**Medidas Repetidas com
Covariáveis Dependentes do Tempo:
Um Exemplo com Dados Incompletos**

Mariana Cúri

Dissertação apresentada ao
Instituto de Matemática e Estatística
da Universidade de São Paulo
para a Obtenção do Grau de
Mestre em Estatística

Área de Concentração: Estatística
Orientador: **Prof. Dr. Julio da Motta Singer**

- São Paulo, junho de 1999 -

Medidas Repetidas com Covariáveis Dependentes do Tempo: Um Exemplo com Dados Incompletos

Este exemplar corresponde à redação final
devidamente corrigida e defendida por
Mariana Cúri e aprovada pela comissão
julgadora.

São Paulo, 25 de junho de 1999.

Banca examinadora:

- Prof. Dr. Júlio da Motta Singer
- Prof. Dr. Antônio Carlos Pedroso de Lima
- Profa. Dra. Luzia Aparecida Trinca

Agradecimentos

Lembro-me que foi assistindo às aulas da professora Lisbeth Cordani que eu comecei a pensar em fazer este curso de mestrado. Depois, nas aulas do professor Dalton de Andrade e do professor Julio Singer acabei concretizando de vez essa idéia. Não foram os únicos professores responsáveis, mas os principais. Ao Julio, com quem tive o prazer de trabalhar e conviver mais de perto, agradeço em especial, pela orientação, dedicação e conversas tão agradáveis de fim de tarde. E desejo que este seja apenas o primeiro dos trabalhos que ainda faremos juntos.

Ao Clovis, quero agradecer pelos dados cedidos e pela força dada no final, quer dizer, no começo.

Aos meus pais, meus irmãos, meus cunhados e minha sobrinha. Sem os quais não conseguiria encontrar a felicidade nem o equilíbrio que tanto me ajudam a reconhecer e alcançar minhas conquistas. Os sábios dizeres de minha mãe que me emocionam e me aguçam a curiosidade sobre tantas coisas, não só nessa vida. O apoio que recebi,

principalmente de meus pais e irmãos, em alguns momentos em que me senti como que anestesiada, sem dar muita importância para coisas que machucam e, conseqüentemente, para outras que alegram e engrandecem também.

Em especial, agradeço ao Fabio pela paciência, incentivo e também pelas n revisões e sugestões, que foram indispensáveis.

Aos amigos que me acompanham pelo caminho que venho percorrendo, passando pelo Bandeirantes, pelo IME e, agora também, pelo InCor. Desses, não posso deixar de citar seis pessoas em particular: Camila, Renata, Paulinho, Raquel, Carine e Rita.

Ao CEA pelo conjunto de dados aqui analisado.

E, finalmente, ao CNPq, pelo suporte financeiro.

*“Imperioso que compreendas que,
se te oferecem o rumo,
a ti pertence a caminhada.”*

Emmanuel

Resumo

Consideramos a análise de dados provenientes de um experimento com medidas repetidas incompletas e covariáveis dependentes do tempo. O experimento consistiu na avaliação da resposta a um estímulo elétrico aplicado em nervos de mexilhões, antes e após tratamento com água do mar com diferentes salinidades (condições de avaliação). O estudo envolveu a modelagem do efeito da salinidade sobre as respostas avaliadas após os tratamentos, tendo como covariáveis as respostas avaliadas antes dos tratamentos. O objetivo foi comparar os resultados obtidos sob modelos de regressão para dados não correlacionados com aqueles em que diferentes estruturas de covariância intraunidades amostrais foram levadas em conta. Discutimos tanto os aspectos associados à escolha da forma da relação entre a resposta e as diferentes salinidades quanto aqueles que focam a escolha da estrutura de covariância. Mostramos que, embora ambas as classes de modelos produzam resultados semelhantes, a última gera estimadores mais precisos.

Summary

We considered the analysis of data from an incomplete repeated measures experiment with time dependent covariates. The study involved the evaluation of the response to an electric stimulus applied in mussel nerves before and after treatment with sea water with different salinities. The study involved the modeling of the effect of the salinity on the response evaluated after the treatments, considering the response evaluated before the treatments as covariates. The objective was to compare the results obtained under regression models for uncorrelated data with those in which different within sample units covariance structures were considered. We discussed not only the choice of the form of the relation between the response and the different salinities, but also the choice of the covariance structure. We show that although both classes of models yield similar results, the latter produces more precise estimators.

