, 1988). Para os autores, a regressão linear fornece estimativas e outros resultados de inferências para os parâmetros $\beta = (\beta_1, \beta_2, \dots, \beta_p)^T$ no modelo:

$$Y_n = \beta_1 x_{n1} + \beta_2 x_{n2} + \dots + \beta_p x_{np} + \xi_n = (x_{n1}, x_{n2}, \dots, x_{np}) \beta + \xi_n \quad (2)$$

Neste modelo, a variável aleatória Y_n , que representa a resposta para o caso n , $n = 1, 2, \dots, N$, tem uma parte determinística e uma parte estocástica. A parte determinística, $(x_{n1}, x_{n2}, \dots, x_{np}) \beta$, depende dos parâmetros e das variáveis preditoras ou variáveis regressoras x_{np} , $p = 1, 2, \dots, P$. A parte estocástica, representada pela variável aleatória ξ_n , é um distúrbio que desordena a resposta para esse caso (alguns autores utilizam o termo erro aleatório em vez de distúrbio). O modelo linear (2) pode ser escrito na forma matricial como:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e} \quad (3)$$

no qual $\mathbf{y}$ é o vetor das variáveis aleatórias representando o vetor com os dados observados, $\mathbf{X}$ é a matriz de dimensões $N \times P$, das variáveis regressoras e tem posto coluna completo, sendo N o número de observações e P o número de parâmetros do modelo

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & x_{13} & \cdots & x_{1P} \\ x_{21} & x_{22} & x_{23} & \cdots & x_{2P} \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ x_{N1} & x_{N2} & x_{N3} & \cdots & x_{NP} \end{bmatrix}$$

e $\mathbf{e}$ é o vetor das variáveis aleatórias não observáveis representando os erros aleatórios.

A parte determinística, $\mathbf{X}\boldsymbol{\beta}$, uma função de parâmetros e variáveis regressoras, fornece o modelo matemático ou a função do modelo para as respostas. Para o modelo (3) supõe-se em geral, que:

$$E[\mathbf{e}] = \mathbf{0}$$

ou

$$E[\mathbf{y}] = \mathbf{X}\boldsymbol{\beta} \quad (4)$$

Assim, $\mathbf{X}\boldsymbol{\beta}$ é chamada a função esperança para o modelo de regressão. A matriz $\mathbf{X}$ é chamada de matriz derivada, tendo em vista que ela é obtida derivando-se o modelo em relação $\boldsymbol{\beta}$.

Supondo que o $\mathbf{e}$ é normalmente distribuído com:

$$\text{Var}[\mathbf{e}] = E[\mathbf{e}\mathbf{e}^T] = \sigma^2 \mathbf{I} \quad (5)$$

em que $\mathbf{I}$ é uma matriz identidade $N \times N$, então, a função densidade de probabilidade conjunta para o $\mathbf{y}$, dado o $\boldsymbol{\beta}$ e a variância σ^2 , é:

$$\begin{aligned} p(\mathbf{y} / \boldsymbol{\beta}, \sigma^2) &= (2\pi\sigma^2)^{-N/2} \exp\left(\frac{-(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^T (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})}{2\sigma^2}\right) \\ &= (2\pi\sigma^2)^{-N/2} \exp\left(\frac{-\|\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\|^2}{2\sigma^2}\right) \end{aligned} \quad (6)$$

em que as duas barras verticais denotam a norma de um vetor. Embora produzida com uma matriz derivada $\mathbf{X}$ e um vetor $\mathbf{y}$ de dados observados, deseja-se fazer inferências sobre σ^2 e os P parâmetros de $\boldsymbol{\beta}$ (Jorgensen, 1993).

2.2.1 Os estimadores de mínimos quadrados e de máxima verossimilhança

A função de verossimilhança, ou, simplesmente, a verossimilhança, $l(\boldsymbol{\beta}, \sigma / \mathbf{y})$, para $\boldsymbol{\beta}$ e σ , é idêntica na forma da densidade de probabilidade conjunta, exceto que $l(\boldsymbol{\beta}, \sigma / \mathbf{y})$ é considerado como uma função de parâmetros

condicional aos dados observados, de preferência como uma função de respostas condicional aos valores dos parâmetros (Mood et al., 1974). Suprimindo-se a constante $(2\pi)^{-N/2}$, escreve-se:

$$l(\boldsymbol{\beta}, \sigma / \mathbf{y}) \propto \sigma^{-N} \exp\left(\frac{-\|\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\|^2}{2\sigma^2}\right) \quad (7)$$

A verossimilhança é maximizada com relação ao $\boldsymbol{\beta}$ quando a soma de quadrados residual:

$$S(\boldsymbol{\beta}) = \|\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\|^2 = \sum_{n=1}^N \left[y_n - \left(\sum_{p=1}^P x_{np} \beta_p \right) \right]^2 \quad (8)$$

é mínima (Jorgensen, 1993). Assim, o valor que torna a verossimilhança máxima $\hat{\boldsymbol{\beta}}$ é o valor de $\boldsymbol{\beta}$ que minimiza $S(\boldsymbol{\beta})$. Dessa forma, o estimador de máxima verossimilhança é também o estimador de mínimos quadrados e pode ser escrito como (Draper & Smith, 1998):

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y} \quad (9)$$

A existência de $(\mathbf{X}^T \mathbf{X})^{-1}$ decorre do fato de $\mathbf{X}$ ter posto coluna completo.

É importante perceber que, embora o estimador de mínimos quadrados não necessita das mesmas pressuposições que o estimador de máxima verossimilhança para o seu cálculo algébrico ou numérico, para a inferência, ele só se torna viável quando o modelo (3) e as suposições no termo do erro

aleatório (4) e (5) são válidas. Expressando de outra maneira, o que se deseja dizer é que, ao se usar as estimativas de mínimos quadrados, supõe-se que (Bates & Watts, 1988):

- i. a função esperança é correta;
- ii. a resposta é a função esperança mais o erro;
- iii. o erro é independente da função esperança;
- iv. cada erro tem uma distribuição normal;
- v. cada erro tem média zero;
- vi. os erros têm variâncias iguais;
- vii. os erros são independentemente distribuídos.

O estimador de mínimos quadrados tem inúmeras propriedades desejáveis, como mostrado, por exemplo, em Seber (1977):

- i. o estimador de mínimos quadrados $\hat{\beta}$ é normalmente distribuído. Isto se dá porque o estimador é uma função linear de y que, por sua vez, é uma função linear de e . Já que o e é suposto ser normalmente distribuído, o $\hat{\beta}$ é normalmente distribuído;
- ii. $E[\hat{\beta}] = \beta$: o estimador de mínimos quadrados é não viesado;
- iii. $\text{var}[\hat{\beta}] = \sigma^2 (\mathbf{X}^T \mathbf{X})^{-1}$: a matriz de covariância do estimador de mínimos quadrados depende da variância dos distúrbios e da matriz derivada $\mathbf{X}$;
- iv. uma região de confiança conjunta $1 - \alpha$ para o parâmetro β é a elipse:

$$(\beta - \hat{\beta})^T \mathbf{X}^T \mathbf{X} (\beta - \hat{\beta}) \leq P s^2 F_{(P, N-P; \alpha)}, \text{ em que } s^2 = \frac{S(\hat{\beta})}{N - P}$$

é o quadrado médio residual ou o estimador da variância associada com $N - P$ número de graus de liberdade e $F_{(P, N-P; \alpha)}$ é o quantil α superior para a distribuição F , com P e $N - P$ graus de liberdade;

- v. um intervalo de confiança marginal $1 - \alpha$ para o parâmetro β_p é:

$$\hat{\beta}_p \pm se(\hat{\beta}_p) t_{(N-P; \alpha/2)}$$

em que $t_{(N-P; \alpha/2)}$ é o quantil $\alpha/2$ superior para a distribuição t de Student com $N - P$ número de graus de liberdade e, o erro padrão do estimador do parâmetro é:

$$se(\hat{\beta}_p) = s \sqrt{\left\{ (\mathbf{X}^T \mathbf{X})^{-1} \right\}_{pp}}$$

com $\left\{ (\mathbf{X}^T \mathbf{X})^{-1} \right\}_{pp}$ igual ao p -ésimo termo diagonal da matriz $(\mathbf{X}^T \mathbf{X})^{-1}$;

- vi. um intervalo de confiança $1 - \alpha$ para a resposta esperada em $\mathbf{x}_0$ é:

$$\mathbf{x}_0^T \hat{\boldsymbol{\beta}} \pm s \sqrt{\mathbf{x}_0^T (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{x}_0} t_{(N-P; \alpha/2)};$$

- vii. uma banda de confiança $1 - \alpha$ para a função da resposta em qualquer vetor $\mathbf{x}$ é dada por:

$$\mathbf{x}^T \hat{\boldsymbol{\beta}} \pm s \sqrt{\mathbf{x}^T (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{x}} \sqrt{PF_{(P, N-P; \alpha)}}$$

Bates & Watts (1988) destacam que a função de verossimilhança $l(\boldsymbol{\beta}, \sigma / \mathbf{y})$ depende do $\boldsymbol{\beta}$ apenas pelo $\|\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\|$, portanto, os contornos da verossimilhança são da forma:

$$\|\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\|^2 = c \quad (10)$$

em que c é uma constante. A região de verossimilhança é limitada pelo contorno para a qual:

$$c = S(\hat{\beta}) \left[1 + \frac{P}{N-P} F_{(P, N-P; \alpha)} \right] \quad (11)$$

é idêntica a uma região de confiança conjunta $1-\alpha$ da abordagem teórica de amostragem. No entanto, a interpretação de uma região de verossimilhança é bem diferente daquela de uma região de confiança.

A linearidade ou a não-linearidade a que se refere o texto são inerentes aos parâmetros. Um modelo é dito linear nos parâmetros se qualquer uma das derivadas parciais desse modelo em relação aos parâmetros não depender especificamente de nenhum deles (Draper & Smith, 1998). O modelo de regressão linear múltipla de hiperplano com quatro variáveis regressoras é um exemplo de modelo linear:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \xi$$

$$\frac{\partial Y}{\partial \beta_0} = 1; \frac{\partial Y}{\partial \beta_1} = X_1; \frac{\partial Y}{\partial \beta_2} = X_2; \frac{\partial Y}{\partial \beta_3} = X_3 \text{ e } \frac{\partial Y}{\partial \beta_4} = X_4$$

2.3 Modelos de regressão não linear

A definição de um modelo não linear apresentada pela maioria dos autores, como Draper & Smith (1998), Bates & Watts (1988), Ratkowsky (1983) e outros, é que pelo menos uma derivada parcial da variável dependente, com relação a algum parâmetro presente no modelo, depende de algum parâmetro.

Gallant (1987) apresentou uma definição mais geral, em que um modelo linear é apenas um caso particular de um modelo não linear, assim como uma distribuição de qui-quadrado central é apenas um caso particular de uma distribuição de qui-quadrado não central. Mesmo assim, o autor trabalha no escopo de seu livro com a definição mais usual apresentada pela maioria dos autores, justificando que a mesma já está consagrada na literatura vigente.

2.3.1 Modelos de regressão intrinsecamente lineares

Gujarati (2006) afirma que alguns modelos podem parecer não lineares nos parâmetros, mas são *inerente* ou *intrinsecamente* lineares porque, com as devidas transformações, podem se tornar modelos de regressão lineares em novos parâmetros. O autor detalha esse conceito, utilizando a função de produção de Cobb-Douglas. Fazendo Y = produção, X_2 = mão-de-obra, X_3 = capital e u_i o componente do erro aleatório, pode-se escrever esta função de três modos diferentes¹:

1º modo:

$$Y_i = \beta_1 X_{2i}^{\beta_2} X_{3i}^{\beta_3} e^{u_i} \quad (a)$$

Aplicando o logaritmo neperiano em ambos os lados da igualdade (a), o modelo passou a:

$$W_i = \alpha + \beta_2 \ln X_{2i} + \beta_3 \ln X_{3i} + u_i \quad (b)$$

com $\alpha = \ln \beta_1$ e $W_i = \ln Y_i$, e as seguintes derivadas parciais em relação aos novos parâmetros

$$\frac{\partial W_i}{\partial \alpha} = 1; \quad \frac{\partial W_i}{\partial \beta_2} = \ln X_{2i} \quad \text{e} \quad \frac{\partial W_i}{\partial \beta_3} = \ln X_{3i}$$

¹ No exemplo (a), e representa o número de Euler, base dos logaritmos neperianos.

não dependem dos parâmetros, o que de acordo com a definição de linearidade, indica que o modelo (b) é linear segundo os parâmetros α , β_2 e β_3 e, portanto, o modelo (a) é um modelo intrinsecamente linear.

2º modo:

$$Y_i = \beta_1 X_{2i}^{\beta_2} X_{3i}^{\beta_3} u_i \quad (c)$$

Aplicando-se o logaritmo neperiano em ambos os lados da igualdade (c), obtém-se o modelo:

$$W_i = \alpha + \beta_2 \ln X_{2i} + \beta_3 \ln X_{3i} + \ln u_i \quad (d)$$

com $\alpha = \ln \beta_1$ e $W_i = \ln Y_i$. As seguintes derivadas parciais em relação aos novos parâmetros são:

$$\frac{\partial W_i}{\partial \alpha} = 1; \quad \frac{\partial W_i}{\partial \beta_2} = \ln X_{2i} \quad \text{e} \quad \frac{\partial W_i}{\partial \beta_3} = \ln X_{3i}$$

E também, de acordo com a definição de linearidade, pode-se concluir que o modelo (d) é linear segundo os parâmetros α , β_2 e β_3 e, portanto, o modelo (c) é um modelo intrinsecamente linear.

3º modo:

$$Y_i = \beta_1 X_{2i}^{\beta_2} X_{3i}^{\beta_3} + u_i \quad (e)$$

As derivadas parciais de Y em relação aos parâmetros dependem dos mesmos:

$$\frac{\partial Y_i}{\partial \beta_1} = X_{2i}^{\beta_2} X_{3i}^{\beta_3}; \quad \frac{\partial Y_i}{\partial \beta_2} = \beta_1 X_{2i}^{\beta_2} X_{3i}^{\beta_3} \ln(X_{2i}) \quad \text{e} \quad \frac{\partial Y_i}{\partial \beta_3} = \beta_1 X_{2i}^{\beta_2} X_{3i}^{\beta_3} \ln(X_{3i})$$

O modelo (e) não satisfaz a definição de um modelo de regressão linear e não há jeito de transformá-lo em um modelo linear nos parâmetros.

2.3.2 O problema das transformações

Na prática, um modelo não-linear é linearizado para facilitar a obtenção das estimativas dos parâmetros (Bates & Watts, 1988). O inconveniente de uma transformação é que, além de o parâmetro perder sua interpretação intrínseca, pode-se alterar a estrutura e a distribuição do erro. Ou seja, se os erros do modelo original satisfazem às suposições usuais de normalidade, independência e homogeneidade da variância, os erros do novo modelo, em geral, não satisfarão tais suposições porque tais transformações podem mudar a forma da distribuição dos mesmos ou fazer com que a variância desta distribuição deixe de ser constante (Currie, 1982).

Pode-se observar isso no 2º modo da função de produção de Cobb-Douglas, equação (d), em que o erro do modelo é o logaritmo neperiano do erro do modelo anterior.

2.3.3 Modelos de regressão intrinsecamente não lineares

Se um modelo de regressão não é linear e nem intrinsecamente linear nos parâmetros, então, ele é chamado de intrinsecamente não-linear nos parâmetros (Draper & Smith, 1998). De agora em diante, toda vez que aqui for citado o termo modelo não-linear, estará subentendido que esse modelo será intrinsecamente não linear nos parâmetros.

Gallant (1987) afirma que uma das mais comuns situações em análise estatística é aquela de dados observados que consistem em respostas univariadas y_t , conhecidas sendo dependentes com contribuição x_t . Essa situação pode ser representada por equações de regressão:

$$y_t = f(x_t, \boldsymbol{\theta}) + e_t, \quad t = 1, 2, \dots, N \quad (12)$$

em que $\boldsymbol{\theta} = (\theta_1, \theta_2, \dots, \theta_p)$ é um vetor P-dimensional de parâmetros desconhecidos e e_t é o erro aleatório aditivo, representando fenômenos não observáveis ou erro experimental. Os erros são assumidos independentes e identicamente distribuídos com média zero e variância desconhecida σ^2 .

A sequência de valores x_t também é chamada de ajuste experimental, de forma que o vetor de parâmetros e o ajuste experimental juntos determinam o valor da função modelo $f(x_t, \boldsymbol{\theta})$ que, sobre as suposições de que $E[e_t] = 0$, é o valor esperado de y_t condicional a x_t e $\boldsymbol{\theta}$ (Bates & Watts, 1980):

$$E[y_t / x_t, \boldsymbol{\theta}] = f(x_t, \boldsymbol{\theta})$$

A função modelo é freqüentemente chamada de o modelo ou o modelo matemático, mas existe uma distinção que deve ser feita, tendo em vista que funções modelos diferentes podem ser associadas com o mesmo modelo. A função modelo é associada com a parametrização particular escolhida para o modelo, então, um conjunto diferente de parâmetros pode produzir uma função modelo diferente. O modelo é a relação que dá suporte entre a resposta e o ajuste experimental gerado pela função modelo e é independente da parametrização.

Define-se a soma de quadrados dos erros para o modelo não linear fornecido pelos dados como (Draper & Smith, 1998):

$$S(\boldsymbol{\theta}) = \sum_{t=1}^N \{y_t - f(x_t, \boldsymbol{\theta})\}^2 \quad (13)$$

Desde que y_t e x_t sejam observações fixas, a soma dos quadrados é uma função de $\boldsymbol{\theta}$. A estimativa de mínimos quadrados $\hat{\boldsymbol{\theta}}$ de $\boldsymbol{\theta}$ é o valor que minimiza $S(\boldsymbol{\theta})$. Pode ser mostrado que a estimativa de mínimos quadrados de $\boldsymbol{\theta}$ é também o estimador de máxima verossimilhança de $\boldsymbol{\theta}$ (Draper & Smith, 1998). Para encontrar o estimador de mínimos quadrados $\hat{\boldsymbol{\theta}}$, é necessário diferenciar a equação (13) com relação a $\boldsymbol{\theta}$. Isso gera P equações normais que precisam ser resolvidas para $\hat{\boldsymbol{\theta}}$. As equações normais podem ser tomadas na forma:

$$\sum_{t=1}^N \{y_t - f(x_t, \boldsymbol{\theta})\}^2 \left[\frac{\partial f(x_t, \boldsymbol{\theta})}{\partial \theta_i} \right]_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} = 0 \quad (14)$$

para $i = 1, 2, \dots, P$, em que a quantidade denotada pelos colchetes são as derivadas de $f(x_t, \boldsymbol{\theta})$ com os respectivos θ_i com todos os θ 's substituídos pelos correspondentes $\hat{\theta}$'s, que têm o mesmo subscrito.

Considerando, para cada ajuste experimental x_t , a resposta esperada condicional:

$$\eta_t(\boldsymbol{\theta}) = E[y_t / \boldsymbol{\theta}] = f(x_t, \boldsymbol{\theta}) \quad (15)$$

e os vetores $\boldsymbol{\eta}(\boldsymbol{\theta}) = (\eta_1(\boldsymbol{\theta}), \eta_2(\boldsymbol{\theta}), \dots, \eta_N(\boldsymbol{\theta}))^T$, $\mathbf{y} = (y_1, y_2, \dots, y_N)^T$ e $\mathbf{e} = (e_1, e_2, \dots, e_N)^T$, então, pode-se escrever o modelo em forma vetorial como:

$$\mathbf{y} = \boldsymbol{\eta}(\boldsymbol{\theta}) + \mathbf{e} \quad (16)$$

e a função soma de quadrados (14) como:

$$S(\boldsymbol{\theta}) = \|\mathbf{y} - \boldsymbol{\eta}(\boldsymbol{\theta})\|^2 \quad (17)$$

na qual as duas barras verticais indicam a norma de um vetor. Assim, $S(\boldsymbol{\theta})$ tem uma interpretação geométrica de ser o quadrado da distância entre o vetor $\mathbf{y}$ e o vetor $\boldsymbol{\eta}(\boldsymbol{\theta})$, em um espaço amostral N -dimensional.

À medida que $\boldsymbol{\theta}$ é modificado, o vetor $\boldsymbol{\eta}(\boldsymbol{\theta})$ descreve uma superfície P -dimensional, chamada de “o local da solução”, nesse espaço amostral (Box & Lucas, 1959). Em geral, $\mathbf{y}$ não será um ponto no local da solução e a estimativa de mínimos quadrados $\hat{\boldsymbol{\theta}}$ é um ponto no local da solução próximo de $\mathbf{y}$.

O método dos mínimos quadrados é utilizado na estimação dos parâmetros em modelos não lineares, da mesma maneira que em modelos lineares. Porém, para o sistema de equação normais não linear, não existe uma solução explícita, mas sim, uma seqüência de soluções aproximadas que são obtidas por meio de processos iterativos por aproximação em série de Taylor, usando um software estatístico (Gallant, 1987).

2.4 Estimativa dos mínimos quadrados não lineares

A maioria dos algoritmos para computação das estimativas de mínimos quadrados $\hat{\boldsymbol{\theta}}$ e a maioria dos métodos inferenciais para modelos não lineares são

baseados em métodos iterativos que consideram uma aproximação linear local para o modelo (Chambers, 1973 apud Bates & Watts, 1980).