Índice

	Pag
Capítulo 1. Introdução	01
1.1. Planejamento da coleta de dados	02
1.2. Modelos estatísticos para estudos com medidas repetidas	03
1.2.1. Estruturas para a matriz de covariâncias intraunidades amostrais Σ_i	04
1.3. Estimação dos parâmetros do modelo	08
1.3.1. Máxima verossimilhança (MV)	08
1.3.2. Máxima verossimilhança restrita (MVR)	10
1.4. Processos iterativos	11
1.5. Comparações de modelos	11
Capítulo 2. Descrição do experimento	14
Capítulo 3. Modelagem dos parâmetros de regressão sob independência dos erros	19
3.1. Análise descritiva da relação entre as variáveis resposta e independentes	19
3.2. Modelagem dos parâmetros de regressão considerando as observações independentes	24
3.2.1. Redução dos coeficientes das medidas controles	26
3.2.2. Redução dos coeficientes das amplitudes médias do PAC ajustadas pelo controle	31

	Pag
Capítulo 4. Modelagem simultânea dos parâmetros de regressão e de covariância	35
4.1. Estratégias de ajuste dos modelos	36
4.2. Análise descritiva da estrutura de covariância amostral	37
4.2.1. Análise descritiva baseada nas matrizes de covariância e correlação	37
4.2.2. Análise descritiva baseada no variograma	42
4.3. Ajuste da estrutura de Σ_{ij} para o modelo saturado nos parâmetros de regressão	44
4.4. Modelagem dos parâmetros de regressão sob a estrutura (4.1) para a matriz de covariância intraunidades amostrais	53
4.4.1. Redução dos coeficientes das medidas controles	53
4.4.2. Redução dos coeficientes das amplitudes médias do PAC ajustadas pelo controle	54
Capítulo 5. Conclusão	60
Apêndice. Aspectos Computacionais	65
Bibliografia	69

Capítulo 1

Introdução

O objetivo principal deste trabalho é estudar um conjunto de dados com medidas repetidas usando modelos de efeitos aleatórios como ferramenta de análise e considerando que a variável resposta é normalmente distribuída. Tais dados já foram analisados pelo Centro de Estatística Aplicada do Instituto de Matemática e Estatística da USP (Singer e Mauro, 1987), porém a metodologia utilizada supôs independência entre as observações obtidas na mesma unidade experimental. Aqui, deseja-se verificar se a incorporação de uma estrutura de covariância mais complexa pode trazer maior precisão à análise. Uma característica importante desse conjunto de dados é que ele é baseado num planejamento com medidas repetidas e envolve o ajuste de curvas de crescimento em que o fator que define as condições de avaliação não é o tempo.

Nas demais seções deste capítulo, está apresentado um resumo da metodologia de análise de dados com medidas repetidas com uma única variável resposta contínua com distribuição de probabilidade normal multivariada. Maiores detalhes sobre a análise desse e de outros tipos de resposta são encontrados em Lindsey (1993), Diggle, Liang and Zeger (1994)

ou Jones (1993), por exemplo. No capítulo 2, estão descritos o planejamento e os objetivos do experimento em questão, além de todo o processo de coleta de dados. No capítulo 3, apresenta-se, com finalidade de comparação, o ajuste de um modelo linear para esses dados sob a suposição de que todas as observações são independentes. No capítulo 4, incorpora-se, ao modelo discutido no capítulo 3, uma estrutura de covariância intraunidades amostrais através de modelos de efeitos aleatórios e autorregressivos. No capítulo 5, comparam-se os ajustes dos dois modelos.

1.1. Planejamento da coleta de dados

A pesquisa científica utiliza métodos estatísticos para entender melhor a variabilidade de certos fenômenos da natureza. Para que essas pesquisas sejam mais eficientes, convém fazer um planejamento, em que a forma de coleta de dados deve estar bem definida. Dois importantes tipos de planejamento para a coleta de dados são o transversal (*cross sectional*) e o de medidas repetidas. O primeiro caracteriza-se por sugerir uma única observação de uma ou mais variáveis respostas para cada unidade amostral. Já o planejamento de medidas repetidas envolve várias observações de uma ou mais variáveis respostas para uma mesma unidade amostral. As observações feitas na mesma unidade experimental são coletadas sob diferentes condições de avaliação, que podem ser definidas por diversas características: diferentes distâncias de uma certa origem, diferentes posições, diferentes concentrações de uma determinada substância, diferentes tempos de exposição a uma certa droga, etc. Vale ressaltar que as observações não precisam ser feitas necessariamente na mesma sequência das condições de avaliação em todas as unidades experimentais, sendo muito comum a aleatorização dessa sequência para cada unidade. Quando o planejamento não permite a aleatorização da ordem com que as observações são obtidas na mesma unidade amostral (por exemplo, quando a característica que determina as condições de avaliação é o tempo), ele é chamado longitudinal.

Planejamentos de medidas repetidas e longitudinais podem, ainda, ser classificados como balanceados ou desbalanceados em relação às condições de avaliação. Planejamento de medidas repetidas balanceados determinam que todas as unidades experimentais sejam avaliadas sob as mesmas condições, podendo ser completos ou incompletos, conforme todas

as observações tenham sido medidas ou não. Já em planejamentos de medidas repetidas desbalanceados, cada unidade experimental pode ser avaliada sob diferentes condições.

Esses estudos, com mais de uma observação da resposta em uma mesma unidade amostral, são interessantes por várias razões. Entre elas, destacam-se:

- o número de unidades experimentais necessárias para a realização do estudo é menor do que nos planejamentos transversais;
- permitem o estudo dos efeitos de variações individuais ao longo das condições de avaliação;
- comparações intraunidades amostrais podem aumentar a precisão da análise estatística, principalmente nos casos em que a variabilidade entre indivíduos é alta.

1.2. Modelos estatísticos para estudos com medidas repetidas

Uma vez coletados os dados, a análise estatística, da mesma maneira que nos estudos transversais, visa estudar a relação entre a variável dependente e uma ou mais variáveis independentes. A única diferença está no fato de que se espera que as medidas obtidas em uma mesma unidade amostral sejam correlacionadas e que apenas as medidas feitas em diferentes unidades sejam independentes¹. Conseqüentemente, os modelos estatísticos apropriados para tais situações são mais complexos. O modelo pode ser escrito na forma linear geral como:

$$y_i = X_i\beta + \varepsilon_i \quad (1.1)$$

$i=1, \dots, n$, em que n é o número de unidades amostrais, y_i é o vetor de respostas da i -ésima unidade amostral com dimensão $(t_i \times 1)$, t_i é o número de condições de avaliação da i -ésima unidade amostral, X_i é a matriz de especificação $(t_i \times p)$ do modelo, β é o vetor $(p \times 1)$ de parâmetros de regressão desconhecidos e ε_i é o vetor de erros não observáveis com dimensão $(t_i \times 1)$. Assumem-se ε_i independentes e com distribuição $N(0, \Sigma_i)$, sendo que Σ_i depende de um vetor de q parâmetros desconhecidos, θ , variando independentemente dos parâmetros β .

As etapas de modelagem são duas: (i) a modelagem dos valores médios da

¹ A suposição de distribuição normal multivariada para a variável resposta torna os conceitos de medidas dependentes e correlacionadas sinônimos. Assim, não é necessária a distinção entre observações dependentes e correlacionadas.

variável resposta e (ii) a modelagem da matriz de covariâncias intraunidades amostrais Σ_i como função do vetor ($q \times 1$) de parâmetros θ . Ainda que o interesse principal esteja nos parâmetros de regressão β , pode-se melhorar consideravelmente suas estimativas, isto é, diminuir seus erros padrão, modelando-se a matriz de covariâncias intraunidades amostrais.

1.2.1. Estruturas para a matriz de covariâncias intraunidades amostrais Σ_i

Das possíveis estruturas para a matriz de covariâncias intraunidades amostrais, algumas estão descritas abaixo.

- Não Estruturada

Todos os parâmetros de covariância são incluídos no modelo, ou seja, nenhuma estrutura é imposta nas covariâncias entre os diferentes pares de observações intraunidades amostrais. Particularmente para uma unidade experimental i que foi observada em todas as T condições de avaliação do estudo, a matriz pode ser representada por

$$\Sigma_i = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} & \cdots & \sigma_{1T} \\ \sigma_{12} & \sigma_2^2 & \sigma_{23} & \cdots & \sigma_{2T} \\ \sigma_{13} & \sigma_{23} & \sigma_3^2 & \cdots & \sigma_{3T} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sigma_{1T} & \sigma_{2T} & \sigma_{3T} & \cdots & \sigma_T^2 \end{pmatrix}.$$