2.4.1 Aproximação Linear

Na vizinhança de um valor de parâmetro $\boldsymbol{\theta}^0$, a função modelo é aproximada por:

$$f(x_t, \boldsymbol{\theta}) \cong f(x_t, \boldsymbol{\theta}^0) + \sum_{i=1}^P (\theta_i - \theta_i^0) v_i(x_t)$$

em que:

$$v_i(x_t) = \left[\frac{\partial f(x_t, \boldsymbol{\theta})}{\partial \theta_i} \right]_{\boldsymbol{\theta}=\boldsymbol{\theta}^0}$$

Reunindo-se os componentes dentro dos vetores $\boldsymbol{\eta}(\boldsymbol{\theta})$, $\boldsymbol{\eta}(\boldsymbol{\theta}^0)$ e $\mathbf{v}_i$, em que:

$$\begin{aligned} \mathbf{v}_i &= (v_i(x_1), v_i(x_2), \dots, v_i(x_n))^T \\ &= \left[\frac{\partial \boldsymbol{\eta}}{\partial \theta_i} \right]_{\boldsymbol{\theta}=\boldsymbol{\theta}^0}, \quad i = 1, 2, \dots, P \end{aligned}$$

tem-se:

$$\boldsymbol{\eta}(\boldsymbol{\theta}) \cong \boldsymbol{\eta}(\boldsymbol{\theta}^0) + \sum_{i=1}^P (\theta_i - \theta_i^0) \mathbf{v}_i$$

ou

$$\boldsymbol{\eta}(\boldsymbol{\theta}) \cong \boldsymbol{\eta}(\boldsymbol{\theta}^0) + \mathbf{V}^0 (\boldsymbol{\theta} - \boldsymbol{\theta}^0) \quad (18)$$

em que $\mathbf{V}^0$ é a matriz de derivadas avaliada em $\boldsymbol{\theta}^0$, tendo como colunas os vetores $\mathbf{v}_i$ s ordenadamente. A matriz $\mathbf{V}^0$ também é chamada de a matriz jacobiana com ordem $N \times P$ de $\boldsymbol{\eta}(\boldsymbol{\theta})$ avaliada em $\boldsymbol{\theta}^0$:

$$\mathbf{V}^0 = \left[\begin{array}{cccc} \frac{\partial \eta_1(\boldsymbol{\theta})}{\partial \theta_1} & \frac{\partial \eta_1(\boldsymbol{\theta})}{\partial \theta_2} & \cdots & \frac{\partial \eta_1(\boldsymbol{\theta})}{\partial \theta_p} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial \eta_n(\boldsymbol{\theta})}{\partial \theta_1} & \frac{\partial \eta_n(\boldsymbol{\theta})}{\partial \theta_2} & \cdots & \frac{\partial \eta_n(\boldsymbol{\theta})}{\partial \theta_p} \end{array} \right]_{\boldsymbol{\theta}=\boldsymbol{\theta}^0}$$

Bates & Watts (1980) afirmam que o efeito dessa aproximação é substituir o local da solução pelo seu plano tangente em $\boldsymbol{\eta}(\boldsymbol{\theta}^0)$ e, simultaneamente, impor um sistema de coordenadas uniformes nesse plano tangente. Esses dois componentes da aproximação linear são chamados “a pressuposição planar” e a “pressuposição de coordenadas uniformes”, respectivamente. Tanto a efetividade do algoritmo de mínimos quadrados quanto a validade das inferências feitas com respeito aos parâmetros de um modelo não linear serão afetadas pela proximidade da linearidade do modelo.

2.4.2 O método de Gauss-Newton

A abordagem sugerida por Gauss foi utilizar uma aproximação linear à função esperança, como em (18), para reiterativamente melhorar o $\boldsymbol{\theta}^0$ como valor inicial para o $\boldsymbol{\theta}$ e continuar melhorando as estimativas, até não existir

mais mudança. Segundo Ratkowsky¹ (1983), esse método expande uma série de Taylor sobre $\boldsymbol{\theta}_i$ (o subscrito i indica a i -ésima iteração), como em (18):

$$\boldsymbol{\eta}(\boldsymbol{\theta}) \cong \boldsymbol{\eta}(\boldsymbol{\theta}_i) + \mathbf{V}_i(\boldsymbol{\theta} - \boldsymbol{\theta}_i)$$

Rearranjando-se (18) com (17), tem-se:

$$\begin{aligned} S(\boldsymbol{\theta}) &= \|\mathbf{y} - \boldsymbol{\eta}(\boldsymbol{\theta})\|^2 = [\mathbf{y} - \boldsymbol{\eta}(\boldsymbol{\theta})]^T [\mathbf{y} - \boldsymbol{\eta}(\boldsymbol{\theta})] = \\ &\cong [\mathbf{y} - (\boldsymbol{\eta}(\boldsymbol{\theta}_i) + \mathbf{V}_i(\boldsymbol{\theta} - \boldsymbol{\theta}_i))]^T [\mathbf{y} - (\boldsymbol{\eta}(\boldsymbol{\theta}_i) + \mathbf{V}_i(\boldsymbol{\theta} - \boldsymbol{\theta}_i))] = \\ &= [(\mathbf{y} - \boldsymbol{\eta}(\boldsymbol{\theta}_i)) - \mathbf{V}_i(\boldsymbol{\theta} - \boldsymbol{\theta}_i)]^T [(\mathbf{y} - \boldsymbol{\eta}(\boldsymbol{\theta}_i)) - \mathbf{V}_i(\boldsymbol{\theta} - \boldsymbol{\theta}_i)] = \\ &= [\mathbf{y} - \boldsymbol{\eta}(\boldsymbol{\theta}_i)]^T [\mathbf{y} - \boldsymbol{\eta}(\boldsymbol{\theta}_i)] \\ &- 2. [\mathbf{y} - \boldsymbol{\eta}(\boldsymbol{\theta}_i)]^T \mathbf{V}_i(\boldsymbol{\theta} - \boldsymbol{\theta}_i) + (\boldsymbol{\theta} - \boldsymbol{\theta}_i)^T \mathbf{V}_i^T \mathbf{V}_i(\boldsymbol{\theta} - \boldsymbol{\theta}_i) \end{aligned}$$

O vetor gradiente $g(\boldsymbol{\theta}) = \left(\frac{\partial S(\boldsymbol{\theta})}{\partial \theta_1}, \frac{\partial S(\boldsymbol{\theta})}{\partial \theta_2}, \dots, \frac{\partial S(\boldsymbol{\theta})}{\partial \theta_p} \right)^T$ é, então:

$$g(\boldsymbol{\theta}) = -2\mathbf{V}_i^T [\mathbf{y} - \boldsymbol{\eta}(\boldsymbol{\theta}_i)] + 2\mathbf{V}_i^T \mathbf{V}_i(\boldsymbol{\theta} - \boldsymbol{\theta}_i)$$

Igualando-se essa expressão a zero e rearranjando-se os valores, obtém-se:

$$\boldsymbol{\theta}_{i+1} = \boldsymbol{\theta}_i + [\mathbf{V}_i^T \mathbf{V}_i]^{-1} \mathbf{V}_i^T [\mathbf{y} - \boldsymbol{\eta}(\boldsymbol{\theta}_i)] \quad (19)$$

Com valores iniciais para $\boldsymbol{\theta}$ com $i=1$, o processo continua até a convergência que ocorre quando $\|\boldsymbol{\theta}_{i+1} - \boldsymbol{\theta}_i\|$ é menor do que alguma quantidade pequena pré-fixada.

¹ As notações apresentadas aqui não são idênticas às utilizadas por Ratkowsky (1983). Algumas adequações foram realizadas.

Draper & Smith (1998) discutem alguns problemas do método da linearização:

- i. ele pode convergir muito lentamente, isto é, um grande número de iterações pode ser requerido porque a solução estabiliza, mesmo que a soma de quadrados $S(\boldsymbol{\theta}_i)$ cresça de forma constante à medida que i cresce. Esse tipo de comportamento não é comum, mas pode ocorrer;
- ii. pode oscilar bastante, continuamente mudando de direção e, muitas vezes, de modo crescente, decrescendo a soma de quadrados. Todavia, a solução pode estabilizar eventualmente;
- iii. pode não convergir sob qualquer condição e, inalteradamente, divergir.

Observe que o método de Gauss-Newton serve também para encontrar as estimativas em modelo de regressão linear, tendo em vista que tal modelo pode ser escrito da forma:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\theta} + \mathbf{e}$$

em que $\mathbf{X}$ é a matriz $N \times P$ de variáveis regressoras, cuja primeira coluna pode ser um vetor unitário, se a função modelo inclui um termo constante. Para a função modelo acima, a matriz jacobiana é:

$$\mathbf{V}_0 = \mathbf{X}$$

Para qualquer valor arbitrário utilizado como estimativa inicial, por exemplo, $\boldsymbol{\theta}_0$, o vetor de estimativas atualizado $\boldsymbol{\theta}_1$ é:

$$\begin{aligned}
\boldsymbol{\theta}_{0+1} &= \boldsymbol{\theta}_0 + [\mathbf{V}_0^T \mathbf{V}_0]^{-1} \mathbf{V}_0^T [\mathbf{y} - \boldsymbol{\eta}(\boldsymbol{\theta}_0)] \\
&= \boldsymbol{\theta}_0 + [\mathbf{X}^T \mathbf{X}]^{-1} \mathbf{X}^T [\mathbf{y} - \mathbf{X}\boldsymbol{\theta}_0] \\
&= \boldsymbol{\theta}_0 + [\mathbf{X}^T \mathbf{X}]^{-1} \mathbf{X}^T \mathbf{y} - [\mathbf{X}^T \mathbf{X}]^{-1} [\mathbf{X}^T \mathbf{X}] \boldsymbol{\theta}_0 \\
&= \boldsymbol{\theta}_0 + [\mathbf{X}^T \mathbf{X}]^{-1} \mathbf{X}^T \mathbf{y} - \boldsymbol{\theta}_0 \\
&= [\mathbf{X}^T \mathbf{X}]^{-1} \mathbf{X}^T \mathbf{y}
\end{aligned}$$

É óbvio que a estimativa fica imutável se o processo iterativo é continuado. O lado direito da igualdade é o estimador de mínimos quadrados para um modelo de regressão linear (Draper & Smith, 1998). Assim, o método de Gauss-Newton, para um modelo linear, converge para o estimador de mínimos quadrados em uma única iteração para qualquer valor inicial.

Alem do método de Gauss-Newton, existem vários métodos iterativos propostos na literatura para a obtenção das estimativas de mínimos quadrados dos parâmetros de um modelo de regressão não-linear. Dentre eles podem-se destacar o método *steepest-descent*, ou método Gradiente, o método de Marquardt e o método de Newton-Raphson. Em geral, esses métodos podem ser utilizados quando o método de Gauss-Newton apresenta problemas com relação à convergência (Bates & Watts, 1988).

2.4.3 Convergência

Souza (1998) destaca que o sucesso na convergência de um algoritmo para um método iterativo no processo de estimação não linear está diretamente associado ao uso de uma função resposta apropriada e de valores iniciais adequados ao procedimento numérico.

Gallant (1987) afirma que o sucesso, em termos de convergência para $\hat{\boldsymbol{\theta}}$, não é garantido por qualquer método. A experiência indica que a falha no

processo de convergência (se as derivadas estiverem corretas) depende da distância do valor inicial à resposta correta e do grau de parametrização da função resposta, relativamente ao conjunto de dados utilizado. Estes problemas estão inter-relacionados, no sentido de que funções resposta mais apropriadas conduzem a uma amplitude maior do que o raio de convergência (no espaço paramétrico). Quando a convergência não ocorre, deve-se tentar encontrar valores iniciais melhores (mais próximos de $\hat{\theta}$) ou utilizar uma função resposta mais parcimoniosa (com menos parâmetros).

Bates & Watts (1988) afirmam que se deve checar se os dados que estão sendo analisados são válidos, isto é, o pesquisador deve sempre se certificar se os valores numéricos corretos da resposta e variáveis preditoras foram introduzidos de forma certa na máquina. Uma boa alternativa é plotar a resposta *versus* variável preditora, e observar o comportamento das variáveis envolvidas. Reparametrizações apropriadas podem também facilitar a convergência. Outro problema que pode acontecer é que o mínimo encontrado pode não ser um mínimo global e, neste caso, toda a inferência estaria comprometida. Não existem garantias matemáticas para verificar esse fato, então, uma sugestão é observar se os valores encontrados condizem com os dados propostos, traçando a curva ajustada com os dados e verificando se não há nenhum resultado absurdo.

2.4.4 Valores iniciais

Em conformidade com Bates & Watts (1988), uma das melhores alternativas para assegurar que uma análise não linear é bem sucedida é obter bons valores iniciais para os parâmetros, valores dos quais a convergência é obtida rapidamente. Para Draper & Smith (1998), os valores iniciais podem ser estimativas preliminares baseadas em qualquer informação disponível, por

exemplo, eles podem ser valores sugeridos por informação ganha em um ajuste similar de equações em um laboratório ou sugerida por experiências anteriores.

Esses valores serão improvisados em iterações sucessivas, com a esperança de serem as mais baixas descrições, ou seja, quanto melhor eles forem, menor será o número de iterações necessárias para se encontrar a convergência. Vários princípios simples, porém úteis para determinar valores iniciais podem ser usados, como sugerido por Bates & Watts (1988) e Draper & Smith (1998):

- i. interpretar o comportamento da função esperança, em termos de parâmetros analiticamente ou graficamente;
- ii. interpretar o comportamento das derivadas da função esperança, em termos dos parâmetros analiticamente ou graficamente;
- iii. transformar a função esperança analiticamente ou geograficamente para obter comportamento mais simples, preferencialmente linear. No caso em que o erro é aditivo, pode-se desconsiderar o mesmo apenas para a obtenção das estimativas iniciais;
- iv. reduzir dimensões substituindo valores para alguns parâmetros ou avaliando a função nos valores do delineamento específico;
- v. usar linearidade condicional (se possível).

2.5 Adequação do modelo

Draper & Smith (1998) sugerem que, em qualquer análise, é necessário avaliar o ajuste do modelo para os dados. Na estimação não linear, é possível convergir para valores do parâmetro que não condizem com resultados aceitáveis na prática. Isto ocorre por que, às vezes, pode-se convergir para um mínimo local.

A avaliação de qualquer modelo deve, então, começar com uma consideração cuidadosa das estimativas dos parâmetros e observar se elas fazem sentido cientificamente. Quando uma função simples e adequada é encontrada, um gráfico de valores ajustados ou preditos é uma excelente maneira de avaliar o ajuste. Caso isso não aconteça, deve-se fazer uma revisão em todas as informações da análise procurando possíveis falhas de análise (Bates & Watts, 1988).

2.5.1 O estimador de σ^2

O estimador de mínimos quadrados e de máxima verossimilhança para σ^2 é dado por (Gallant, 1987):

$$s^2 = \frac{S(\hat{\theta})}{N - P} \quad (20)$$

em que $S(\hat{\theta})$ é a soma de quadrados do resíduo avaliada em $\hat{\theta}$, N é o número de elementos na amostra e P é o número de parâmetros.

2.5.2 Coeficiente de determinação - R^2

Para a avaliação do ajuste, pode ser considerado o coeficiente de determinação R^2 para o modelo de regressão, que é dado por:

$$R^2 = 1 - \frac{S(\hat{\theta})}{(\mathbf{y} - \bar{y} \cdot \mathbf{1})^T (\mathbf{y} - \bar{y} \cdot \mathbf{1})}$$

em que $\mathbf{1} = [1 \ 1 \ \dots \ 1]^T$ de ordem $N \times 1$.

Ressalta-se, porém, que o R^2 não é, isoladamente, um critério adequado para discussão de ajuste de modelos, pois, geralmente, em ajuste de modelos não-lineares, é comum a obtenção de R^2 assintóticos altos e similares (Rezende et al., 2007).

2.6 Considerações importantes sobre $\hat{\theta}$

Uma importante consideração a respeito de um estimador de mínimos quadrados em um modelo não linear é que certas propriedades que são válidas para os modelos lineares agora só são consideradas assintoticamente, isso é, com o tamanho da amostra tendendo ao infinito (Gallant, 1987).

Para amostras finitas, $\hat{\theta}$ não é uma combinação linear de y e, conseqüentemente, em geral, não é normalmente distribuído, é viesado para θ e não é um estimador de variância mínima. Assim, ao contrário de um estimador de mínimos quadrados de um parâmetro em um modelo linear, o caso não linear possui propriedades desconhecidas para amostras de tamanho finito e pequenas (Ratkowsky, 1983).

Não obstante, quando o tamanho n da amostra cresce, pode-se prever como $\hat{\theta}$ tende a se comportar, contanto que os erros sejam normalmente, independentemente e identicamente distribuídos (pois, nesse caso, o estimador de mínimos quadrados será também o estimador de máxima verossimilhança). Sob essas circunstâncias, a teoria assintótica afirma que o estimador de máxima verossimilhança se torna mais e mais não viesado, mais e mais normalmente distribuído e, aproximadamente, de menor variância possível. Entretanto, não é possível apresentar alguma diretriz de quão grande uma amostra deve ser para que tais propriedades sejam alcançadas.

2.7 Regiões de Inferência aproximada para os parâmetros

Bates & Watts (1988) apresentam as regiões de confiança conjuntas aproximadas para um modelo não linear, que estão definidas por analogia aos modelos lineares:

$$(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}})^T \hat{\mathbf{V}}^T \hat{\mathbf{V}} (\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}) \leq Ps^2 F_{(P, N-P; \alpha)} \quad (21)$$

ou

$$(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}})^T \hat{\mathbf{R}}_1^T \hat{\mathbf{R}}_1 (\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}) \leq Ps^2 F_{(P, N-P; \alpha)} \quad (22)$$

em que a matriz $\hat{\mathbf{V}}$ é a matriz de derivadas avaliada em $\hat{\boldsymbol{\theta}}$ e a matriz $\hat{\mathbf{R}}_1$ vem de uma decomposição $\mathbf{QR}$, como a que está no anexo 4, de $\hat{\mathbf{V}} = \hat{\mathbf{Q}}_1 \hat{\mathbf{R}}_1$. A fronteira desta região de confiança conjunta é:

$$\left\{ \boldsymbol{\theta} = \hat{\boldsymbol{\theta}} + \sqrt{Ps^2 F_{(P, N-P; \alpha)}} \hat{\mathbf{R}}_1^{-1} \mathbf{d} / \|\mathbf{d}\| = 1 \right\} \quad (23)$$

Os intervalos de confiança, o erro padrão aproximado para um determinado parâmetro θ_p , as correlações aproximadas e os erros padrões para os parâmetros são calculados considerando a aproximação linear e substituindo a matriz $\mathbf{X}$ nas equações resultantes da análise de um modelo linear pela matriz de derivadas $\hat{\mathbf{V}}$ avaliada em $\hat{\boldsymbol{\theta}}$.

As bandas de inferência de aproximação lineares para a resposta esperada também são obtidas por analogia aos modelos lineares, substituindo-se o valor estimado $\mathbf{x}_0^T \hat{\boldsymbol{\beta}}$ por $f(\mathbf{x}_0, \hat{\boldsymbol{\theta}})$, a matriz $\mathbf{X}$ por $\hat{\mathbf{V}}$ e o vetor derivado $\mathbf{x}_0$ por:

$$\mathbf{v}_0 = \frac{\partial f(\mathbf{x}_0, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}^T} \bigg|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}}$$

O intervalo aproximado $1-\alpha$ é, então:

$$f(\mathbf{x}_0, \hat{\boldsymbol{\theta}}) \pm s \left\| \mathbf{v}_0^T \hat{\mathbf{R}}_1^{-1} \right\| t_{(N-P; \alpha/2)} \quad (24)$$

e a banda de inferência aproximada $1-\alpha$ é:

$$f(\mathbf{x}, \hat{\boldsymbol{\theta}}) \pm s \left\| \mathbf{v}^T \hat{\mathbf{R}}_1^{-1} \right\| \sqrt{PF_{(P, N-P; \alpha)}} \quad (25)$$

2.8 Medidas de não linearidade

À medida que o tamanho da amostra cresce, o comportamento dos estimadores de mínimos quadrados se torna mais próximo do comportamento dos estimadores lineares, mas, até recentemente, não havia prontamente métodos aplicáveis para quantificar o comportamento não linear. As expressões que indicam a adequação de uma aproximação linear e seus efeitos nas inferências são chamadas medidas de não linearidade (Bates & Watts, 1980).

Beale (1960) fez a primeira tentativa teórica de medir a não linearidade, mas, suas medidas tendem a predizer que um modelo se comportará linearmente, até mesmo quando uma não linearidade considerável estiver presente (Guttman & Meeter, 1965).

Bates & Watts (1980) desenvolveram novas medidas de não linearidade baseadas no conceito geométrico de curvatura. Eles mostraram que a não linearidade de um modelo pode ser separada em dois componentes: (i) uma não linearidade “intrínseca” associada com a curvatura no local da solução e (ii) um

“efeito-parâmetro” de não linearidade associado com o fato de que as projeções das linhas de parâmetros no plano tangente ao local da solução não são, em geral, nem retas, nem paralelas, nem eqüidistantes.