Para experimentos com muitas observações e poucas unidades experimentais ou nos casos em que os dados são incompletos ou desbalanceados, a estruturação de Σ_i em termos de um número menor de parâmetros é, geralmente, mais conveniente. Entre as várias alternativas, pode-se mencionar:

- Correlação Exponencial

Considera variância constante e correlação decrescente com o aumento do intervalo de tempo entre as observações. A covariância entre duas medidas realizadas em uma mesma unidade amostral nos instantes z_{ij} e z_{ik} é dada por $\sigma^2 \exp(-\phi|z_{ij} - z_{ik}|)$, onde z_{ij} é a j -ésima

condição de avaliação da i -ésima unidade amostral ($j=1, \dots, t_i$). Para uma unidade experimental observada em T condições de avaliação, a matriz é dada por

$$\Sigma_i = \sigma^2 \begin{pmatrix} 1 & \rho^{|z_{i2}-z_{i1}|} & \rho^{|z_{i3}-z_{i1}|} & \dots & \rho^{|z_{iT}-z_{i1}|} \\ \rho^{|z_{i2}-z_{i1}|} & 1 & \rho^{|z_{i3}-z_{i2}|} & \dots & \rho^{|z_{iT}-z_{i2}|} \\ \rho^{|z_{i3}-z_{i1}|} & \rho^{|z_{i3}-z_{i2}|} & 1 & \dots & \rho^{|z_{iT}-z_{i3}|} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{|z_{iT}-z_{i1}|} & \rho^{|z_{iT}-z_{i2}|} & \rho^{|z_{iT}-z_{i3}|} & \dots & 1 \end{pmatrix},$$

em que $\rho = \exp(-\phi)$ é a correlação entre duas medidas de uma mesma unidade experimental distanciadas de uma unidade de tempo. A taxa de decréscimo da correlação é mais rápida quanto maior o valor de ϕ .

Um caso particular, para situações em que as condições de avaliação são equidistantes, é o modelo autorregressivo de ordem 1 (AR1). A covariância entre duas medidas observadas em z_{ij} e z_{ik} é $\sigma^2 \rho^{|j-k|}$ e, para uma unidade experimental avaliada em T momentos,

$$\Sigma_i = \sigma^2 \begin{pmatrix} 1 & \rho & \rho^2 & \dots & \rho^{|T-1|} \\ \rho & 1 & \rho & \dots & \rho^{|T-2|} \\ \rho^2 & \rho & 1 & \dots & \rho^{|T-3|} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{|T-1|} & \rho^{|T-2|} & \rho^{|T-3|} & \dots & 1 \end{pmatrix}.$$

Note que, independentemente da quantidade de momentos de avaliação e de unidades experimentais, consegue-se representar a matriz de covariância por apenas dois parâmetros, σ^2 e ρ .

- Médias Móveis

Outra alternativa para Σ_i é o modelo baseado em médias móveis de ordem 1 (MA1), onde as observações avaliadas em momentos não consecutivos têm correlação nula. Neste caso, a matriz é dada por

$$\Sigma_i = \begin{pmatrix} \sigma^2 & \rho & 0 & \dots & 0 \\ \rho & \sigma^2 & \rho & \dots & 0 \\ 0 & \rho & \sigma^2 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \rho & \sigma^2 \end{pmatrix}.$$

- Efeitos Aleatórios

Uma alternativa que merece destaque é aquela baseada em modelos de efeitos aleatórios. Sob modelos desse tipo, a forma da relação entre as variáveis exploratórias e a variável dependente é a mesma para todas as unidades experimentais, mas os parâmetros de regressão que expressam essa relação variam de uma unidade para outra. Para exemplificar, considere o estudo de Frimm *et al.* (1998), onde se avaliou o efeito da diminuição do *stress* sobre a função cardíaca sistólica em pacientes hipertensos com função cardíaca alterada. A função sistólica e o *stress* foram mensurados para cada paciente em um momento inicial denominado pré tratamento. Após o tratamento com uma droga para diminuir a pressão arterial, novas aferições da função sistólica e do *stress* foram obtidas. O gráfico dos perfis individuais desses dados está apresentado na Figura 1.1.

Figura 1.1. Perfis individuais da relação entre a função sistólica e o *stress* (Frimm *et al.*, 1998).