2.8.1 O local da solução

A expressão da soma de quadrados em um modelo linear ou não linear:

$$S(\boldsymbol{\theta}) = \|\mathbf{y} - \boldsymbol{\eta}(\boldsymbol{\theta})\|^2$$

tem uma interpretação geométrica de ser o quadrado da distância entre o vetor $\mathbf{y}$ e $\boldsymbol{\eta}(\boldsymbol{\theta})$ em um espaço amostral N -dimensional. À medida que $\boldsymbol{\theta}$ é modificado, o vetor $\boldsymbol{\eta}(\boldsymbol{\theta})$ descreve uma superfície P -dimensional, chamada de “o local da solução” ou “superfície esperada” nesse espaço amostral (Box & Lucas, 1959). Em um modelo linear,

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e}$$

essa superfície é um hiperplano no espaço da resposta gerado pelas colunas de $\mathbf{X}$ em casos bem específicos, podendo ser uma reta ou um plano usual (Draper & Smith, 1998). Para exemplificar o local da solução para um modelo não linear, considere o exemplo apresentado por Ratkowsky (1983), analisando a função modelo não linear com um único parâmetro θ :

$$y_t = f(x_t, \theta) + e_t \text{ com } f(x_t, \theta) = x_t^\theta$$

Considere, ainda, que tem-se uma amostra de tamanho dois dada pelos valores da Tabela 2.1:

TABELA 2.1: Ilustração do conjunto de dados

Nº de observações	x_t	y_t
1	2	2,5
2	3	10,0

Fonte: Adaptada de Ratkowsky (1983)

Neste caso, $\boldsymbol{\eta}(\theta)$ é uma função em um domínio $\Theta \subseteq \Re$, com um contradomínio $\Re^2$, definida pela lei:

$$\boldsymbol{\eta}(\theta) = \begin{bmatrix} 2^\theta \\ 3^\theta \end{bmatrix}$$

Pode-se notar que à medida que θ varia, uma curva é gerada em um espaço de resposta bidimensional. Esta curva é denominada curva esperada ou superfície esperada e a estimativa de mínimos quadrados ou de máxima verossimilhança é um valor nesta superfície e, por isso, a mesma é denominada de o “local da solução”. No caso de um modelo linear, esta superfície seria uma reta para o caso de um parâmetro ou um plano para o caso de mais parâmetros. A Figura 2.1 representa o “local da solução”, para θ variando em $\Re$.

Observe que a superfície esperada é curvilínea e que variações iguais em θ não produzem variações iguais em $\boldsymbol{\eta}(\theta)$. O estimador $\hat{\theta}$ é aquele ponto do “local da solução” mais próximo de $\mathbf{y}$. Isso pode ser obtido geometricamente descendo uma perpendicular a “ponta do vetor” $\mathbf{y}$ para a linha tangente ao

“local da solução” para os vários pontos ao longo do mesmo e verificar o tamanho mínimo da distância perpendicular.

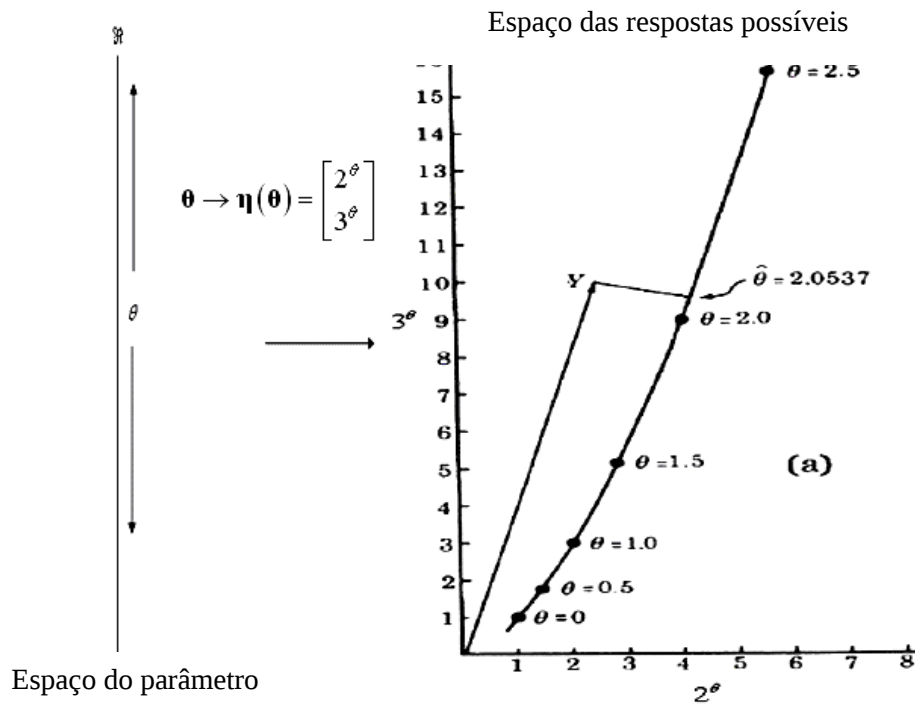


FIGURA 2.1: Espaço do parâmetro, espaço resposta e local da solução para $\eta(\theta)$ (Adaptado de Ratkowsky, 1983).

A forma do espaço solução na vizinhança de $\hat{\theta}$ e o espaço de valores da constante $\Delta\theta$ nessa vizinhança foram usadas por Bates & Watts (1980) para a medida do grau para qual um modelo não linear difere de um modelo linear. Essas possibilidades seguem do fato de que, em um modelo linear, o espaço

solução é linear, sendo uma linha reta para um modelo de parâmetro único e que os pontos do espaço solução linear correspondentes nos incrementos de θ são igualmente espaçados.

2.8.2 Linhas de parâmetro

Considerando um modelo com mais de um parâmetro, define-se uma linha de parâmetro no plano do parâmetro como sendo a linha associada ao parâmetro que está variando enquanto os demais se mantêm fixos (Bates & Watts, 1988). Por exemplo, em um modelo com dois parâmetros β_1 e β_2 , conforme Figura 2.2, as linhas correspondentes ao $\beta_1 = -3, \dots, 5$ (linhas pontilhadas) são chamadas linhas de β_2 :

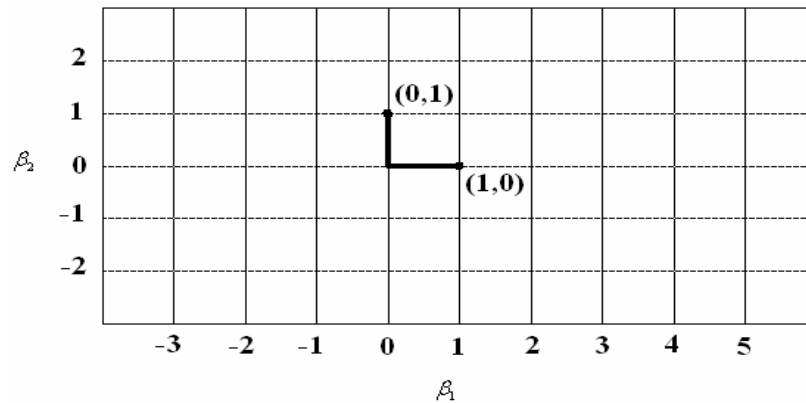


FIGURA 2.2: Linhas de parâmetro em um espaço do parâmetro bidimensional

Em um modelo linear, se as linhas de parâmetros são retas paralelas e também são equidistantes, como na Figura 2.2, as suas imagens no plano da resposta esperada são também retas, paralelas e equidistantes, como ilustrado na

Figura 2.3 (considerando um modelo com 2 parâmetros e um espaço de resposta tridimensional):

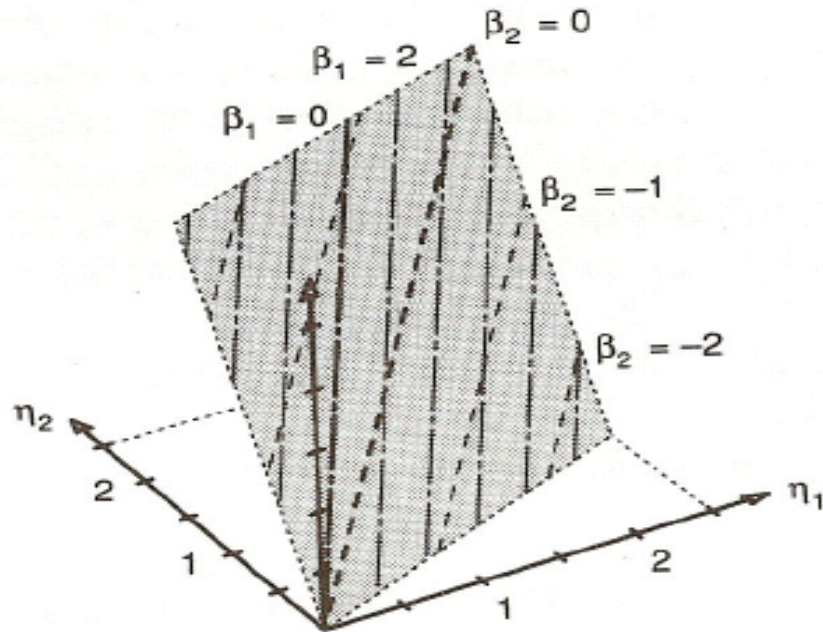


FIGURA 2.3: Imagens das linhas de parâmetro em um espaço tridimensional para um modelo linear (Adaptado de Bates & Watts, 1988).

Em um modelo não linear, as imagens no plano da resposta esperada não são, em geral, nem retas, nem paralelas e nem equidistantes, como ilustrado na Figura 2.4 (considerando um modelo com 2 parâmetros e um espaço de resposta tridimensional):

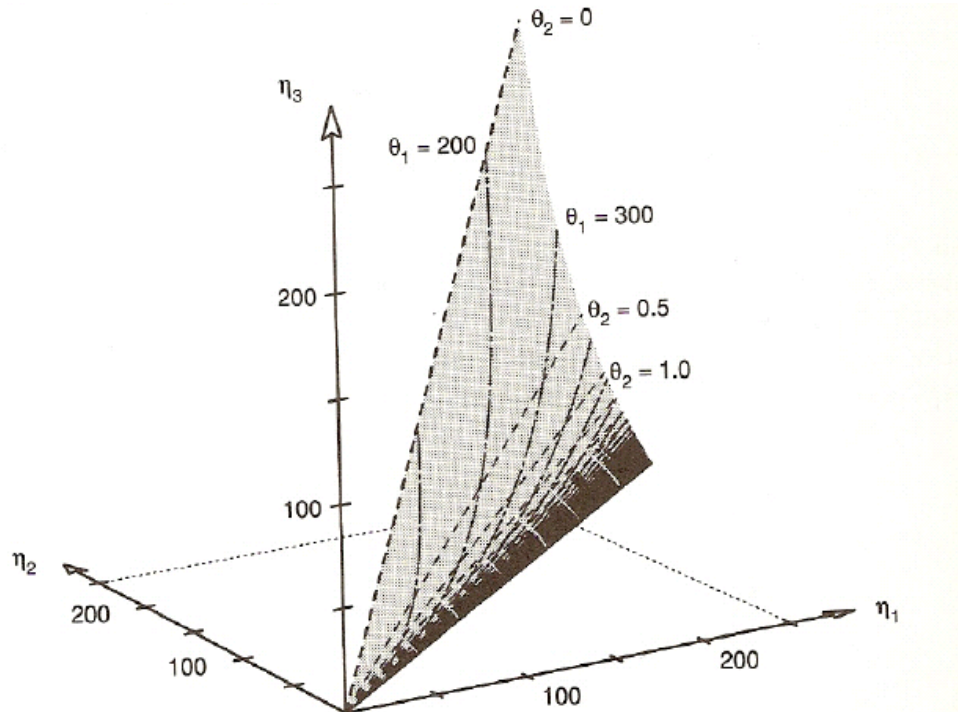


FIGURA 2.4: Imagens das linhas de parâmetro em um espaço tridimensional para um modelo linear (Adaptado de Bates & Watts, 1988).

2.8.3 Medidas de curvatura de Bates e Watts

Antes de dar sequência, convém explicar algumas notações utilizadas na literatura vigente. Para um dado modelo não linear, a sua matriz derivada $N \times P$ será escrita por $\dot{\mathbf{V}}$, com elementos:

$$\{\dot{\mathbf{V}}\}_{tp} = \frac{\partial f(x_t, \boldsymbol{\theta})}{\partial \theta_p}$$

e a segunda derivada um *array* $N \times P \times P$, denotado por $\ddot{\mathbf{V}}$ com elementos:

$$\{\ddot{\mathbf{V}}\}_{tpq} = \frac{\partial^2 f(x_t, \boldsymbol{\theta})}{\partial \theta_p \partial \theta_q}$$

com t seguindo de 1 a N , enquanto p e q seguem de 1 a P . Na notação matricial:

$$\dot{\mathbf{V}} = \frac{\partial \boldsymbol{\eta}}{\partial \boldsymbol{\theta}^T} \text{ e } \ddot{\mathbf{V}} = \frac{\partial^2 \boldsymbol{\eta}}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T}$$

em que cada linha de $\dot{\mathbf{V}}$ é o gradiente de uma coordenada de $\boldsymbol{\eta}(\boldsymbol{\theta})$ com relação a $\boldsymbol{\theta}$ e cada face $\ddot{\mathbf{V}}_t$ de $\ddot{\mathbf{V}}$ é uma matriz de derivada segunda $P \times P$ ou hessiana de um elemento de $\boldsymbol{\eta}(\boldsymbol{\theta})$ com relação a $\boldsymbol{\theta}$. Outra alternativa é considerar a matriz $\dot{\mathbf{V}}$ consistindo de vetores $\dot{\mathbf{V}}_p$ formado pelos elementos $\{\dot{\mathbf{V}}\}_{tp}$ com p fixo e t variando de 1 até N e a matriz $\ddot{\mathbf{V}}$ consistindo de vetores $\ddot{\mathbf{V}}_{pq}$ formados pelos elementos $\{\ddot{\mathbf{V}}\}_{tpq}$ com p e q fixos e t variando de 1 até N .

Os vetores $\dot{\mathbf{V}}_p$ são, evidentemente, os vetores tangentes e eles também são chamados de vetores velocidade, já que fornecem a taxa de variação instantânea de $\boldsymbol{\eta}$ com relação a cada parâmetro. Consequentemente, os vetores $\ddot{\mathbf{V}}_{pq}$ são chamados vetores aceleração, já que eles fornecem a taxas de variação instantânea dos vetores velocidades com relação aos parâmetros.

Para exemplificar o uso dessas notações, considere a função esperança do modelo apresentado por Marchi (2005), com alterações na notação:

$$f(x_t, \boldsymbol{\theta}) = \theta_1 (1 - e^{-\theta_2 x_t})$$

Para este modelo, os elementos de $\dot{\mathbf{V}}$ são:

$$\left\{ \dot{\mathbf{V}} \right\}_{n1} = 1 - e^{-\theta_2 x_t}$$

$$\left\{ \dot{\mathbf{V}} \right\}_{n2} = \theta_1 x_t e^{-\theta_2 x_t}$$

assim,

$$\dot{\mathbf{V}}_1 = \begin{bmatrix} 1 - e^{-\theta_2 x_1} \\ 1 - e^{-\theta_2 x_2} \\ \vdots \\ 1 - e^{-\theta_2 x_N} \end{bmatrix}, \quad \dot{\mathbf{V}}_2 = \begin{bmatrix} \theta_1 x_1 e^{-\theta_2 x_1} \\ \theta_1 x_2 e^{-\theta_2 x_2} \\ \vdots \\ \theta_1 x_N e^{-\theta_2 x_N} \end{bmatrix}, \quad \dot{\mathbf{V}} = \begin{bmatrix} 1 - e^{-\theta_2 x_1} & \theta_1 x_1 e^{-\theta_2 x_1} \\ 1 - e^{-\theta_2 x_2} & \theta_1 x_2 e^{-\theta_2 x_2} \\ \vdots & \vdots \\ 1 - e^{-\theta_2 x_N} & \theta_1 x_N e^{-\theta_2 x_N} \end{bmatrix}$$

Os elementos de $\ddot{\mathbf{V}}$ são:

$$\left\{ \ddot{\mathbf{V}} \right\}_{n11} = 0$$

$$\left\{ \ddot{\mathbf{V}} \right\}_{n12} = x_t e^{-\theta_2 x_t}$$

$$\left\{ \ddot{\mathbf{V}} \right\}_{n21} = \left\{ \ddot{\mathbf{V}} \right\}_{n12}$$

$$\left\{ \ddot{\mathbf{V}} \right\}_{n22} = -x_t^2 \theta_1 e^{-\theta_2 x_t}$$

O array $\ddot{\mathbf{V}}$ é, então, formado pelas N matrizes (ou faces) $P \times P$:

$$\begin{bmatrix} 0 & x_1 e^{-\theta_2 x_1} \\ x_1 e^{-\theta_2 x_1} & -x_1^2 \theta_1 e^{-\theta_2 x_1} \end{bmatrix}, \begin{bmatrix} 0 & x_2 e^{-\theta_2 x_2} \\ x_2 e^{-\theta_2 x_2} & -x_2^2 \theta_1 e^{-\theta_2 x_2} \end{bmatrix}, \dots, \begin{bmatrix} 0 & x_N e^{-\theta_2 x_N} \\ x_N e^{-\theta_2 x_N} & -x_N^2 \theta_1 e^{-\theta_2 x_N} \end{bmatrix}$$

A função modelo fornece um mapeamento do espaço paramétrico P-dimensional sobre uma superfície P-dimensional no espaço amostral N-dimensional. Cada ponto $\boldsymbol{\theta}_0$ mapeia para um ponto $\boldsymbol{\eta}(\boldsymbol{\theta}_0)$ no local da

solução; similarmente, linhas no espaço paramétrico pelo ponto $\boldsymbol{\theta}_0$ mapeiam para curvas no local da solução por meio de $\boldsymbol{\eta}(\boldsymbol{\theta}_0)$.

Uma linha reta arbitrária no espaço paramétrico por meio de $\boldsymbol{\theta}_0$ pode ser expressa utilizando-se o parâmetro geométrico b por:

$$\boldsymbol{\theta}(b) = \boldsymbol{\theta}_0 + b\mathbf{h}$$

em que $\mathbf{h} = (h_1, h_2, \dots, h_p)^T$ é qualquer vetor não nulo. Essa “linha” gera uma curva, $\boldsymbol{\eta}_h(b)$ no local da solução, em que:

$$\boldsymbol{\eta}_h(b) = \boldsymbol{\eta}(\boldsymbol{\theta}_0 + b\mathbf{h})$$

Bates & Watts (1980) mostram que o vetor tangente à curva $\boldsymbol{\eta}_h(b)$ em $b=0$ é $\dot{\boldsymbol{\eta}}_h = \dot{\mathbf{V}}\mathbf{h}$, ou seja, $\dot{\boldsymbol{\eta}}_h$ é uma combinação linear dos vetores velocidades e o conjunto de todas essas combinações lineares é o plano tangente em $\boldsymbol{\eta}(\boldsymbol{\theta}_0)$. A aceleração da curva $\boldsymbol{\eta}_h(b)$ em $b=0$ é $\ddot{\boldsymbol{\eta}}_h = \mathbf{h}^T \ddot{\mathbf{V}}\mathbf{h}$, ou seja, o vetor $\ddot{\boldsymbol{\eta}}_h$ é precisamente uma combinação linear dos vetores aceleração. Isto é, pré e pós-multiplicando o *array* $\ddot{\mathbf{V}}$ por um vetor $(1 \times P)$ e um vetor $(P \times 1)$ produz um vetor $(N \times 1)$, cada elemento do qual é um termo da forma $\mathbf{h}^T \mathbf{V}_t \mathbf{h}$ é a t -ésima face do *array* $\ddot{\mathbf{V}}$.

Os vetores $\dot{\boldsymbol{\eta}}_h$ e $\ddot{\boldsymbol{\eta}}_h$ têm uma interpretação física, imaginando um ponto se movendo através do espaço amostral, tal que, no “tempo” b , ele está em

$\mathbf{\eta}_h(b)$. A tangente $\dot{\mathbf{\eta}}_h$ é a velocidade instantânea e $\ddot{\mathbf{\eta}}_h$ é a aceleração instantânea quando $b = 0$.

O vetor aceleração $\ddot{\mathbf{\eta}}_h$ pode ser escrito como três componentes, $\ddot{\mathbf{\eta}}_h^N$ normal ao plano tangente, $\ddot{\mathbf{\eta}}_h^P$ paralelo ao $\dot{\mathbf{\eta}}_h$, e $\ddot{\mathbf{\eta}}_h^G$ paralelo ao plano tangente e normal ao $\dot{\mathbf{\eta}}_h$. Então:

$$\ddot{\mathbf{\eta}}_h = \ddot{\mathbf{\eta}}_h^N + \ddot{\mathbf{\eta}}_h^P + \ddot{\mathbf{\eta}}_h^G$$

Fisicamente, a aceleração normal $\ddot{\mathbf{\eta}}_h^N$ determina a mudança de direção do vetor $\dot{\mathbf{\eta}}_h$ normal ao plano tangente, enquanto a aceleração geodésica $\ddot{\mathbf{\eta}}_h^G$ determina a mudança de direção do vetor $\ddot{\mathbf{\eta}}_h$ paralelo ao plano tangente. O componente tangente paralelo $\ddot{\mathbf{\eta}}_h^P$ determina a mudança de velocidade do movimento pontual e, consequentemente, determina se o ponto move-se uniformemente sobre o local da solução.

2.8.4 Curvaturas

Os componentes de aceleração podem ser convertidos em curvaturas; por exemplo:

$$k_h^N = \frac{\|\ddot{\mathbf{\eta}}_h^N\|}{\|\dot{\mathbf{\eta}}_h\|^2}$$

fornece a curvatura normal na direção $\mathbf{h}$. Sua interpretação geométrica é que a mesma é a inversa do raio do círculo que melhor aproxima o local da solução na

direção $\dot{\mathbf{\eta}}_{\mathbf{h}}$ em $\boldsymbol{\theta}_0$. Visto que a curvatura normal é uma propriedade do local da solução e depende somente do modelo matemático – delineamento experimental – combinação linear dos vetores de parâmetros (Lipschuts, 1969), ela é uma propriedade intrínseca da superfície. Assim, $k_{\mathbf{h}}^N$ será chamada a curvatura intrínseca na direção associada com $\mathbf{h}$.

Os componentes de aceleração tangencial, que são causados pela parametrização do modelo, podem ser combinados em uma aceleração tangencial total:

$$\ddot{\mathbf{\eta}}_{\mathbf{h}}^T = \ddot{\mathbf{\eta}}_{\mathbf{h}}^P + \ddot{\mathbf{\eta}}_{\mathbf{h}}^G$$

Isso pode ser convertido como uma curvatura do efeito parâmetro na direção $\mathbf{h}$ como:

$$k_{\mathbf{h}}^T = \frac{\|\ddot{\mathbf{\eta}}_{\mathbf{h}}^T\|}{\|\dot{\mathbf{\eta}}_{\mathbf{h}}\|^2}$$

em contraste com a curvatura intrínseca, a curvatura efeito parâmetro depende da particular parametrização escolhida e, então, uma reparametrização não linear do modelo para $\boldsymbol{\theta}$, dita $\boldsymbol{\beta} = G(\boldsymbol{\theta})$, pode mudar consideravelmente a curvatura do efeito parâmetro.

As curvaturas intrínsecas e de efeito parâmetro afetam as regiões de confiança pela aproximação linear de modo diferente. A aproximação linear combina uma pressuposição planar e uma pressuposição de coordenadas uniformes. A pressuposição planar envolve substituição do local da solução curvado pelo plano tangente e essa aproximação não será aceitável se a

curvatura intrínseca máxima é grande em $\boldsymbol{\eta}(\hat{\boldsymbol{\theta}})$. Similarmente, a pressuposição de coordenadas uniformes envolve substituição da linha de parâmetro curvada na aproximação do plano tangente por uma grade de retas, paralelas, linhas eqüidistantes e essa aproximação não será aceitável se a máxima curvatura de efeito parâmetro é grande em $\boldsymbol{\eta}(\hat{\boldsymbol{\theta}})$. Assim, a pressuposição planar tem que ser satisfeita em um estágio anterior para, depois, verificar se a pressuposição de coordenadas é aceitável.

2.8.5 Calculando curvaturas

Bates & Watts (1988) descrevem algoritmos para se encontrar as medidas de curvaturas que possuem interesses práticos na avaliação da linearidade de um modelo não linear.

As medidas k_h^N e k_h^T , respectivamente, a não linearidade intrínseca no local da solução na direção de $\dot{\boldsymbol{\eta}}_h$ e a não linearidade aparente causada pela parametrização do modelo, mas não feita por ele, indica se a não linearidade afeta seriamente as inferências baseadas em aproximações lineares. Para acessar o efeito da não linearidade na inferência, é conveniente converter k_h^N e k_h^T para curvaturas relativas padronizadas γ_h^N e γ_h^T , que são invariantes sob mudança de escalas na resposta. Recomenda-se utilizar curvaturas relativas para comparar não somente as diferentes parametrizações para um dado problema, mas também diferentes conjuntos de dados para o mesmo ou modelos diferentes.

Bates & Watts (1980) recomendam o padrão escalar dado pelo raio padrão, $\rho = s\sqrt{P}$. A justificativa é que, no caso linear, uma região de confiança conjunta $(1-\alpha)$ é dada pelo conjunto de $\boldsymbol{\theta}$, tal que

$\|\boldsymbol{\eta}(\boldsymbol{\theta}) - \boldsymbol{\eta}(\hat{\boldsymbol{\theta}})\| \leq \rho \sqrt{F_{(P, N-P; \alpha)}}$. Assim, a curvatura intrínseca relativa na direção

$\dot{\boldsymbol{\eta}}_{\mathbf{h}}$ é definida como:

$$\gamma_{\mathbf{h}}^N = k_{\mathbf{h}}^N \cdot \rho$$

e a curvatura efeito-parâmetro é definida como:

$$\gamma_{\mathbf{h}}^T = k_{\mathbf{h}}^T \cdot \rho$$

Visto que $k_{\mathbf{h}}^N$ é a curvatura intrínseca, o número $1/k_{\mathbf{h}}^N$ fornecerá o raio da curvatura no local da solução na direção tangente $\dot{\boldsymbol{\eta}}_{\mathbf{h}}$ em $\boldsymbol{\eta}(\hat{\boldsymbol{\theta}})$. Assim, o raio $1/k_{\mathbf{h}}^N$ e $1/k_{\mathbf{h}}^T$ são comparados de acordo com o raio padrão ρ , então, valores pequenos de $\gamma_{\mathbf{h}}$ indicam que o raio padrão ρ é menor, relativamente, do que o raio $1/k_{\mathbf{h}}$ e valores grandes de $\gamma_{\mathbf{h}}$ indicam que o raio padrão é maior relativamente do que o raio $1/k_{\mathbf{h}}$. Para determinar o impacto da não linearidade sob uma região de confiança particular, a curvatura escalada poderia ser comparada com $1/\sqrt{F_{(P, N-P; \alpha)}}$, que pode ser observada como o raio da curvatura para região de confiança $100(1-\alpha)\%$. Se, para todas as direções $\mathbf{h}$, $\gamma_{\mathbf{h}}^N$ é pequeno comparado com $1/\sqrt{F_{(P, N-P; \alpha)}}$, então, o local da solução é relativamente plano sobre o disco da região de confiança e, conseqüentemente, pode-se usar a pressuposição planar com razoável segurança. Similarmente, se $\gamma_{\mathbf{h}}^T$ é pequena comparada com $1/\sqrt{F_{(P, N-P; \alpha)}}$, então, a pressuposição de coordenadas uniformes é boa sobre a região de interesse.

Como existem, em geral, infinitas direções na qual se pode tomar as medidas de curvatura intrínseca ou de efeito parâmetro, se torna complicado e pouco útil avaliar as medidas de curvatura em uma determinada direção particular.

Bates & Watts (1980) sugerem uma alternativa conservadora de se tomar as medidas de curvaturas máximas no local da solução em $\boldsymbol{\eta}(\hat{\boldsymbol{\theta}})$. Assim, se a “maior” das curvaturas for considerada pequena, por transitividade matemática, todas as outras assim também o serão. Dessa maneira, eles consideram todas as curvaturas em todas as direções possíveis no local da solução em $\boldsymbol{\eta}(\hat{\boldsymbol{\theta}})$ para a realização das análises. Um problema levantado por eles é a não existência de fórmulas analíticas para o cômputo de tais curvaturas, sendo sugerido, então, algoritmos alternativos que viabilizem esses resultados.

2.8.6 Curvaturas RMS

Bates & Watts (1980, 1988) sugerem uma medida simplificadora para o estudo da não linearidade por meio de curvatura. Os mesmos autores apresentam a medida RMS (curvatura quadrática média), dada pela relação:

$$c^2 = \frac{1}{A} \int_{\|u\|=1} \sum_t (\mathbf{u}^T \mathbf{C}_t \mathbf{u})^2$$

em que A é área de uma esfera unitária P -dimensional e $\mathbf{C}_t$ é a t -ésima face de um *array* de curvatura que armazena as curvaturas relativas já escaladas pelo fator escala ρ . A notação utilizada pelos mesmos para a curvatura RMS de efeitos paramétricos é c^0 e a curvatura intrínseca RMS por c^I . O cálculo dessas

medidas não é necessariamente muito simples, mas, felizmente, os autores elaboraram um programa computacional, que resolve o problema do cálculo das mesmas.

Pressupondo que os cálculos já foram realizados, surge, então, o problema do que se considera uma curvatura grande. Bates & Watts (1988) sugerem uma escala referencial conveniente que pode ser estabelecida para essas comparações, que é estabelecida pela comparação da curvatura RMS com aquela do disco de confiança em um nível de confiança específico. Assim, uma curvatura RMS será considerada pequena se ela é muito menor do que a curvatura do disco de confiança, a $100(1-\alpha)\%$, isto é, $c \ll 1/\sqrt{F}$, ou equivalentemente se $c\sqrt{F} \ll 1$, em que $F = F_{(p, N-p; \alpha)}$, como mencionado anteriormente.

Segundo Bates & Watts (1980), deve-se considerar uma superfície esperada com raio $1/c$ e determinar o desvio da superfície a partir do plano tangencial em uma distância $\sqrt{F}$ do ponto tangencial. Este desvio expressado como uma percentagem do raio do disco de confiança é $100\left(1 - \sqrt{1 - c\sqrt{F}}\right) / c\sqrt{F}$, portanto, o valor de $c\sqrt{F} = 0,1$ causa um desvio da superfície por 5% de raio de confiança; um valor de $c\sqrt{F} = 0,2$ causa um desvio de 10%; desvio de 0,3 de 15% e desvio de 0,4 de 21%. Daí, se, por exemplo, aceita-se um desvio de não mais do que 15%, pode-se declarar que uma análise é inaceitável se pelo menos um valor de $c\sqrt{F}$ for maior do que 0,3, quando substitui-se c por c^0 ou c^l .

2.9 Estudos de simulação

Os estimadores em regressão não-linear não possuem certas propriedades que os estimadores em regressão linear. No entanto, o estudo de não linearidade tem como foco verificar se o modelo não linear possui propriedades próximas dos modelos lineares, de modo que, se o estimador tem um viés pequeno, uma distribuição próxima da normal e uma variância próxima da variância mínima, parece razoável falar que o estimador tem um comportamento próximo do linear.

Quando as hipóteses de normalidade e da aproximação linear assintótica no modelo de regressão não-linear se tornam questionáveis devido, possivelmente, ao tamanho reduzido da amostra ou à curvatura excessiva da superfície resposta, a técnica “bootstrap” de estimação se torna atrativa, não somente como alternativa para o processo inferencial, como também como ferramenta de diagnóstico (Souza, 1998). Para o autor, o princípio “bootstrap” estabelece que M amostras de tamanho n são simuladas para um modelo não linear como definido neste estudo, então, $\hat{\theta}_m$ e s_m^2 em repetidas amostras do modelo amostral, com $m=1,...,M$, se comportam em relação a $\hat{\theta}$ e s^2 do mesmo modo que $\hat{\theta}$ e s^2 , em repetidas amostras do modelo original, se comportam em relação aos elementos desconhecidos θ e σ^2 . Desse modo, torna-se possível estudar as propriedades distribucionais dos estimadores de mínimos quadrados a partir de simulações do modelo de regressão não linear em estudo. Esse método é conhecido como “bootstrap” paramétrico.

Segundo Ratkowsky (1983), o estudo de simulação constitui a melhor maneira para analisar as propriedades amostrais dos estimadores de mínimos quadrados. O autor sugere o uso do “bootstrap” paramétrico para a realização dos estudos de simulação e sugere que, para se estudar o comportamento

amostral em modelos de regressão não linear, deve-se usar um mínimo de 1.000 conjuntos de dados simulados pseudo-aleatoriamente, especialmente para estatísticas baseadas em momentos amostrais grandes, tais como coeficientes de assimetria e curtose.

De acordo com Souza (1998), a forma mais simples para o modelo de regressão não linear envolve a hipótese de normalidade, em que é assumido que ambas as formas do componente determinístico do modelo e a natureza do componente estocástico ϵ são conhecidos. Na simulação, os dados são gerados utilizando-se um conjunto de valores pré-determinados dos parâmetros, permitindo somente os valores de ϵ mudarem pseudo-aleatoriamente de conjunto a conjunto. Dessa forma, muitos conjuntos de valores simulados são produzidos. Cada qual fornece um conjunto de estimativas de mínimos quadrados para os parâmetros que podem, então, ser examinados para o viés deles, variância e propriedades distribucionais.

Mood et al. (1974) mostram que duas variáveis aleatórias possuem a mesma densidade se elas possuírem a mesma função geradora de momentos, ou seja, se elas possuem todos os seus momentos iguais. Com base nessa idéia, pode-se verificar o comportamento não normal dos estimadores de mínimos quadrados, mesmo que como uma aproximação, analisando-se, pelo menos, os quatro primeiro momentos amostrais dos mesmos, que estão diretamente ligados com as medidas da média amostral m_1 , variância amostral viesada m_2 , coeficiente de assimetria $g_1 = m_3/m_2^{3/2}$ e coeficiente de curtose m_4/m_2^2 , em que m_3 e m_4 são o terceiro e o quarto momentos amostrais em relação a média, respectivamente.

O coeficiente de assimetria e curtose pode ser comparado com o valor esperado deles sob amostragem para distribuição normal (0 e 3, respectivamente) para testar se o conjunto de estimativas do parâmetro é

comportado normalmente. Para um grande estudo de simulação, g_1 é aproximadamente normalmente distribuído com média zero e desvio padrão $(6/N)^{1/2}$ (Snedecor & Cochran, 1980). Para o coeficiente de curtose, é melhor subtrair o valor 3 esperado para uma distribuição normal para se obter o coeficiente então chamado de “excesso de curtose” $g_2 = (m_4/m_2^2) - 3$. Para N grande, g_2 é aproximadamente normalmente distribuído com média zero e desvio padrão $(24/N)^{1/2}$ (Snedecor & Cochran, 1980).

De fundamental importância em regressão não linear é o estabelecimento de intervalos de confiança para os parâmetros do modelo. Segundo Souza (1998), a técnica “bootstrap” pode fornecer de modo automático a escala de mensuração adequada relativamente à qual é possível determinar intervalos confiáveis. Dentre estes, o autor destaca os intervalos obtidos via método percentil, que têm por base unicamente os quantis e outras medidas da distribuição “bootstrap”.

Após a obtenção desses intervalos, deve-se comparar os resultados obtidos com os resultados obtidos via teoria assintótica e verificar se os mesmos não diferem muito um do outro. Para Machado (2006), essa comparação pode ser realizada considerando a amplitude dos mesmos. Aquele que possuir uma amplitude menor apresentará maior precisão na estimação dos parâmetros do modelo. No entanto, esta análise também tem o papel de diagnosticar se os resultados obtidos pela aproximação linear não apresentam problemas teóricos com relação a inferência, devido a um tamanho reduzido da amostra.

3 MATERIAL E MÉTODOS

3.1 Material

Os dados¹ deste trabalho foram obtidos do trabalho de Marchi (2005) e provêm de experimentos realizados nos laboratórios de ciências ambientais da Universidade da Califórnia, em Riverside, entre os anos de 2003 e 2004. O solo usado neste experimento - classificado como Kandiudox – foi coletado de um experimento realizado em Jaguariúna, São Paulo, Brasil, 22°41'S, 47°W. Estas amostras foram analisadas para avaliar suas características físicas e químicas: $\text{pH}_{\text{H}_2\text{O}} = 5,5$; Ca^{2+} , Mg^{2+} , e Al^{3+} ($1 \text{ mol L}^{-1} \text{ KCl}$) = 27,5, 8,5 e $1 \text{ mmol}_c \text{ dm}^{-3}$, respectivamente; P (Mehlich I) = $3,5 \text{ mg dm}^{-3}$; K^+ = $1,51 \text{ mmol}_c \text{ dm}^{-3}$; a acidez titulável extraída com $0,5 \text{ mol L}^{-1}$ de acetato de cálcio com $\text{pH } 7,0$ (H^+Al^3) = $35 \text{ mmol}_c \text{ dm}^{-3}$, matéria orgânica (método Walkley – Black) = $25,5 \text{ g kg}^{-1}$; argila = 450 g kg^{-1} .

Um lodo de esgoto obtido da Estação de Tratamento de Lodo de Esgoto da cidade de Franca, SP, Brasil, foi aplicado e essa área delimitada em 5 anos consecutivos (24,11, 26,03, 30,13, 35,45 e $34,80 \text{ t ha}^{-1}$ massa seca, a cada ano, respectivamente), acumulando $150,53 \text{ t ha}^{-1}$. Amostras do solo tratado com lodo de esgoto foram coletadas e digeridas utilizando-se o método USEPA 3052 ($\text{HCl-HNO}_3\text{-HF}$, digestão de microondas) (United States Environmental Protection Agency - USEPA, 1996). O teor total do elemento-traço² Zn foi de $71,74 \text{ mg kg}^{-1}$.

¹ Os dados deste estudo foram gentilmente cedidos por Giuliano Marchi, pesquisador da Embrapa Cerrados, DF-Brasil.

² Em seu trabalho, o autor estudou cinco elementos-traço, Cd, Zn, Cr, Ni e Cu. Neste estudo, as análises foram realizadas somente para o Zn, para cada um dos sete extratores.

Sete extratores foram usados para extrair, sucessivamente, os elementos-traço das amostras de solo. Quatro extratores foram compostos por ácidos orgânicos de baixo peso molecular ($0,1 \text{ mol}_e \text{ L}^{-1}$): i) mistura de ácidos orgânicos modificada (MOAS); ii) mistura de ácidos orgânicos simplificada (SOAS); iii) ácido láctico e iv) ácido acético. O pH dessas soluções foi ajustado para $4,8 \pm 0,1$. Um grama e meio de solo foi colocado em tubos de 50 mL contendo 30 mL para cada extrator. Os tubos foram agitados a 25°C , por 24 horas a 2 rpm usando um misturador giratório. Um por cento (v/v) de clorofórmio foi adicionado para cada volume de solução para o controle da atividade microbial e para prevenir a decomposição da acidez orgânica durante a extração.

Os extratores para as análises rotineiras, Mehlich I, acetato de amônio (NH_4OAc) e DTPA, foram também utilizados para extração dos elementos-traço. O procedimento para cada extrator foi descrito como segue:

- i) Mehlich I, 30 mL de $0,05 \text{ mol L}^{-1} \text{ HCl}$, e $0,0125 \text{ mol L}^{-1} \text{ H}_2\text{SO}_4$ foram misturados com 7,8 g da amostra e agitados por 5 minutos;
- ii) acetato de amônio (NH_4OAc) – 30 mL de 1 mol L^{-1} de acetato de amônio, pH 7, foram misturados com 1,875 g da amostra e agitados por 14 horas;
- iii) DTPA – 20 mL de $0,005 \text{ mol L}^{-1}$ ácido dietilenotriaminopentacético, $0,1 \text{ mol L}^{-1} \text{ CaCl}_2$, pH 7,3; foram misturados com 10 gramas de solo e agitados por 2 horas.

Tanto, a mistura de ácidos orgânicos e os extratores de análises rotineiras, foram utilizados para, sucessivamente, extrair os elementos-traço do solo.

A tendência da remoção cumulativa do metal com respeito ao número de extrações foi ajustada através de um modelo de cinética de primeira ordem

(modelo exponencial de elevação máxima), em que a dessorção cumulativa de um íon das partículas do solo é dado por Bermond & Ghestem (2001):

$$C_t = C_l (1 - e^{-kt})$$

em que C_t representa os metais cumulativos removidos (mg kg^{-1}) para a extração t ; C_l é a quantidade de metal extraído até o equilíbrio (mg kg^{-1}); t é o número de extrações (ε); k é uma constante positiva representando a taxa de extração (ε^{-1}), ou seja, a constante k indica a velocidade com que a extração é realizada.

Para as análises estatísticas, foi considerado o erro aditivo (ξ_t):

$$C_t = C_l (1 - e^{-kt}) + \xi_t$$

com ξ_t sendo o erro aleatório identicamente, normalmente e independentemente distribuído, com média zero e variância constante σ^2 .

O delineamento utilizado foi o inteiramente ao acaso com três repetições, devido à homogeneização do solo realizada em laboratório. As extrações sucessivas e cumulativas representam os tratamentos.

3.2 Métodos

3.2.1 Estimativas iniciais

As estimativas iniciais para cada um dos sete conjuntos de dados, referentes aos sete extratores, foram encontradas com base em análises na parte determinística do modelo:

$$C_t = C_l (1 - e^{-kt})$$

Primeiramente, observa-se que:

$$\lim_{t \rightarrow \infty} (C_t) = \lim_{t \rightarrow \infty} (C_l (1 - e^{-kt})) = \lim_{t \rightarrow \infty} (C_l - C_l e^{-kt}) = C_l$$

Considerando que a expressão só terá sentido prático se $C_l > 0$ e $k > 0$ e que os dados experimentais são valores pré-estabelecidos com $t \geq 0$, optou-se por uma estimativa inicial para C_l como o valor máximo para C_t , fornecido pelos dados. Isso porque a estabilização ocorrerá para o modelo próximo de C_l e que o experimento foi feito com a intenção de se alcançar esta estabilização que, provavelmente, ocorrerá próximo do máximo de C_t fornecido pela amostra, como mostrado na Figura 3.1. Tal estimativa inicial será denotada por C_l^0 .