Considere o modelo que relaciona a função sistólica e o *stress* através de uma reta para cada paciente, mas com coeficientes linear e angular variando entre os pacientes. Neste contexto, os coeficientes da reta associada a cada paciente podem ser vistos como a soma de

um termo fixo (coeficiente populacional) e um termo aleatório (efeito de paciente). Expressando essa relação através do modelo linear geral (1.1), tem-se

$$y_{ij} = \beta_{0i} + \beta_{1i}X_{ij} + \varepsilon_{ij}, \quad (1.2)$$

$i=1, \dots, 11$, $j=1, 2$ (pois $t_i=2$ para qualquer i), y_{ij} e X_{ij} representam, respectivamente, a função sistólica e o *stress* do i -ésimo paciente na j -ésima avaliação (pré ou pós tratamento), β_{0i} e β_{1i} são os coeficientes (aleatórios) linear e angular da reta associados ao i -ésimo paciente e ε_{ij} é um erro não observável. Supondo que o vetor dos coeficientes aleatórios (β_{0i}, β_{1i}) tenha distribuição normal multivariada com vetor de médias (β_0, β_1) e matriz de covariância $\mathbf{G} = \begin{pmatrix} \sigma_0^2 & \sigma_{01} \\ \sigma_{01} & \sigma_1^2 \end{pmatrix}$, pode-se escrever (1.2) como:

$$\begin{aligned} y_{ij} &= (\beta_0 + \gamma_{0i}) + (\beta_1 + \gamma_{1i})X_{ij} + \varepsilon_{ij} = \\ &= \beta_0 + \beta_1 X_{ij} + (\gamma_{0i} + \gamma_{1i}X_{ij}) + \varepsilon_{ij} \end{aligned} \quad (1.3)$$

em que γ_{0i} e γ_{1i} são os coeficientes (ou efeitos) aleatórios associados ao i -ésimo paciente e β_0 e β_1 são os parâmetros populacionais da relação entre função sistólica e *stress*. Observe que a relação **populacional** entre função sistólica e *stress* não precisa, necessariamente, ser expressa por uma reta, podendo ser quadrática ou cúbica, por exemplo.² Assim, pode-se reescrever (1.3) de forma genérica como

$$\mathbf{y}_i = \mathbf{X}_i \boldsymbol{\beta} + \mathbf{Z}_i \boldsymbol{\gamma}_i + \boldsymbol{\varepsilon}_i, \quad (1.4)$$

para $\mathbf{y}_i$, $\mathbf{X}_i$ e $\boldsymbol{\beta}$ definidos em (1.1), e em que $\mathbf{Z}_i$ é uma matriz de especificação dos efeitos aleatórios com dimensão $(t_i \times q)$, $\boldsymbol{\gamma}_i$ é um vetor $(q \times 1)$ de efeitos aleatórios e $\boldsymbol{\varepsilon}_i$ é um vetor de dimensão $(t_i \times 1)$ de erros aleatórios não observáveis.

Assumem-se $\boldsymbol{\gamma}_i$ e $\boldsymbol{\varepsilon}_i$ independentes e com distribuições $N(\mathbf{0}, \mathbf{G})$ e $N(\mathbf{0}, \mathbf{R}_i)$, respectivamente. Desta maneira, os vetores $\mathbf{y}_i$ são normalmente distribuídos com média $\mathbf{X}_i \boldsymbol{\beta}$ e matriz de covariância

$$\boldsymbol{\Sigma}_i = \mathbf{Z}_i \mathbf{G} \mathbf{Z}_i^t + \mathbf{R}_i, \quad (1.5)$$

com $\mathbf{G}$ e $\mathbf{R}_i$ podendo assumir qualquer estrutura.

² Neste exemplo, em particular, a relação populacional pode ser expressa por um polinômio com grau maior que um, mas a relação individual só pode ser representada por uma reta, pois só há duas observações para cada paciente.

Quando $\mathbf{R}_i = \sigma^2 \mathbf{I}_{t_i}$, onde $\mathbf{I}_{t_i}$ é a matriz identidade de dimensão t_i , (1.4)-(1.5) é chamado de modelo de efeitos aleatórios com independência condicional. Se além disso, $\mathbf{Z}_i = \mathbf{1}_{t_i}$, sendo $\mathbf{1}_{t_i}$ um vetor de dimensão t_i com todos os elementos iguais a 1, o modelo é vulgarmente rotulado de uniforme; sob tal modelo, todas as variâncias são iguais e todas as covariâncias intraunidades experimentais são constantes.