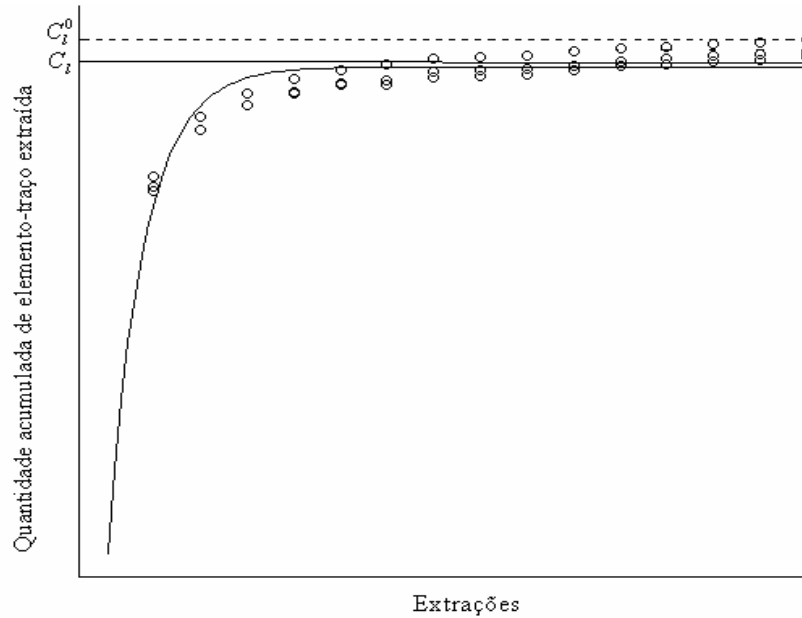


FIGURA 3.1: Representação gráfica para a escolha da estimativa inicial C_l^0

Para encontrar as estimativas iniciais para k , utilizou-se a estimativa inicial C_l^0 e reescreveu-se a parte determinística do modelo como:

$$C_t = C_l^0 (1 - e^{-kt})$$

Como C_t e t nulos não oferece muita informação prática para o pesquisador, considerou-se, então, $t > 0$ e $C_t > 0$. Dessa forma, pode-se reorganizar a expressão para:

$$\begin{aligned} C_t = C_l^0 (1 - e^{-kt}) &\Rightarrow C_t = C_l^0 - C_l^0 e^{-kt} \Rightarrow C_l^0 e^{-kt} = C_l^0 - C_t \Rightarrow \\ &\Rightarrow e^{-kt} = (C_l^0 - C_t) / C_l^0 \Rightarrow -kt = \ln[(C_l^0 - C_t) / C_l^0] \Rightarrow \\ &\Rightarrow k = -\frac{\ln[(C_l^0 - C_t) / C_l^0]}{t} \end{aligned}$$

Essa expressão apresenta problemas matemáticos quando $C_t \leq C_l^0$, mas como $C_l^0 = \max\{C_t\}$ para os dados experimentais, só haverá problemas quando $C_t = C_l^0$ nos dados. Então, sugere-se uma adequação na expressão, substituindo-se C_l^0 por $C_l^* = C_l^0 + \delta$, com δ positivo e pequeno com relação à escala utilizada para os dados. Observando, ainda, que para cada t tem-se um k diferente, sugere-se, então, que a estimativa inicial para k seja dada pela média de todos os k obtidos pelos dados experimentais, ou seja:

$$k^0 = \frac{\sum_{t=1}^n -\frac{\ln[(C_l^* - C_t) / C_l^*]}{t}}{n}$$

Para todos os dados do experimento, considerou-se, empiricamente, $\delta = 0,05$.

3.2.2 Estimativas

Para a estimação dos parâmetros do modelo não linear foi utilizada a metodologia estatística clássica, com o uso do método iterativo de Gauss-Newton (Bates & Watts, 1988). As estimativas para os parâmetros foram encontradas utilizando-se rotina ou procedimento específico desenvolvida para o software R versão 2.7.0 (R Development Core Team, 2008), com o emprego da função “*nls*” que apresenta tal algoritmo como *default* e que requer estimativas iniciais, que foram fornecidas de acordo com a metodologia exposta anteriormente (anexo 1). Essa função apresenta como resultados as estimativas dos parâmetros do modelo, o erro padrão das mesmas, o erro padrão residual, os valores de p (para o teste t) amostrais para cada parâmetro e os seus respectivos número de graus de liberdade considerados.

3.2.3 Bandas de confiança para a resposta esperada

Considerando a notação utilizada por Bates & Watts (1988), a parte determinística do modelo é reescrita da seguinte forma:

$$f(x, \boldsymbol{\theta}) = \theta_1 (1 - e^{-\theta_2 x}), \text{ com } \boldsymbol{\theta} = \begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix}$$

Para encontrar a banda de confiança para a resposta esperada foi utilizada a equação (25), adequada para o modelo proposto no estudo:

$$f(\mathbf{x}, \hat{\boldsymbol{\theta}}) \pm s \|\mathbf{v}^T \mathbf{R}_1^{-1}\| \sqrt{PF_{(P, N-P; \alpha)}}$$

ou

$$f(\mathbf{x}, \hat{\boldsymbol{\theta}}) \pm \|\mathbf{v}^T \mathbf{R}_1^{-1}\| \sqrt{s^2 PF_{(P, N-P; \alpha)}}$$

Com:

$$f(x, \hat{\boldsymbol{\theta}}) = \hat{\theta}_1 (1 - e^{-\hat{\theta}_2 x}), \text{ com } \hat{\boldsymbol{\theta}} = \begin{bmatrix} \hat{\theta}_1 \\ \hat{\theta}_2 \end{bmatrix}$$

vetor de derivadas em relação aos parâmetros:

$$\mathbf{v}^T = [v_1 \quad v_2], \text{ sendo } v_1 = \left[\frac{\partial f(x, \boldsymbol{\theta})}{\partial \theta_1} \right]_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} \text{ e } v_2 = \left[\frac{\partial f(x, \boldsymbol{\theta})}{\partial \theta_2} \right]_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}}$$

Assim,

$$v_1 = 1 - e^{-\hat{\theta}_2 x} \text{ e } v_2 = \hat{\theta}_1 x e^{-\hat{\theta}_2 x}$$

Logo,

$$\mathbf{v}^T = \begin{bmatrix} 1 - e^{-\hat{\theta}_2 x} & \hat{\theta}_1 x e^{-\hat{\theta}_2 x} \end{bmatrix}$$

Agora, tomando-se a matriz de derivada $\dot{\mathbf{V}}$ avaliada em $\hat{\boldsymbol{\theta}}$ e fazendo-se uma decomposição $\mathbf{QR}$, tem-se $\dot{\mathbf{V}} = \mathbf{QR} = \mathbf{Q}_1 \mathbf{R}_1$. Essa decomposição pode ser calculada diretamente no software R versão 2.7.0, usando a função “qr”, que faz parte do pacote “base” do mesmo. Realizada essa decomposição (conforme rotina no anexo 2) e considerando:

$$\mathbf{A} = \mathbf{R}_1^{-1} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

Tem-se:

$$\begin{aligned} \mathbf{v}^T \mathbf{R}_1^{-1} = \mathbf{v}^T \mathbf{A} &= \begin{bmatrix} 1 - e^{-\hat{\theta}_2 x} & \hat{\theta}_1 x e^{-\hat{\theta}_2 x} \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \\ &= \left[\left(1 - e^{-\hat{\theta}_2 x}\right)a + \left(\hat{\theta}_1 x e^{-\hat{\theta}_2 x}\right)c \quad \left(1 - e^{-\hat{\theta}_2 x}\right)b + \left(\hat{\theta}_1 x e^{-\hat{\theta}_2 x}\right)d \right] \end{aligned}$$

Dessa maneira:

$$\|\mathbf{v}^T \mathbf{R}_1^{-1}\| = \sqrt{\left[\left(1 - e^{-\hat{\theta}_2 x}\right)a + \left(\hat{\theta}_1 x e^{-\hat{\theta}_2 x}\right)c \right]^2 + \left[\left(1 - e^{-\hat{\theta}_2 x}\right)b + \left(\hat{\theta}_1 x e^{-\hat{\theta}_2 x}\right)d \right]^2}$$

Portanto, a banda de confiança para resposta esperada, no nível $1 - \alpha$ de confiança, é dada pela expressão abaixo:

$$\hat{\theta}_1(1 - e^{-\hat{\theta}_2 x}) \pm k \sqrt{\left[(1 - e^{-\hat{\theta}_2 x})a + (\hat{\theta}_1 x e^{-\hat{\theta}_2 x})c \right]^2 + \left[(1 - e^{-\hat{\theta}_2 x})b + (\hat{\theta}_1 x e^{-\hat{\theta}_2 x})d \right]^2}$$

com

$$k = \sqrt{s^2 P F_{(P, N-P; \alpha)}}$$

Os gráficos das bandas de confiança foram montados em rotinas elaboradas no software R versão 2.7.0, utilizando estas expressões como referência, conforme pode ser visto no Anexo 2.

3.2.4 Região de confiança conjunta para os parâmetros

Para encontrar a região de confiança conjunta para os parâmetros utilizando a aproximação linear, considerou-se a equação (23):

$$\left\{ \boldsymbol{\theta} = \hat{\boldsymbol{\theta}} + \sqrt{P s^2 F_{(P, N-P; \alpha)}} \hat{\mathbf{R}}_1^{-1} \mathbf{d} / \|\mathbf{d}\| = 1 \right\}$$

ou

$$\left\{ \boldsymbol{\theta} = \hat{\boldsymbol{\theta}} + k \mathbf{A} \mathbf{d} / \|\mathbf{d}\| = 1 \right\}, \text{ com } k = \sqrt{P s^2 F_{(P, N-P; \alpha)}} \text{ e } \mathbf{A} = \hat{\mathbf{R}}_1^{-1} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

Como o caso em questão envolve apenas dois parâmetros, é necessário apenas considerar um vetor $\mathbf{d}$ com norma igual a um, como requisitado na expressão (23). Se $\mathbf{d} = [\text{cost} \quad \text{sent}]^T$, tem-se, $\|\mathbf{d}\| = \sqrt{(\text{cost})^2 + (\text{sent})^2} = 1$ e, então, a expressão (23) pode ser utilizada corretamente. Desse modo:

$$\begin{aligned} \begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix} &= \begin{bmatrix} \hat{\theta}_1 \\ \hat{\theta}_2 \end{bmatrix} + k \cdot \begin{bmatrix} a & b \\ c & d \end{bmatrix} \cdot \begin{bmatrix} \text{cost} \\ \text{sent} \end{bmatrix} \Rightarrow \begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix} = \begin{bmatrix} \hat{\theta}_1 \\ \hat{\theta}_2 \end{bmatrix} + \begin{bmatrix} ka & kb \\ kc & kd \end{bmatrix} \cdot \begin{bmatrix} \text{cost} \\ \text{sent} \end{bmatrix} \Rightarrow \\ \begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix} &= \begin{bmatrix} \hat{\theta}_1 \\ \hat{\theta}_2 \end{bmatrix} + \begin{bmatrix} ka \text{ cost} + kb \text{ sent} \\ kc \text{ cost} + kd \text{ sent} \end{bmatrix} \Rightarrow \begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix} = \begin{bmatrix} \hat{\theta}_1 + ka \text{ cost} + kb \text{ sent} \\ \hat{\theta}_2 + kc \text{ cost} + kd \text{ sent} \end{bmatrix}. \end{aligned}$$

Assim, encontrar a região de confiança conjunta é encontrar todos os pontos do plano $\theta_1\theta_2$, tais que $\theta_1 = \hat{\theta}_1 + ka \text{ cost} + kb \text{ sent}$ e $\theta_2 = \hat{\theta}_2 + kc \text{ cost} + kd \text{ sent}$. Ou seja, o que se deseja é o gráfico da função $f : \Re \rightarrow \Re^2$, tal que $f(t) = (\hat{\theta}_1 + ka \text{ cost} + kb \text{ sent}, \hat{\theta}_2 + kc \text{ cost} + kd \text{ sent})$, que é o gráfico de uma elipse centrada em $(\hat{\theta}_1, \hat{\theta}_2)$. Para encontrar o traço de tal função, foi elaborada uma rotina no software R versão 2.7.0, de forma semelhante ao que foi exposto acima. Assim, foram encontradas, para cada conjunto de dados, as regiões, por meio da aproximação linear.

3.2.5 Intervalos de confiança, correlação entre os parâmetros e R^2

Os intervalos de confiança marginais para os parâmetros e as estimativas de correlações entre eles foram encontrados utilizando-se a aproximação linear por meio de funções próprias do software R versão 2.7.0, tais como “*confint*” para o intervalo de confiança e “*vcov*” para a matriz de covariâncias, como mostrado no anexo 1. O coeficiente de determinação foi encontrado montando-se rotina própria considerando a fórmula:

$$R^2 = 1 - \frac{S(\hat{\theta})}{(\mathbf{y} - \bar{y} \cdot \mathbf{1})^T (\mathbf{y} - \bar{y} \cdot \mathbf{1})}$$

em que $\mathbf{1} = [1 \ 1 \ \dots \ 1]^T$ de dimensões $N \times 1$.

3.2.6 Estudos de simulação: “bootstrap” paramétrico

O estudo de simulação foi realizado usando o modelo apresentado por Marchi (2005), envolvendo os dados amostrais de seu estudo, tendo dois parâmetros (C_l e k),

$$C_t = C_l (1 - e^{-kt}) + \xi_t$$

em que ξ_t é assumido ter uma distribuição normal com média zero e variância σ^2 . O “verdadeiro” valor de C_l , k e σ^2 foi considerado como sendo suas estimativas de mínimos quadrados $\hat{C}_l$, $\hat{k}$ e $\hat{\sigma}^2$ (com quatro casas decimais sendo consideradas para grande precisão nos cálculos subseqüentes). Como $N = 45$, cada conjunto de dados consiste de 45 pares de elementos, a saber, (t_1, C_1) , (t_2, C_2) , ..., (t_{45}, C_{45}) , e respectivamente os erros $\xi_1, \xi_2, \dots, \xi_{45}$ que foram obtidos pseudo-aleatoriamente, utilizando a função “*rnorm*” do software computacional R versão 2.7.0. Esse novo conjunto de dados foi, então, ajustado pelos mínimos quadrados para a obtenção de uma “nova” estimativa de C_l e k . Cinco mil conjuntos de dados do mesmo modo foram gerados e forneceram 5.000 estimativas $\hat{C}_l^*$ e $\hat{k}^*$ de $\hat{C}_l$ e $\hat{k}$ (parâmetros populacionais no estudo de simulação).

Com essas 5000 estimativas foram construídos histogramas de distribuição de frequência, calculado o viés e os intervalos de confiança “bootstrap”. Foram também encontrados os primeiros quatro momentos amostrais centrado na média dessas 5000 estimativas, e as quantidades derivadas deles para serem apresentadas em tabelas, com o objetivo da realização do teste de assimetria e curtose assintótico.

Os momentos amostrais foram encontrados com o uso da fórmula apresentada por Mood et al. (1974):

$$m_r = \frac{\sum_{i=1}^n (X_i - \bar{X})^r}{n}$$

em que $n = 5000$, $X_i - \bar{X}$ é o desvio de cada estimativa em relação a média “bootstrap” e m_r o r – ésimo momento amostral. Depois de encontrados os momentos amostrais foram calculadas as quantidades de interesse $g_1 = m_3/m_2^{3/2}$ para avaliar a assimetria e $g_2 = m_4/m_2^2 - 3$ para avaliar o excesso de curtose, os testes foram realizados tendo em vista que $g_1 \sim N\left(0, (6/5000)^{\frac{1}{2}}\right)$ e $g_2 \sim N\left(0, (24/5000)^{\frac{1}{2}}\right)$, conforme Snedecor & Cochran (1980).

O viés foi obtido como se segue, visto que o verdadeiro valor para os parâmetros assumido nas simulações foram as estimativas $\hat{C}_l$ e $\hat{k}$ calculadas sobre os dados reais. O viés, então, seria a diferença entre as médias dos valores obtidos via simulação e os valores $\hat{C}_l$ e $\hat{k}$. O resultado do viés “bootstrap” foi apresentado percentualmente e testado com base na distribuição exposta em Souza (1998), adaptada para estudo da seguinte forma:

$$z_i = \sqrt{5000} \frac{\bar{\hat{C}}_l^* - \hat{C}_l}{\sqrt{\text{Var}(\hat{C}_l)}} \sim N(0,1) \text{ e } z_i = \sqrt{5000} \frac{\bar{\hat{k}}^* - \hat{k}}{\sqrt{\text{Var}(\hat{k})}} \sim N(0,1)$$

Para verificar se as estimativas seguiam uma distribuição normal, além dos histogramas, foram realizados também o teste de normalidade de Shapiro-Wilk e o teste de assimetria e de excesso de curtose assintótico.

3.2.6.1 Intervalos de confiança “bootstrap”

Os intervalos de confiança “bootstrap” foram gerados por meio do método percentil t. Este método tem por base, unicamente, os quantis e outras medidas da distribuição “bootstrap” do estimador de interesse. Considerando uma amostra aleatória de tamanho n e $F(x)$ a função de distribuição para a população referente a essa amostra, com $\hat{F}(x)$ sendo um estimador de $F(x)$ com base nessa amostra, então, denotando por θ_p^* e $\hat{\sigma}^*$ as quantidades $\hat{\theta}_p$ e $\hat{\sigma}$ calculadas para as amostras (“bootstrap”) de $\hat{F}(x)$, um tipo de intervalo de confiança percentil t para θ_p , a $100(1-\alpha)\%$, tem a forma (Souza, 1998):

$$\left(\hat{\theta}_p - t_{1-\alpha/2}^{(1)} \frac{\hat{\sigma}}{\sqrt{n}}, \hat{\theta}_p - t_{\alpha/2}^{(1)} \frac{\hat{\sigma}}{\sqrt{n}} \right)$$

em que $t_{\beta}^{(1)}$ satisfaz

$$P_{\hat{F}} \left\{ \sqrt{n} \left(\frac{\hat{\theta}_p^* - \hat{\theta}_p}{\hat{\sigma}^*} \right) \leq t \right\} = \beta$$

Dessa forma, foram calculadas 5.000 estimativas “bootstrap” para cada parâmetro e, em seguida, essas estimativas foram ordenadas e padronizadas para se encontrar o valor $t_{\beta}^{(1)}$ referente a cada parâmetro. Considerou-se $\alpha = 5\%$, Assim os valores $t_{\beta}^{(1)}$ foram $t_{0,025}^{(1)}$ e $t_{0,975}^{(1)}$, que são os valores da amostra ordenada condizentes com as seguintes igualdades:

$$P_{\hat{F}} \left\{ \sqrt{5000} \left(\frac{\hat{\theta}_p^* - \hat{\theta}_p}{\hat{\sigma}^*} \right) \leq t_{0,975}^{(1)} \right\} = 0,975 \text{ e } P_{\hat{F}} \left\{ \sqrt{5000} \left(\frac{\hat{\theta}_p^* - \hat{\theta}_p}{\hat{\sigma}^*} \right) \leq t_{0,025}^{(1)} \right\} = 0,025$$

Os intervalos de confiança “bootstrap” para os parâmetros foram encontrados substituindo os resultados na expressão:

$$\left(\hat{\theta}_p - t_{0,975}^{(1)} \frac{\hat{\sigma}}{\sqrt{5000}}, \hat{\theta}_p - t_{0,025}^{(1)} \frac{\hat{\sigma}}{\sqrt{5000}} \right)$$

3.2.7 Estudo da curvatura de Bates e Watts

O estudo da curvatura de Bates e Watts foi realizado com o uso da função “*rms.curv*” do pacote “*MASS*” do software estatístico R versão 2.7.0. Essa função fornece os valores de $c\sqrt{F}$, considerando c igual a c^0 para a curvatura RMS de efeito parâmetro e c^l para a curvatura RMS intrínseca. Para uso dessa função, é necessário também utilizar a função “*deriv3*”, que armazena os “*arrays*” de hessianos mostrados no item 2.8.3. A rotina para o uso dessa função consta no anexo 3.

4 RESULTADOS E DISCUSSÃO

4.1 Valores iniciais

Os valores iniciais fornecidos produziram convergência para todos os ajustes requeridos, os quais estão dispostos na Tabela 4.1. Isso indica que a metodologia apresentada para esse modelo pode também ser aproveitada em outras situações similares que o utilizem. Ressalta-se que a metodologia foi específica para tal modelo e suas peculiaridades, não podendo ser aplicado em qualquer outro modelo ou situações que difiram muito da apresentada.

TABELA 4.1: Estimativas iniciais dos parâmetros C_l e k no ajuste do modelo não linear para descrição da extração de zinco (Zn) por diferentes extratores.

Parâmetro	Extratores						
	Acético	DTPA	Lático	Mehlich	MOAS	NH ₄ OAc	SOAS
C_l^0	28,40	23,64	37,93	37,37	34,91	9,85	35,62
k^*	0,29	0,42	0,31	0,49	0,29	0,16	0,29

4.2 Estimativas e resultados inferenciais

Na Tabela 4.2 são apresentadas as estimativas encontradas para os parâmetros do modelo, conforme a metodologia estabelecida, bem como os valores do erro padrão da estimativa, os intervalos de confiança a 95%, a estimativa de correlação entre os parâmetros para cada conjunto de dados, o valor do coeficiente de determinação R^2 e o erro padrão residual. Percebe-se que todos os parâmetros da função foram significativos, pois se encontram dentro do intervalo de confiança assintótico de 95% e estes, por sua vez, não incluem a constante zero.

TABELA 4.2: Estimativas dos parâmetros, erro padrão da estimativa (E.P.E), valor de p para o teste t , intervalo de confiança (IC) de 95% para os parâmetros, estimativa do coeficiente de correlação entre as estimativas dos parâmetros, coeficiente de determinação (R^2) e erro padrão residual (E.P.R.) para a descrição da extração de zinco em um solo tratado com lodo de esgoto por diferentes extratores.

Extrator	Parâmetro	Estimativa	E.P.E.	Valor-P	IC(95%)	Correlação	R^2	E.P.R.
Acético	C_l	26,9449	0,2110	$< 2.10^{-16}$	[26,5106;27,3960]	-0,4072	0,9716	0,8697
	k	0,3673	0,0118	$< 2.10^{-16}$	[0,3427;0,3940]			
DTPA	C_l	21,5538	0,1993	$< 2.10^{-16}$	[21,1442;21,9681]	-0,2736	0,5177	1,2240
	k	1,3313	0,1150	$8,55.10^{-15}$	[1,0959;1,6645]			
Lático	C_l	34,9237	0,3914	$< 2.10^{-16}$	[34,1246;35,7534]	-0,3605	0,9005	1,8920
	k	0,4829	0,0265	$< 2.10^{-16}$	[0,4293;0,5456]			
Mehlich I	C_l	35,4071	0,1878	$< 2.10^{-16}$	[35,0239;35,7922]	-0,2756	0,8080	1,1470
	k	1,2680	0,0606	$< 2.10^{-16}$	[1,1439;1,4161]			
MOAS	C_l	31,9402	0,2967	$< 2.10^{-16}$	[31,3276;32,5752]	-0,3716	0,9376	1,3840
	k	0,4493	0,0195	$< 2.10^{-16}$	[0,4085;0,4956]			
NH₄OAc	C_l	10,9697	0,3897	$< 2.10^{-16}$	[10,2316;11,9048]	-0,6683	0,9725	0,3855
	k	0,1193	0,0081	$< 2.10^{-16}$	[0,1026;0,1370]			
SOAS	C_l	32,2656	0,3482	$< 2.10^{-16}$	[31,5488;33,0124]	-0,3697	0,9157	1,6340
	k	0,4547	0,0231	$< 2.10^{-16}$	[0,4069;0,5101]			

A quantidade de ajuste fornecido pelo coeficiente de determinação para os extratores ácido acético e NH_4OAc foram altos, com R^2 igual a 97,16% e 97,25%, respectivamente; para os extratores de ácido láctico, MOAS e SOAS, pode-se dizer que o ajuste não foi tão alto, com R^2 igual a 90,05%, 93,76% e 91,57% respectivamente. No entanto, os extratores DTPA e Mehlich I apresentaram um ajuste baixo, com R^2 igual a 51,77% e 80,80% respectivamente.

Todos os valores das correlações entre os parâmetros foram negativos, indicando que à medida que um determinado parâmetro cresce o outro tende a diminuir, ou seja, se determinado extrator tem uma elevada taxa de extração, então, a quantidade de metal extraído deve estabilizar mais rapidamente.

4.3 Resultados gráficos

Nas Figuras de 1 a 7 são apresentados os ajustes para cada extrator. Na Figura 4.1 é mostrado o ajuste para o extrator ácido acético, tendo bandas de confiança para resposta esperada delimitando uma região relativamente pequena entre elas, tendendo a diminuir à medida que o número de extrações cresce. Todos os pontos pertencentes à região de confiança conjunta fazem sentido prático. A taxa de extração para o Zn com o ácido acético é de 0,3673, apresentando uma estabilidade das extrações acumuladas de forma moderada, durante as 15 extrações.

As bandas de confiança para o DTPA foram quase coincidentes, conforme Figura 4.2. Os intervalos de confiança marginais delimitaram um retângulo que é quase coincidente com a região de confiança conjunta. Todos os valores da região conjunta fazem sentido prático. Este extrator retirou quase toda a fração extraível na primeira vez, alcançando a estabilidade muito rapidamente.

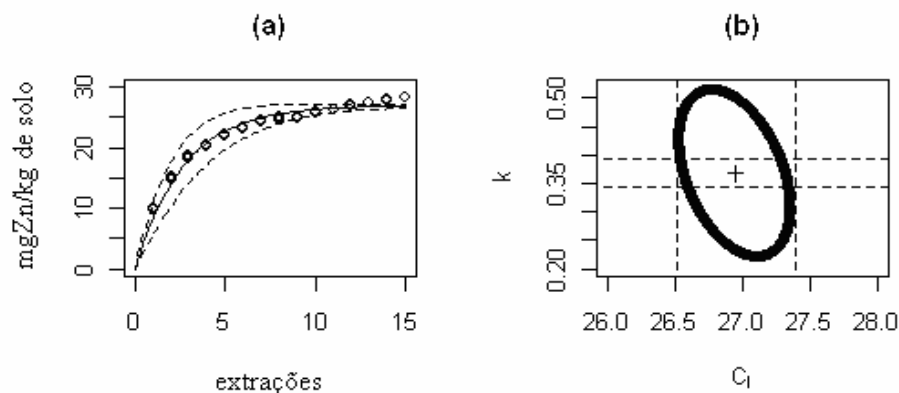


FIGURA 4.1: Representação gráfica da quantidade de zinco extraída em um solo tratado com lodo de esgoto pelo extrator **ácido acético** em relação ao número de extrações, bandas de confiança a 95% para a resposta esperada (a), e região de confiança conjunta a 95% para os parâmetros C_1 e k (b). O símbolo “+” indica a posição da estimativa e as linhas pontilhadas representam os intervalos de confiança marginais.

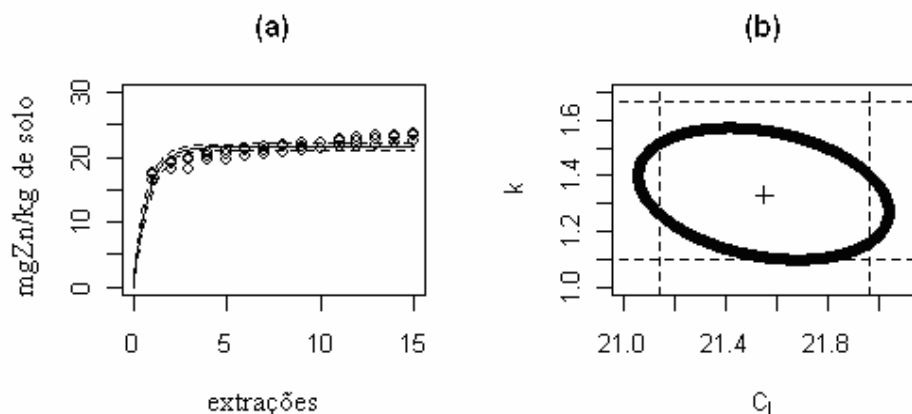


FIGURA 4.2: Representação gráfica da quantidade de zinco extraída em um solo tratado com lodo de esgoto pelo extrator **DTPA** em relação ao número de extrações, bandas de confiança a 95% para a resposta esperada (a), e região de confiança conjunta a 95% para os parâmetros C_1 e k (b). O símbolo “+” indica a posição da estimativa e as linhas pontilhadas representam os intervalos de confiança marginais.

Observa-se que, para o extrator de ácido láctico, a estabilização ocorre próximo da sétima extração, conforme Figura 4.3, apresentando resultados próximos aos do ácido acético (Figura 4.1). No entanto, o valor para a estabilização C_i deste extrator foi bem superior ao do ácido acético. As bandas de confiança estão bem distribuídas por toda a curva, indicando que previsões dentro dos limites estabelecidos das extrações apresentarão boa precisão. A região de confiança conjunta apresenta todos os resultados com sentido prático, tanto para o parâmetro que representa a taxa de extração quanto para o que representa a estabilização desta extração.

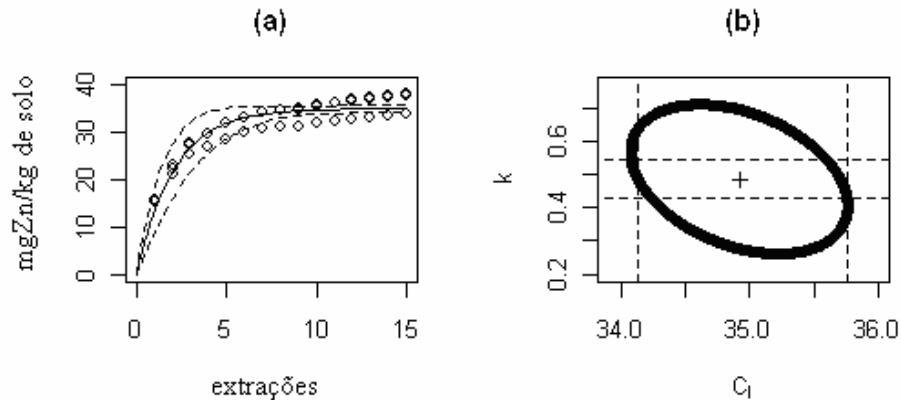


FIGURA 4.3: Representação gráfica da quantidade de zinco extraída em um solo tratado com lodo de esgoto pelo extrator **ácido láctico** em relação ao número de extrações, bandas de confiança a 95% para a resposta esperada (a), e região de confiança conjunta a 95% para os parâmetros C_i e k (b). O símbolo “+” indica a posição da estimativa e as linhas pontilhadas representam os intervalos de confiança marginais.

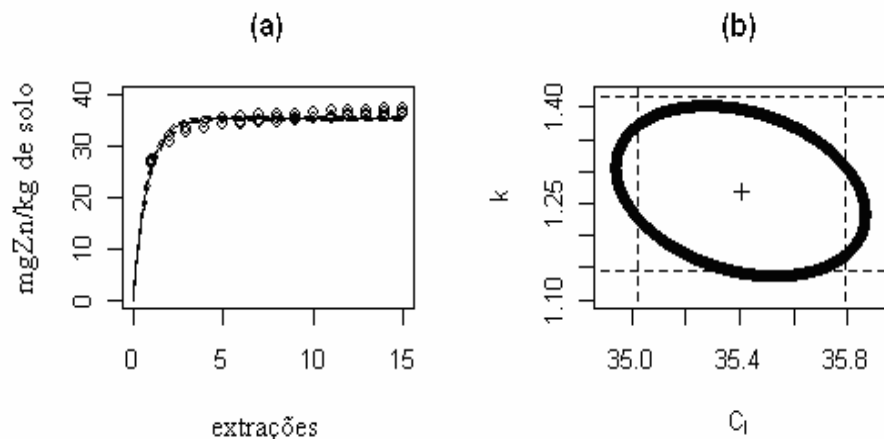


FIGURA 4.4: Representação gráfica da quantidade de zinco extraída em um solo tratado com lodo de esgoto pelo extrator **Mehlich I** em relação ao número de extrações, bandas de confiança a 95% para a resposta esperada (a), e região de confiança conjunta a 95% para os parâmetros C_l e k (b). O símbolo “+” indica a posição da estimativa e as linhas pontilhadas representam os intervalos de confiança marginais.

Como se observa na Figura 4.4, o extrator Mehlich I extraiu quase toda a porção disponível nas primeiras extrações, assim como tinha ocorrido com o DTPA (figura 4.2). Os resultados para estes dois extratores foram bem semelhantes, também com relação às bandas de confiança para a resposta esperada. No entanto, a constante de estabilidade C_l para o Mehlich I é bem maior do que para o DTPA.

O extrator MOAS (Figura 4.5) apresentou resultados semelhantes aos do ácido acético e do ácido láctico, com relação aos ajustes e às bandas de confiança, sendo muito semelhante com o láctico e diferindo um pouco do acético com relação à constante de estabilidade C_l .

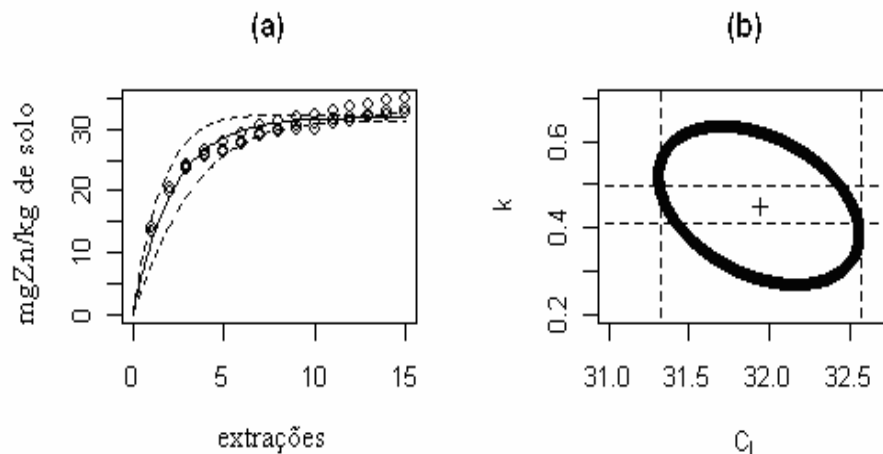


FIGURA 4.5: Representação gráfica da quantidade de zinco extraída em um solo tratado com lodo de esgoto pelo extrator **MOAS** em relação ao número de extrações, bandas de confiança a 95% para a resposta esperada (a), e região de confiança conjunta a 95% para os parâmetros C_l e k (b). O símbolo “+” indica a posição da estimativa e as linhas pontilhadas representam os intervalos de confiança marginais.

O caso mais crítico dos sete extratores é o do NH_4OAc , que apresentou valores experimentais ajustando-se aproximadamente em linha reta, como visto na Figura 4.6. Mesmo que o modelo tenha fornecido um bom ajuste com relação ao coeficiente de determinação R^2 , percebem-se alguns problemas com as análises referentes a esse extrator. Por exemplo, as bandas de confiança determinaram regiões muito amplas para as respostas esperadas, indicando que uma possível previsão para a extração do Zn teria uma resposta pouco confiável. Além disso, este extrator extrai o elemento traço Zn muito vagarosamente, não chegando a indicar, com os dados experimentais, um valor de estabilização C_l próximo do estimado.

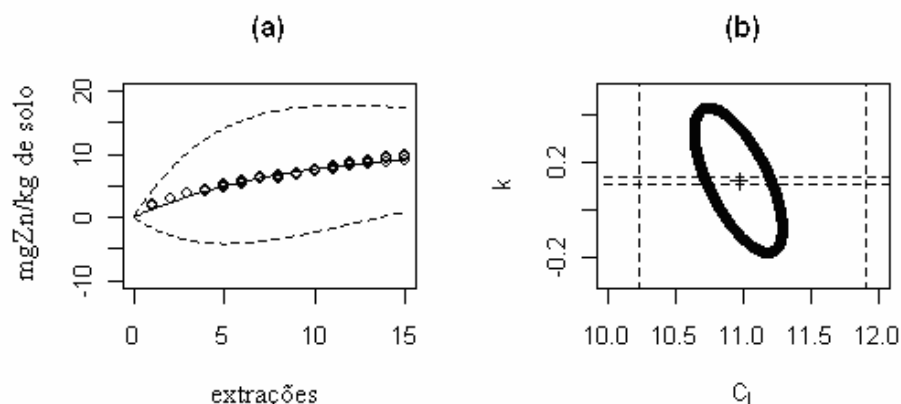


FIGURA 4.6: Representação gráfica da quantidade de zinco extraída em um solo tratado com lodo de esgoto pelo extrator NH_4OAc em relação ao número de extrações, bandas de confiança a 95% para a resposta esperada (a), e região de confiança conjunta a 95% para os parâmetros C_1 e k (b). O símbolo “+” indica a posição da estimativa e as linhas pontilhadas representam os intervalos de confiança marginais.

Os resultados do SOAS, MOAS e ácido láctico foram muito semelhantes, indicando que esses extratores podem prever a disponibilidade do elemento Zn com igual eficiência, conforme pode ser visto na Figura 4.7.

Nas Figuras de 4.1 a 4.7, pode-se notar, ainda, que a quantia de Zn extraída pelo ácido láctico foi mais alta do que a extraída pelo ácido acético. Isto ocorreu, provavelmente, devido ao fato de que as constantes de estabilidade C_1 entre o Zn e os ácidos são mais altas para o láctico que para o acético (Gustafsson, 2006)

Os níveis de Zn extraídos por Mehlich I foram altamente significativos, quando comparados aos outros extratores. Nota-se, ainda, que a transferência do Zn do solo para a solução de ácidos orgânicos é mais lenta do que para a de DTPA e de Mehlich I e a quantidade extraída de Zn, pelo extrator NH_4OAc , durante as 15 extrações sucessivas, foi mais baixa e mais lenta do que para todos os outros extratores.

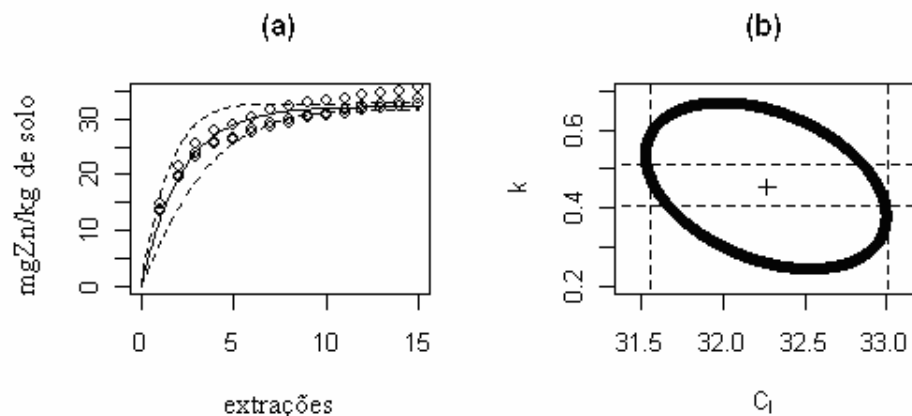


FIGURA 4.7: Representação gráfica da quantidade de zinco extraída em um solo tratado com lodo de esgoto pelo extrator **SOAS** em relação ao número de extrações, bandas de confiança a 95% para a resposta esperada (a), e região de confiança conjunta a 95% para os parâmetros C_1 e k (b). O símbolo “+” indica a posição da estimativa e as linhas pontilhadas representam os intervalos de confiança marginais.

A quantia cumulativa de Zn removida na décima quinta extração ($C_{t_{15}}$) foi maior para o Mehlich I nas quinze extrações sucessivas do que para qualquer outro extrator. Considerando apenas uma única extração para o Zn, no solo tratado com lodo de esgoto, nota-se que, para o extrator Mehlich I, ela foi maior do que para os outros extratores estudados. Apesar da grande quantia de Zn extraída pelo Mehlich I, a taxa de extração (k) relatada para a extração pelo DTPA foi a maior em relação aos demais extratores.

Os resultados cinéticos mostraram que o extrator de ácido acético extraiu menores concentrações de Zn do que os outros ácidos orgânicos.

Os resultados para todos os extratores com relação ao Zn foram idênticos aos apresentados por Marchi (2005), no entanto, em seu trabalho não constava à relação entre as bandas de confiança para a resposta esperada, de

modo que, não foi notado que as respostas esperadas para o extrator NH_4OAc apresentam alta imprecisão com relação às predições a serem realizadas, ou seja, ao se realizar determinada predição da quantidade de elemento-traço extraída para uma determinada extração, tem-se um intervalo de confiança amplo para este extrator, quando se considera um nível de significância fixo, como, por exemplo, 5%.

Os ajustes e as inferências realizados até aqui foram todos realizados por meio da aproximação linear. Para uma validação ou não desses resultados foi realizado também uma simulação “bootstrap” e a verificação de uma boa aproximação linear foi realizada por meio do estudo da curvatura de Bates e Watts, como apresentado a seguir.

4.4 Simulação “bootstrap”

As Figuras de 4.8 a 4.14 são os histogramas referentes à simulação “bootstrap” para as estimativas de cada um dos parâmetros. A distribuição das estimativas é visualmente próxima da distribuição normal.

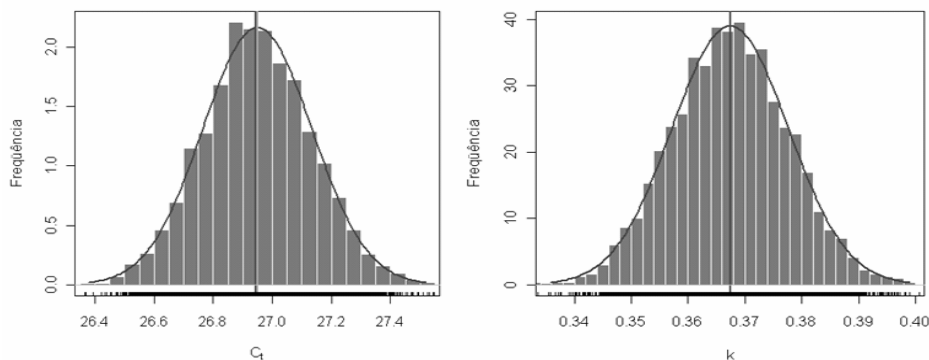


FIGURA 4.8: Histograma das estimativas de C_l e k fornecido via simulação “bootstrap” (paramétrico), para as concentrações cumulativas de Zn obtidas pelas extrações sucessivas com **ácido acético**.

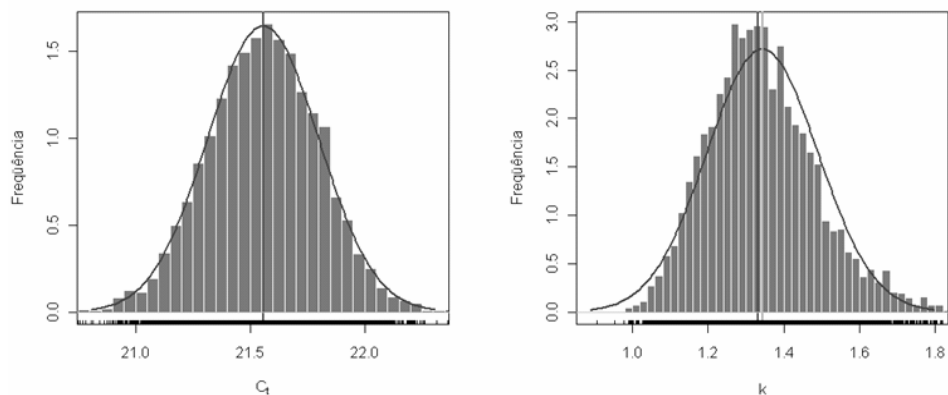


FIGURA 4.9: Histograma das estimativas de C_l e k fornecido via simulação “bootstrap” (paramétrico), para as concentrações cumulativas de Zn obtidas pelas extrações sucessivas com **DTPA**.