Mais detalhes sobre as estruturas de Σ_i podem ser encontrados em Singer and Andrade (1999), Diggle, Liang and Zeger (1994), Box and Jenkins (1976) e Laird and Ware (1982).

1.3. Estimação dos parâmetros do modelo

Os métodos mais usados de estimação dos parâmetros dos modelos lineares tratados aqui são o método de máxima verossimilhança (MV) e o método de máxima verossimilhança restrita (MVR). É importante ressaltar que, para que as análises descritas nas próximas seções sejam válidas nos casos em que o conjunto de dados é incompleto, deve-se assumir que o mecanismo gerador das observações perdidas é ignorável (Rubin, 1976).

1.3.1. Máxima verossimilhança (MV)

O núcleo da função log-verossimilhança do modelo (1.4) é dada por:

$$l(\beta, \theta) = \sum_{i=1}^n \left[\ln |\Sigma_i(\theta)| + (\mathbf{y}_i - \mathbf{X}_i \beta)' \Sigma_i^{-1}(\theta) (\mathbf{y}_i - \mathbf{X}_i \beta) \right] \quad (1.6)$$

O método de MV consiste em maximizar (1.6) em relação a β e θ . Esse valor máximo pode ser determinado através do sistema de equações gerado quando a primeira derivada de $l(\beta, \theta)$ com respeito a cada um dos parâmetros desconhecidos é igualada a zero, ou seja, os estimadores de máxima verossimilhança $\hat{\beta}$ e $\hat{\theta}$ de β e θ , respectivamente, são obtidos através das equações

$$\begin{cases} \hat{\beta} = \left(\sum_{i=1}^n \mathbf{X}_i' \Sigma_i^{-1}(\hat{\theta}) \mathbf{X}_i \right)^{-1} \left(\sum_{i=1}^n \mathbf{X}_i' \Sigma_i^{-1}(\hat{\theta}) \mathbf{y}_i \right) \\ \sum_{i=1}^n \left[tr(\Sigma_i^{-1}(\hat{\theta}) \dot{\Sigma}_i(\hat{\theta})) + \left[\frac{\partial \left((\mathbf{y}_i - \mathbf{X}_i \hat{\beta})' \Sigma_i^{-1}(\theta) (\mathbf{y}_i - \mathbf{X}_i \hat{\beta}) \right)}{\partial \theta_k} \right] \right]_{\theta=\hat{\theta}} = 0, \quad k = 1, \dots, q. \end{cases} \quad (1.7)$$

$$\text{em que } \dot{\Sigma}_i(\hat{\theta}) = \left[\frac{\partial \Sigma_i(\theta)}{\partial \theta_k} \right] \bigg|_{\theta=\hat{\theta}}.$$

Supondo $\Sigma_i(\theta)$ conhecida, $\hat{\beta}$ corresponde ao estimador de mínimos quadrados generalizados de β , digamos $\tilde{\beta}$, gerado através da minimização da distância de Mahalanobis

$$\sum_{i=1}^n (\mathbf{y}_i - \mathbf{X}_i \beta)' \Sigma_i^{-1}(\theta) (\mathbf{y}_i - \mathbf{X}_i \beta) \quad (1.8)$$

Além disso, $\tilde{\beta}$ tem distribuição normal com média β e variância $\left(\sum_{i=1}^n \mathbf{X}_i' \Sigma_i^{-1}(\theta) \mathbf{X}_i \right)^{-1}$. Como $\hat{\theta}$ é um estimador consistente de θ , sob algumas condições de regularidade e para um tamanho de amostra suficientemente grande, pode-se demonstrar que $\hat{\beta}$ segue a mesma distribuição de $\tilde{\beta}$ (Sen and Singer, 1993).