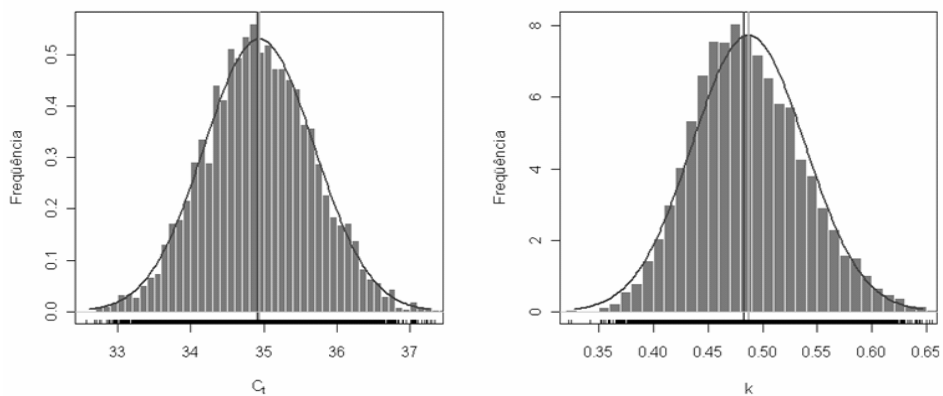


FIGURA 4.10: Histograma das estimativas de C_l e k fornecido via simulação “bootstrap” (paramétrico), para as concentrações cumulativas de Zn obtidas pelas extrações sucessivas com **ácido láctico**.

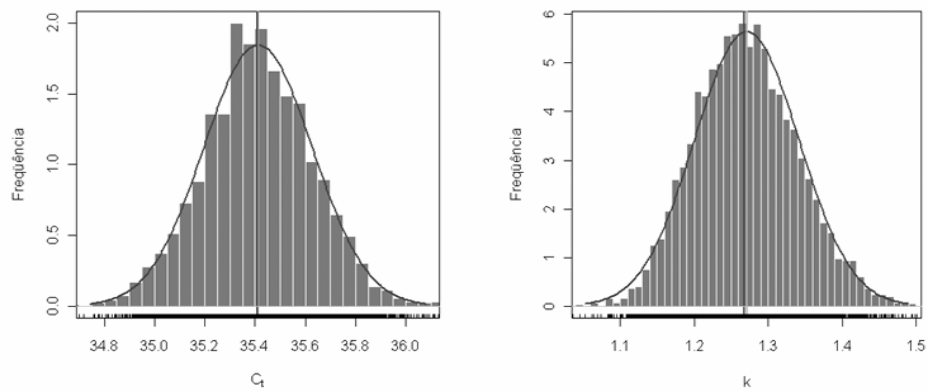


FIGURA 4.11: Histograma das estimativas de C_l e k fornecido via simulação “bootstrap” (paramétrico), para as concentrações cumulativas de Zn obtidas pelas extrações sucessivas com **Mehlich I**

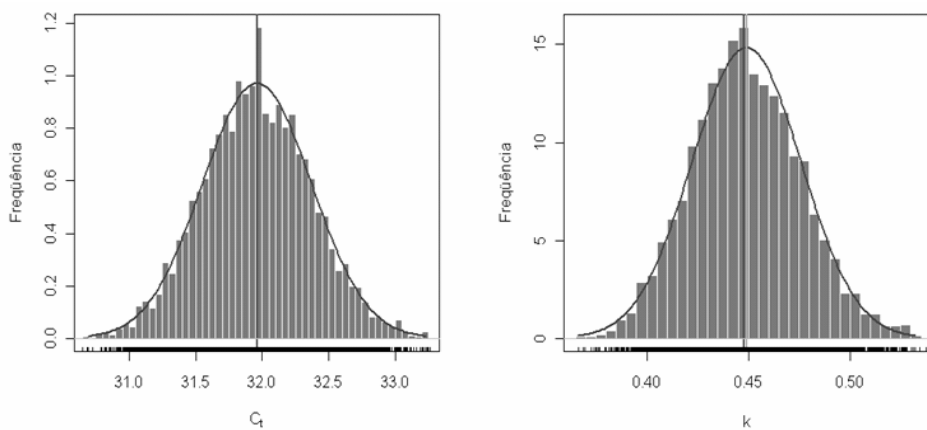


FIGURA 4.12: Histograma das estimativas de C_l e k fornecido via simulação “bootstrap” (paramétrico), para as concentrações cumulativas de Zn obtidas pelas extrações sucessivas com **MOAS**.

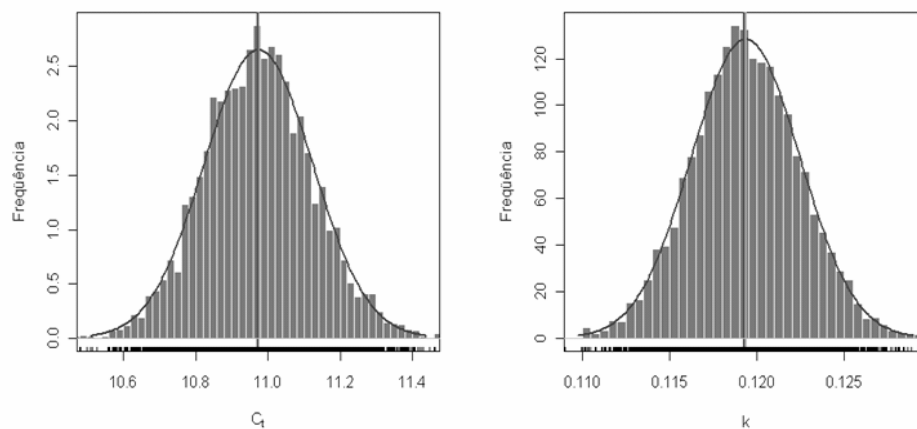


FIGURA 4.13: Histograma das estimativas de C_l e k fornecido via simulação “bootstrap” (paramétrico), para as concentrações cumulativas de Zn obtidas pelas extrações sucessivas com **NH₄OAC**.

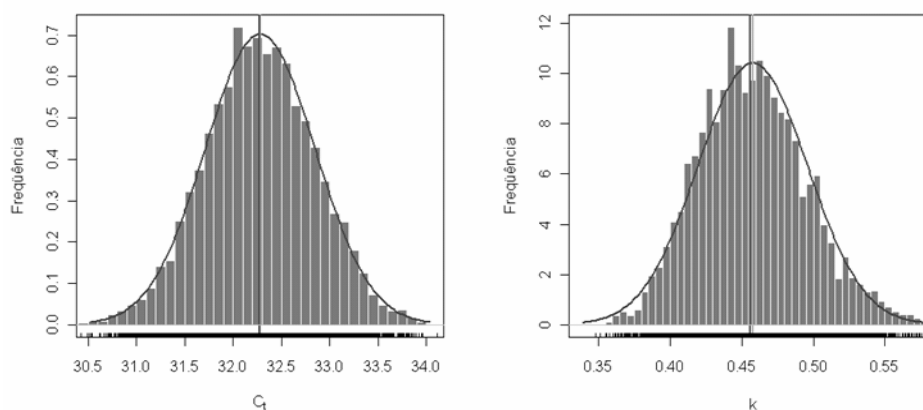


FIGURA 4.14: Histograma das estimativas de C_l e k fornecido via simulação “bootstrap” (paramétrico), para as concentrações cumulativas de Zn obtidas pelas extrações sucessivas com **SOAS**.

A análise visual dos histogramas é apenas uma forma empírica de avaliar a distribuição das estimativas em modelos não lineares. Em uma análise mais elaborada, foram realizados testes de assimetria e curtose assintótico e o teste de Shapiro-Wilk (Tabela 4.4). Os testes de assimetria e curtose indicam um distanciamento da distribuição normal para as estimativas do parâmetro k para os dados referentes às quantidades de Zn extraídos com as soluções: SOAS, Mehlich I, Lático e DTPA. Esses mesmos resultados são reforçados pelo teste de Shapiro-Wilk. O teste de assimetria e curtose não destacou em absoluto a normalidade para as estimativas do parâmetro C_l , para nenhum dos extratores, no entanto, o teste de Shapiro-Wilk indicou normalidade apenas para os dados de Zn extraídos com as soluções: acético, DTPA, MOAS e SOAS.

Confrontando-se os aspectos gráficos e os testes propostos, pode-se dizer que, embora para alguns extratores, as estimativas dos parâmetros referentes aos dados da extração cumulativa de Zn possam não ter distribuição normal, esse desvio de normalidade não é tão grande, como observado pelos histogramas.

Segundo os resultados apresentados na Tabela 4.4, o estimador para C_l é viesado para os dados de Zn provenientes das extrações sucessivas com lático, Mehlich I, MOAS e SOAS e, para o estimador de k daquelas com DTPA, lático, Mehlich I, e SOAS. Ou seja, os estimadores referente aos dados de Zn extraídos sucessivamente por todos extratores apresentam vício. No entanto, o viés percentual não alcançou 1% o que, segundo Ratkowsky (1983), é considerado baixo, podendo ser desprezado.

Comparando-se os resultados da Tabela 4.4 com os da Tabela 4.2, pode-se constatar que os intervalos de confiança “bootstrap” são muito semelhantes aos intervalos de confiança obtidos pela aproximação linear, validando ainda mais os resultados obtidos anteriormente.

TABELA 4.4: Resultados das inferências “bootstrap”: estimativas dos parâmetros obtidos pela aproximação linear (E.A.L.), média “bootstrap”, Estatística W de Shapiro-Wilk, viés “bootstrap”, intervalo de confiança “bootstrap” (IC) de 95% para os parâmetros, estatística g_1 para o teste de assimetria assintótico, estatística g_2 para o teste de excesso de curtose assintótico.

Extrator	P	E.A.L.	Média	W	Viés%	IC (95%)	Assimetria (g_1)	E.C. (g_2)
Acético	C_l	26,9449	26,9493	0,9996 ^(ns)	0,0164 ^(ns)	[26,5736;27,3272]	0,0688*	0,0353 ^(ns)
	k	0,3673	0,3674	0,9996 ^(ns)	0,0416 ^(ns)	[0,3477;0,3889]	0,0543 ^(ns)	-0,0277 ^(ns)
DTPA	C_l	21,5538	21,5571	0,9970 ^(ns)	0,01526 ^(ns)	[21,0732;22,0433]	-0,0413 ^(ns)	0,0073 ^(ns)
	k	1,3313	1,3434	0,9812*	0,9056*	[1,1006;1,6948]	0,5968*	0,7114*
Lático	C_l	34,9237	34,9383	0,9990*	0,0417*	[33,4804;36,5268]	0,1105*	0,0682 ^(ns)
	k	0,4829	0,4872	0,9906*	0,8902*	[0,3948;0,6053]	0,4140*	0,4082*
Mehlich-I	C_l	35,4071	35,4119	0,9993*	0,01391*	[34,9748;35,8516]	0,0551 ^(ns)	0,1137 ^(ns)
	k	1,2680	1,2711	0,9951*	0,2472*	[1,1413;1,4240]	0,2851*	0,2852*
MOAS	C_l	31,9402	31,9617	0,9996 ^(ns)	0,0670*	[31,1327;31,7900]	0,0615*	0,0205 ^(ns)
	k	0,4493	0,4491	0,9977*	0,1850 ^(ns)	[0,3999;0,5092]	0,185*	-0,0047 ^(ns)
NH₄OAc	C_l	10,9697	10,9739	0,9989*	0,0388 ^(ns)	[10,6735;11,2892]	0,1135*	0,0783 ^(ns)
	k	0,1193	0,1193	0,9996 ^(ns)	0,0071 ^(ns)	[0,1132;0,1257]	0,0508 ^(ns)	0,0334 ^(ns)
SOAS	C_l	32,2656	32,2805	0,9998 ^(ns)	0,0464*	[31,1786;33,4789]	0,0244 ^(ns)	0,0009 ^(ns)
	k	0,4547	0,4577	0,9929*	0,6647*	[0,3866;0,5397]	0,3503*	0,1606*

P: parâmetro; **(ns):** não significativo a 5%; *****: significativo a 5%; **E.C.:** excesso de curtose;

4.5 Medidas de curvatura de Bates e Watts

Mesmo após a avaliação da simulação “bootstrap” para uma possível validação dos resultados inferenciais obtidos via aproximação linear, foi realizado também um estudo da não linearidade do modelo para cada conjunto de dados. O objetivo foi o de avaliar o quão distante estão as propriedades dos estimadores não lineares de suas propriedades assintóticas, devido ao tamanho reduzido da amostra.

Na tabela 4.5 é apresentado o valor das curvaturas RMS, propostas por Bates & Watts (1980) para analisar o quão próximo está um modelo não linear de um modelo linear. Na mesma Tabela, se encontram os resultados para os dois tipos de curvatura: a curvatura de efeito parâmetro e a curvatura intrínseca. A curvatura intrínseca foi menor do que 0,1 para todos os extratores, indicando que o local da solução na vizinhança das estimativas é relativamente plano, tornando, assim, verdadeira a aproximação linear por um plano. Como essa medida não ultrapassou 0,1, pode-se dizer que, nas proximidades das estimativas, o “local da solução” é, aproximadamente, 95% plano, o que é um bom resultado, segundo Bates & Watts (1988).

Observa-se também que a curvatura de efeito parâmetro apresentou resultado entre 0,2 e 0,3, para os dados de Zn obtidos com todos os extratores, exceto para o extrator NH_4OAc . Esse resultado indica que, para os extratores acético, DTPA, lático, Mehlich I, MOAS e SOAS, a aproximação das “coordenadas uniformes” é em torno de 90% para um modelo linear, o que também é considerado baixo para os autores. Isso indica que a forma como parâmetros aparecem no modelo não prejudica demasiadamente as inferências pela aproximação linear, mesmo que o tamanho da amostra seja pequeno.

TABELA 4.5: Medidas de curvatura quadrática média (RMS) para a descrição da extração de zinco em um solo tratado com lodo de esgoto por diferentes extratores.

Extrator	Curvatura de efeito parâmetro	Curvatura intrínseca
Acético	0,1407	0,0342
DTPA	0,2606	0,0976
Lático	0,2009	0,0655
Mehlich-I	0,1429	0,0551
MOAS	0,1659	0,0506
NH₄OAc	0,9164	0,0286
SOAS	0,1928	0,0595

O extrator NH₄OAc apresentou curvatura intrínseca baixa, de apenas 0,0286, indicando que a aproximação por um plano na vizinhança das estimativas pode ser considerada boa. No entanto, este extrator apresentou curvatura de efeito parâmetro muito alta, em torno de 0,9164, indicando que ou a forma com que os parâmetros aparecem no modelo são inadequadas ou o modelo pode ser inadequado para este conjunto de dados, tendo em vista que as inferências obtidas via aproximação linear e teoria assintótica podem ser inválidas, pois as propriedades dos estimadores referente a este modelo estão muito distantes em relação às propriedades dos estimadores lineares.

O resultado ruim da curvatura de efeito parâmetro para o extrator NH₄OAc pode também ser justificado pelo comportamento dos dados com relação ao modelo. Nota-se que, embora o ajuste tenha sido em torno de 97,25% (Tabela 4.2), os dados amostrais não apresentaram uma assíntota e uma estabilização nas quinze extrações sucessivas, como ocorreu com os demais extratores. Isso pode ter ocorrido, ou por algum problema experimental, ou

porque o modelo não é adequado para este extrator ou, ainda, porque o modelo precisa ser reajustado para fornecer resultados inferenciais melhores. O reajuste do modelo pode ser realizado por uma reparametrização adequada, pois segundo Bates & Watts (1980), as reparametrizações não mudam a forma da superfície esperada, nem o efeito da curvatura intrínseca, que é uma propriedade do local da solução. Assim, uma boa reparametrização pode mudar consideravelmente a curvatura de efeito parâmetro, tornando mais válidos os resultados das inferências que foram obtidos pela aproximação linear e pela teoria assintótica.

No trabalho de Marchi (2005) não foi realizado o estudo da curvatura de Bates & Watts (1980), não sendo possível dessa forma, identificar os problemas inferenciais apresentados pelo extrator NH_4OAc , bem como tornar mais confiável os resultados obtidos para os outros extratores.

Uma melhor reparametrização poderá ser estudada em trabalhos futuros, preferencialmente, com mais conjuntos de dados referentes ao mesmo extrator, tendo em vista que cada análise é realizada de forma particular para cada modelo, delineamento e conjunto de dados.

5 CONCLUSÕES

- 1) O modelo teve ajuste superior a 80% para todos os extratores, utilizando a metodologia da aproximação linear, exceto para o extrator DTPA, considerando como critério o R^2 e apresentou ajustes mais altos para os extratores de ácido acético e NH_4OAc .
- 2) O extrator NH_4OAc apresentou bandas de confiança com dimensões muito amplas, tornando, assim, imprecisas as predições realizadas no intervalo de 0 a 15 extrações.
- 3) A metodologia “bootstrap” apresentou resultados similares aos da aproximação linear, para todos os extratores.
- 4) Todos os estimadores dos parâmetros apresentaram viés “bootstrap” para algum extrator. No entanto, o viés foi inferior a 1%, para todos eles, o que é considerado baixo, segundo a literatura.
- 5) Todos os parâmetros apresentaram distribuição não normal, considerando os testes assintóticos e a metodologia “bootstrap”. No entanto, a visualização dos histogramas mostra que essa falta de normalidade foi pequena.
- 6) Os extratores apresentaram valores baixos para as medidas de curvatura de Bates e Watts, tornando, assim, válidas as inferências realizadas por meio da aproximação linear, com exceção do extrator NH_4OAc , que apresentou medida de curvatura de efeito parâmetro alta, tornando os resultados inferenciais para este extrator questionáveis.
- 7) Estudos futuros devem ser realizados para o extrator NH_4OAc , com o objetivo de avaliar se ele segue realmente o modelo de cinética de primeira ordem ou se é necessária uma reparametrização apropriada para tal modelo.

- 8) Os extratores ácido acético, láctico, MOAS e SOAS apresentaram resultados inferenciais semelhantes. Assim, em estudos futuros, o melhor extrator será aquele que mais se assemelhar aos resultados obtidos pela planta.
- 9) Os extratores de DTPA e Mehlich I absorvem os elementos-traço quase que nas primeiras extrações e, então, não devem ser indicados para representar a planta, tendo em vista que a mesma absorve os elementos-traço ao longo do tempo.
- 10) A metodologia da aproximação linear, a técnica “bootstrap” e as medidas de curvatura de Bates e Watts tornaram os resultados inferenciais de Marchi (2005) válidos e com previsões eficientes para todos os extratores, exceto para o extrator NH_4OAc .

REFERÊNCIAS BIBLIOGRÁFICAS

ANDERSON, A. Relative efficiency of nine different soil extractants. **Swedish Journal of Agricultural Research**, Stockholm, v. 5, n. 2, p. 125-135, 1975.

BATES, D. M.; WATTS, D. G. **Nonlinear regression analysis and its applications**. New York: Wiley series in probability e mathematical statistics, 1988. 365p.

BATES, D. M.; WATTS, D. G. Relative curvature measures of nonlinearity (with discussion). **Journal of the Royal Statistical Society**, Ser. B, v. 42, n. 1, p. 1-25, 1980.

BEALE, E. M. L. Confidence regions in nonlinear estimation (with discussion). **Journal of the Royal Statistical Society**, London, v. 22, p. 41-88, 1960.

BERMOND, A.; GHESTEM, J. P. Kinetic study of trace metal EDTA desorption from contaminated soils. In: SELIM, H. M.; SPARKS, D. L. **Heavy metals release in soils**. New York: Lewis Publishers, 2001. p. 131-148.

BOX, G. E. P.; LUCAS, H. L. Design of experiments in nonlinear situations. **Biometrika**, Oxford, n. 46, p. 77-90, 1959.

CHARNET, R.; BONVINO, H.; FREIRE, C. A. L.; CHARNET, E. M. R. **Análise de modelos de regressão linear com aplicação**. Campinas: Unicamp, 1999. 354 p.

CURRIE, D. J. Estimating Michaelis-Menten parameters: bias, variance e o experimental design. **Biometrics**, Oxford, v. 38, n. 4, p. 907-919, 1982.

DRAPER, N. R.; SMITH, H. **Applied regression analysis**. 3. ed. New York: J. Wiley, 1998. 706 p.

EVANS, L. J.; SPIERS, G. A.; ZHAO, G. Chemical aspects of heavy metal solubility with reference to sewage sludge amended soils. **International Journal of Environmental Analytical Chemistry**, London, v. 59, n. 2/4, p. 291-302, 1995.

GALLANT, A. R. **Nonlinear statistical models**. New York: Wiley series in probability e mathematical statistics, 1987. 624 p.

GUJARATI, D. **Econometria básica**. 4. ed. São Paulo: Campus, 2006. 860 p.

GUSTAFFSON, J. P. **Visual Minteq, ver.2.50, Kungliga Tekniska högskolan**. Stockolm: Royal Institute of Technology. Department of Land and Water Resources Engineering, 2006.

GUTTMAN, I.; MEETER, D. A. On Beale's measure of nonlinearity. **Technometrics**, Washington, v. 7, n. 4, p. 623-637, 1965.

INTERNATIONAL UNION OF PURE AND APPLIED CHEMISTRY. **Compendium of chemical terminology, electronic version**. 1979. Disponível em: <<http://goldbook.iupac.org/T06421.html>>. Acesso em: 04 abr. 2008.

JORGENSEN, B. **The theory of linear models**. London: Champman & Hall, 1993. 226 p.

KABATA-PENDIAS, A.; PENDIAS, H. **Trace elements in soils and plants**. Florida: CRC, 2001. 413 p.

LIPSCHUTZ, M. M. **Schaum's outline of theory and problems of differential geometry**. New York: McGraw-Hill, 1969. 269 p.

MACHADO, E. J. **Uso do bootstrap na estimação de parâmetros em modelos não-lineares**: uma aplicação em mecanismos cinéticos de Michaelis-Menten. 2006. 131p. Dissertação (Mestrado em Estatística e Experimentação Agropecuária) – Universidade Federal de Lavras, MG.

MARCHI, G. **Organic acids and trace elements extractability in sewage sludge-treated soils**. 2005. 110 p. Tese (Doutorado em Agronomia/ Solos e Nutrição de Plantas) – Universidade Federal de Lavras, MG.

MAZUCHELI, J.; ACHCAR, J.A. Algumas considerações importantes em regressão não linear. **Acta Scientiarum**, Maringá, v. 24, n. 6, p. 1761-1770, 2002.

MOOD, A. M.; GRAYBILL, F. A.; BOES, D. C. **Introduction to the theory of statistics**. 3. ed. Singaporé: McGraw-Hill Book Company, 1974. 562 p.

PEREIRA, J. M.; MUNIZ, J. A. M.; SILVA, C. A. Nonlinear models to predict nitrogen mineralization in an oxisol. **Scientia Agricola**, Piracicaba, v. 62, n. 4, p. 395-400, 2005.

PEREIRA, J. **Predição do nitrogênio mineralizado em latossolo pelo uso de modelos não-lineares**. 2004. 82 p. Dissertação (Mestrado em Agronomia/Estatística e Experimentação Agropecuária) – Universidade Federal de Lavras, MG.

RATKOWSKY, D. A. **Nonlinear regression modeling**: a unified practical approach. New York: Marcel Dekker, 1983. 276 p.

R DEVELOPMENT CORE TEAM. **R**: A language and environment for statistical computing; versão 2.7.0.v. Vienna: R Foudation for Statistical Computing. Disponível em: <<http://www.R-project.org>>. Acesso em: 23 fev. 2008.

REZENDE, D. M. L. C.; MUNIZ, J. A.; FERREIRA, D. F.; SILVA, F. F; AQUINO, L. H. Ajuste de modelo de platô de resposta para a exigência de zinco em frangos de corte. **Ciência e Agrotecnologia**, Lavras, v. 31, n. 2, p. 468- 478, 2007.

SEBER, G.A.F. **Linear regression analysis**. New York: Wiley series in probability e mathematical statistics, 1977. 465 p.

SNEDECOR, G. W.; COCHRAN, W. G. **Statistical methods**. 7. ed. Ames: Iowa State University, 1980. 507 p.

SOUZA, G. S. **Introdução aos modelos de regressão linear e não linear**. Brasília: EMBRAPA, SP/ SEA, 1998. 505 p.

UNITED STATES ENVIRONMENTAL PROTECTION AGENCY - USEPA 3052. **Microwave assisted acid digestion of silicious and organically based matrices**. 1996. 20 p. Disponível em: <<http://www.epa.gov/epaoswer/hazwaste/test/pdfs/3052>>. Pdf>. Acesso em: 02 jan. 2008.

ANEXOS

ANEXO	Pag.
ANEXO 1 Rotina R, programa utilizado para a obtenção das estimativas dos parâmetros do modelo não-linear via aproximação assintótica, banda de confiança para a resposta esperada e região de confiança conjunta.....	85
ANEXO 2 Rotina R, programa utilizado para a obtenção das estimativas dos parâmetros do modelo não-linear, utilizando a técnica “bootstrap”.....	89
ANEXO 3 Rotina R, programa utilizado para a obtenção das medidas de curvatura de Bates e Watts.....	92
ANEXO 4 Transformação QR	93

ANEXO 1

Rotina R, programa utilizado para a obtenção das estimativas dos parâmetros do modelo não-linear via aproximação assintótica, banda de confiança para a resposta esperada e região de confiança conjunta:

```
#Programa para realizar uma análise de regressão não linear
#Utilizando funções próprias do R e funções programadas de acordo com as
#fórmulas apresentadas na literatura
# x = Número de extrações
# y = Acético
# X matriz das derivadas parciais
#Entrando com os dados
x=c(1,1,1, ... ,15,15,15)
y=c(9.69,9.89,9.67, ... ,28.19,28.40,28.38)
# Encontrado os chutes iniciais conforme metodologia apresentada na
#dissertação
chute1=max(y)
chute2=mean(-log((chute1-y+0.05)/(chute1+0.05))/x)
vetor.chutes=c(chute1,chute2)
round(vetor.chutes,4) # resultado dos chutes iniciais
# Ajustando o modelo (utilizando a função nls do R)
funcao=y~theta1*(1-exp(-theta2*x))
exponencial=nls(funcao,start=c(theta1=chute1,theta2=chute2))
resul=summary(exponencial)
resul # resultado das estimativas
# Guardando os resultados separadamente
mat.res=as.matrix(resul$parameters) #vendo os resultados como uma matriz
```

```

theta1=mat.res[1,1] # estimativa para theta1
theta2=mat.res[2,1] # estimativa para theta2
DPtheta1=mat.res[1,2] # desvio padrão de theta1
DPtheta2=mat.res[2,2] # desvio padrão de theta 2
s=resul$sigma # erro padrão residual
s2=s^2 # estimativa da variância
theta=c(theta1,theta2) # vetor de estimativas
# valor ajustado e residuo (programando com as fórmulas)
yhat=(theta1*(1-exp(-theta2*x))) # valores preditos
residuo=y-yhat # residuo
# matriz das derivadas
x1=(1-exp(-theta2*x)) # colocar a derivada em relação a theta1
x2=(theta1*theta2*exp(-theta2*x)) # colocar a derivada em relação a theta2
X=cbind(x1,x2) # matriz das derivadas
N=nrow(X) # Número de linhas de X
P=ncol(X) # Número de colunas de X
# Encontrando o  $R^2$ 
SQR=s^2*(N-P)
SQR
# Encontrando a soma de quadrados do total corrigido
SQTcorrigida=sum((y-mean(y))^2)
# Encontrando o  $R^2$  (programando com as fórmulas)
R2 = 1 - SQR/SQTcorrigida
round(R2,4) # resultado para o  $R^2$ 
# Intervalos de confiança marginais (usando direto uma função R)
IC=confint(exponencial, level = 0.95)
round(IC,4)
# Matriz de Covariâncias (usando direto uma função R)

```

```

vcov(exponencial)
# Matriz de correlações (programando de acordo com as fórmulas)
R1=qr.R(qr(X), complete = FALSE) # matriz R1 da decomposição QR de X
F=solve(t(X)%*%X)
r1=as.real(F[1,1])
r2=as.real(F[2,2])
vetor=c(sqrt(r1),sqrt(r2))
Z=diag(vetor)
L=solve(Z)%*%solve(R1)
corr=L%*%t(L)
round(corr,4) # matriz de correlação entre os parâmetros
#Apresentação gráfica do ajuste, banda de confiança e região de Inferência #
alfa=0.05 # necessário entrar com alfa
Ftab=qf(1-alfa,P,N-P)
k=sqrt(P*s2*Ftab)
A=solve(R1)
a=as.real(A[1,1])
b=as.real(A[1,2])
c=as.real(A[2,1])
d=as.real(A[2,2])
# Banda de Confiança para a resposta esperada
# Gráficos dos ajustes e das bandas de confiança
# Aqui tem-se que modificar as funções de acordo com os dados
plot(x,y,xlab="extrações",ylab="mgZn/Kg de solo",ylim=c(0,30),xlim=c(0,15))
curve((theta1*(1-exp(-theta2*x))),0,max(x),add=T)
curve(((theta1*(1-exp(-theta2*x)))-sqrt((((1-exp(-theta2*x))*a+(theta1*x*exp(-
theta2*x))*c)^2+((1-exp(-theta2*x))*b+(theta1*x*exp(-theta2*x))*d)^2))*k, 0,
max(x),add=T,lty=2,col=2)

```

```

curve(((theta1*(1-exp(-theta2*x)))+sqrt((((1-exp(-theta2*x))*a+(theta1*x*exp(-
theta2*x))*c)^2+((1-exp(-theta2*x))*b+(theta1*x*exp(-theta2*x))*d)^2))*k,0,
max(x),add=T,lty=2,col=2)
# guardando os limites dos ICs separadamente
LItheta1=IC[1,1] # limite inferior de theta1
LStheta1=IC[1,2] # limite superior de theta1
LItheta2=IC[2,1] # limite inferior de theta2
LStheta2=IC[2,2] # limite superior de theta2
# Gráfico da região de confiança conjunta para o parâmetro
t=seq(-10000,10000,len=10000)
y1=theta1+k*a*cos(t)+k*b*sin(t)
y2=theta2+k*c*cos(t)+k*d*sin(t)
plot(y1,y2,xlab=expression(C[1]),ylab=expression(k),lwd=1,xlim=c(26,28),ylim
=c(0.2,0.52))
points(theta1,theta2,pch=3)
a1=c(LItheta1,LItheta1)
a2=c(0,LStheta2+1000)
lines(a1,a2,lty=2)
a3=c(LStheta1,LStheta1)
a4=c(0,LStheta2+1000)
lines(a3,a4,lty=2)
a5=c(0,LStheta1+1000)
a6=c(LItheta2,LItheta2)
lines(a5,a6,lty=2)
a7=c(0,LStheta1+1000)
a8=c(LStheta2,LStheta2)
lines(a7,a8,lty=2)

```

ANEXO 2

Rotina R, programa utilizado para a obtenção das estimativas dos parâmetros do modelo não-linear, utilizando a técnica “bootstrap”

Simulação: Bootstrap Paramétrico (5000 simulações)

```
zb = theta1; zb1 = c(zb); hb=theta2; hb1 = c(hb)
DPzb=DPtheta1; DPzb1=c(DPzb); DPhb=DPtheta2; DPhb1=c(DPhb)
options(show.error.messages = FALSE)
cont=1
exponencialsimu=0;contf=0
while (cont<=4999)
{
  e=rnorm(N,0,s2)
  ysimu=(theta1*(1-exp(-theta2*x)))+e
  funcaosimu=ysimu~the1*(1-exp(-the2*x))
  exponencialsimu=as.null(exponencialsimu)

  exponencialsimu=try(nls(funcaosimu,start=c(the1=chute1,the2=chute2),trace =
FALSE))

  if (!is.null(names(exponencialsimu)))
  {
    resulsimu=summary(exponencialsimu)
    mat.res=as.matrix(resulsimu$parameters)
    theta11=mat.res[1,1]
    theta22=mat.res[2,1]
    DPtheta11=mat.res[1,2]
```

```

DPtheta22=mat.res[2,2]
zb = theta11
hb=theta22
DPzb=DPtheta11
DPhb=DPtheta22
DPzb1=c(DPzb1,DPzb)
DPhb1=c(DPhb1,DPhb)
zb1 = c(zb1,zb) # vetor armazenando os theta1
hb1 = c(hb1,hb) # vetor armazenando os theta2
cont = cont + 1
} else contf=contf+1
}
options(show.error.messages = TRUE)
library(fBasics) # O pacote fBasics deve estar instalado (plotar histogramas)
#Histograma para as estimativas simuladas
histPlot(zb1,xlab=expression(C[t]),ylab="Frequência",skipZeros = FALSE)
histPlot(hb1,xlab=expression(k),ylab="Frequência",skipZeros = FALSE)
# Fazendo os estudos bootstraps para o parâmetro theta1
mean(zb1) # média dos valores estimados
var(zb1) # variância amostral
m2z=(sum((zb1-mean(zb1))^2))/length(zb1) #segundo momento amostral
centrado na média
m3z=(sum((zb1-mean(zb1))^3))/length(zb1) #Terceiro momento amostral
centrado na média
m4z=(sum((zb1-mean(zb1))^4))/length(zb1) #Quarto momento amostral
centrado na média
m2z;m3z;m4z # imprimindo os valores
g1z=m3z/(m2z^(3/2)) #coeficiente de assimetria

```

```

curtose=m4z/(m2z^2) #coeficiente de curtose
g2z=curtose-3 #excesso de curtose
round(g1z,4);round(g2z,4)
vies=mean(zb1)-theta1 #estimativa do vies
(vies/theta1)*100 # percentual do vies com relação ao parâmetro
g1z.tab=qnorm(0.95,0,(6/5000)^(1/2)) # Coef. assimetria tabelado
g2z.tab=qnorm(0.95,0,(24/5000)^(1/2)) # Excesso de curtose tabelado
round(g1z.tab,4);round(g2z.tab,4)
# Estudando a significância do viés para theta1
zi=sqrt(5000)*((mean(zb1)-theta1)/DPtheta1)
zi
vies.tab=qnorm(0.95,0,1)
vies.tab
# Intervalo de Confiança Bootstrap usando o método percentil t
qpiv=sqrt(N)*((zb1-theta1)/DPzb1) #quantidade necessária para encontrar t
qpiv.ord=sort(qpiv) # ornenando as quantidades
t1inf=qpiv.ord[4875] #quantil superior para o intervalo percentil bootstrap
t1sup=qpiv.ord[125] #quantil inferior para o intervalo percentil bootstrap
Linferior=mean(zb1)-t1inf*(sd(zb1)/sqrt(N))
Lsuperior=mean(zb1)-t1sup*(sd(zb1)/sqrt(N))
Linferior;Lsuperior
# Para theta2 realizar procedimento semelhante

```

ANEXO 3

Rotina R, programa utilizado para a obtenção das medidas de curvatura de Bates e Watts

#Estudando a curvatura de Bates e Watts

Ver capítulo 07 do livro - 1988

```
expcurve <- deriv3(~theta1*(1-exp(-theta2*x)), c("theta1", "theta2"),
```

```
function(theta1, theta2, x) NULL)
```

```
(ajuste.expon <- nls(y ~ expcurve(theta1, theta2, x),
```

```
start = list(theta1=chute1, theta2=chute2)))
```

```
rms.curv(ajuste.expon) # função que fornece os valores para medir o efeito da  
curvatura
```

ANEXO 4

Transformação QR

Para computar a decomposição **QR** de uma matriz **X** de ordem $N \times P$, podem-se usar transformações Householder (Householder, 1958 apud Bates & Watts, 1988), uma generalização de reflexos no plano. Estas são matrizes da forma:

$$\mathbf{H}_u = \mathbf{I} - 2\mathbf{u}\mathbf{u}^T$$

em que **I** é a matriz identidade $N \times N$ e **u** é um vetor de unidade N -dimensional (isto é, $\|\mathbf{u}\| = \sqrt{\mathbf{u}^T \mathbf{u}} = 1$). $\mathbf{H}_u$ é simétrica e ortogonal, já que:

$$\mathbf{H}_u^T = \mathbf{I}^T - 2\mathbf{u}\mathbf{u}^T = \mathbf{H}_u \quad \text{e} \quad \mathbf{H}_u^T \mathbf{H}_u = \mathbf{I} - 4\mathbf{u}\mathbf{u}^T + 4\mathbf{u}\mathbf{u}^T \mathbf{u}\mathbf{u}^T = \mathbf{I}$$

Multiplcando um vetor **y** por $\mathbf{H}_u$:

$$\mathbf{H}_u \mathbf{y} = \mathbf{y} - 2\mathbf{u}\mathbf{u}^T \mathbf{y}$$

corresponde a refletir **y** próximo da linha que passa pela origem perpendicular ao **u**, como mostrado na Figura A4.1. Escolhendo:

$$\mathbf{u} = \frac{\mathbf{y} - \|\mathbf{y}\| \mathbf{e}_1}{\|(\mathbf{y} - \|\mathbf{y}\| \mathbf{e}_1)\|} \quad (\text{E4.1})$$

ou

$$\mathbf{u} = \frac{\mathbf{y} + \|\mathbf{y}\| \mathbf{e}_1}{\|(\mathbf{y} + \|\mathbf{y}\| \mathbf{e}_1)\|} \quad (\text{E4.2})$$

em que $\mathbf{e}_1 = (1, 0, \dots, 0)^T$, a transformação Householder pode ser feita para levar o vetor **y** ao eixo *x*, isto é, no novo sistema coordenado (girado), o vetor $\mathbf{H}_u \mathbf{y}$ tem coordenadas $(\pm \|\mathbf{y}\|, 0, \dots, 0)^T$.

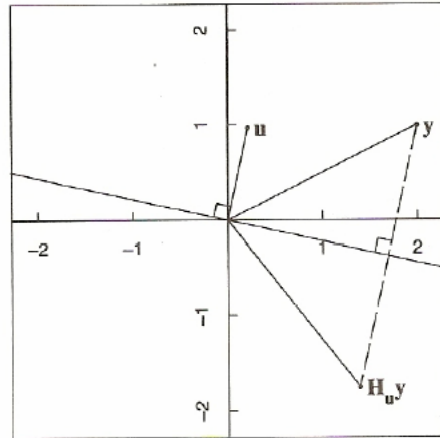


FIGURA A4.1 Uma transformação Householder mostrando a reflexão sobre a linha perpendicular de $\mathbf{u}$ para o vetor $\mathbf{y}$ para a forma $\mathbf{H}_u \mathbf{y}$.

Exemplo: Para realizar a decomposição QR da matriz

$$\mathbf{X} = \begin{bmatrix} 1 & 1,26 \\ 1 & 1,82 \\ 1 & 2,22 \end{bmatrix}$$

Escolhe-se uma transformação $\mathbf{H}_{u_1}$ para levar a primeira coluna $\mathbf{x}_1$ de $\mathbf{X}$ usando (E4.1) e obtém-se:

$$\mathbf{u}_1 = \frac{(1,1,1)^T - \sqrt{3}(1,0,0)^T}{1,5925} = \frac{(-0,7321; 1; 1)^T}{1,5925}$$

de modo que:

$$\mathbf{X}_1 = \mathbf{H}_{u_1} \mathbf{X} = \begin{bmatrix} 1,7321 & 3,0600 \\ 0 & -0,6388 \\ 0 & -0,2388 \end{bmatrix}$$

Agora realiza-se uma segunda rotação, ortogonal à primeira, escolhendo $\mathbf{u}_2$ de modo que $\mathbf{H}_{u_2}$ zere as linhas abaixo da diagonal da segunda coluna de $\mathbf{X}_1$

sem mudar a primeira coluna. Para assegurar que não se muda a primeira coluna, faz-se o primeiro elemento de $\mathbf{u}_2$ zero. O vetor $\mathbf{u}_2$ é assim escolhido por ser:

$$\mathbf{u}_2 = \frac{(0; -0,6388; -0,2388)^T - 0,6820(0,1,0)^T}{1,3422}$$

que resulta:

$$\mathbf{R} = \mathbf{H}_{u_2} \mathbf{H}_{u_1} \mathbf{X} = \begin{bmatrix} 1,7321 & 3,0600 \\ 0 & 0,6820 \\ 0 & 0 \end{bmatrix}$$

Como mostrado no exemplo acima, a matriz $\mathbf{R}$ é produzida ao mesmo tempo em que os vetores de reflexão Householder $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_p$ são obtidos. O $\mathbf{Q}$ correspondente é determinado a partir do $\mathbf{R} = \mathbf{Q}^T \mathbf{X}$, isto é, $\mathbf{Q}^T = \mathbf{H}_{u_p} \dots \mathbf{H}_{u_1}$. Pela simetria das transformações Householder:

$$\mathbf{Q} = \mathbf{H}_{u_1}^T \dots \mathbf{H}_{u_p}^T = \mathbf{H}_{u_1} \dots \mathbf{H}_{u_p}$$

Observe que o $\mathbf{Q}$ quase nunca é calculado explicitamente, quando quaisquer multiplicações por $\mathbf{Q}$ ou $\mathbf{Q}^T$ são realizadas aplicando-se transformações Householder na ordem adequada. Uma transformação Householder $\mathbf{H}_u$ é aplicada a um vetor $\mathbf{y}$ subtraindo-se $2\mathbf{u}\mathbf{u}^T \mathbf{y}$ do $\mathbf{y}$ e só requer calcular um produto interno e subtrair um múltiplo de um vetor de um outro.

Para regressão linear e não linear, às vezes, é conveniente realizar uma decomposição $\mathbf{QR}$ da matriz $\mathbf{X}$ aumentada com o vetor $\mathbf{y}$. A decomposição $\mathbf{QR}$ de $(\mathbf{X} | \mathbf{y})$ produz o vetor $\mathbf{w}_1$ diretamente nas P primeiras linhas da $(P+1)$ -ésima coluna de $\mathbf{R}$, da qual a estimativa de mínimos quadrados ou o aumento podem ser obtidos. A $(P+1)$ -ésima entrada dessa coluna resulta $\|\mathbf{w}_2\|$, que pode ser usado para teste da convergência em regressão não linear e, após a convergência, calcular o desvio padrão residual.