A matriz de covariância assintótica de $(\hat{\beta}, \hat{\theta})$ pode ser obtida através da inversa da matriz de informação de Fisher associada à função de log-verossimilhança, que é igual à esperança da matriz hessiana $\mathbf{H}$ com sinal negativo (Jennrich and Schluchter, 1986). Especificamente,

$$\begin{aligned} \begin{bmatrix} \text{Var}(\hat{\beta}) & \text{Cov}(\hat{\beta}, \hat{\theta}) \\ \text{Cov}(\hat{\beta}, \hat{\theta}) & \text{Var}(\hat{\theta}) \end{bmatrix} &= \mathbf{E}(-\mathbf{H})^{-1} = -\mathbf{E} \begin{bmatrix} \frac{\partial^2 l(\beta, \theta)}{\partial \beta \partial \beta} & \frac{\partial^2 l(\beta, \theta)}{\partial \beta \partial \theta} \\ \frac{\partial^2 l(\beta, \theta)}{\partial \beta \partial \theta} & \frac{\partial^2 l(\beta, \theta)}{\partial \theta \partial \theta} \end{bmatrix}^{-1} = \\ &= \begin{bmatrix} \left(\sum_{i=1}^n \mathbf{X}_i' \Sigma_i^{-1}(\theta) \mathbf{X}_i \right)^{-1} & \mathbf{0} \\ \mathbf{0} & (-\mathbf{H}_\theta)^{-1} \end{bmatrix}, \end{aligned} \quad (1.9)$$

$$\text{onde } -[\mathbf{H}_\theta]_{ks} = \frac{1}{2} \sum_{i=1}^n \text{tr}(\Sigma_i^{-1}(\theta)) \dot{\Sigma}_{ik}(\theta) \Sigma_i^{-1}(\theta) [2(\mathbf{y}_i - \mathbf{X}_i \beta)(\mathbf{y}_i - \mathbf{X}_i \beta)' - \Sigma_i(\theta)] \Sigma_i^{-1}(\theta) \dot{\Sigma}_{is}(\theta) \\ - \frac{1}{2} \sum_{i=1}^n \text{tr}(\Sigma_i^{-1}(\theta)) [(\mathbf{y}_i - \mathbf{X}_i \beta)(\mathbf{y}_i - \mathbf{X}_i \beta)' - \Sigma_i(\theta)] \Sigma_i^{-1}(\theta) \ddot{\Sigma}_{i,ks}(\theta)$$

$$k, s=1, \dots, q, \dot{\Sigma}_{ik}(\theta) = \frac{\partial \Sigma_i(\theta)}{\partial \theta_k} \text{ e } \ddot{\Sigma}_{i,ks}(\theta) = \frac{\partial^2 \Sigma_i(\theta)}{\partial \theta_k \partial \theta_s}. \text{ Pode-se mostrar que } \text{Cov}(\hat{\beta}, \hat{\theta}) = \mathbf{0}.$$

Os estimadores das variâncias dadas em (1.9) podem ser obtidos substituindo-se β e θ , respectivamente, por $\hat{\beta}$ e $\hat{\theta}$.

O estimador de MV de θ não leva em conta a perda de graus de liberdade devido à estimação de β e é viesado. Por esse motivo, um outro método, conhecido como máxima verossimilhança restrita (Laird and Ware, 1982), pode ser mais adequado para a estimação de θ .

1.3.2. Máxima verossimilhança restrita (MVR)

No caso do modelo linear geral, o estimador de máxima verossimilhança restrita é o estimador de máxima verossimilhança baseado em uma transformação linear do conjunto de dados, $y^* = Ay$, definida de maneira que a distribuição de y^* não dependa de β . A matriz de transformação A não é única, mas pode ser definida como a matriz que transforma y nos resíduos de mínimos quadrados ordinários, ou seja, $A = I - X(X'X)^{-1}X'$.

O núcleo da função log-verossimilhança restrita é

$$\begin{aligned} l_r(\theta) &= \sum_{i=1}^n \left[\ln |\Sigma_i(\theta)| + (y_i - X_i \tilde{\beta})' \Sigma_i^{-1}(\theta) (y_i - X_i \tilde{\beta}) \right] + \ln \left| \sum_{i=1}^n X_i' \Sigma_i^{-1}(\theta) X_i \right| = \\ &= l(\tilde{\beta}, \theta) + \ln \left| \sum_{i=1}^n X_i' \Sigma_i^{-1}(\theta) X_i \right| \end{aligned} \quad (1.10)$$

Maximizando-se (1.10) com respeito a θ , obtém-se o estimador de máxima verossimilhança restrita $\hat{\theta}_r$ através da solução da equação

$$\begin{aligned} \sum_{i=1}^n \left[\text{tr}(\Sigma_i^{-1}(\hat{\theta}_r) \dot{\Sigma}_i(\hat{\theta}_r)) + \left[\frac{\partial ((y_i - X_i \tilde{\beta})' \Sigma_i^{-1}(\theta_r) (y_i - X_i \tilde{\beta}))}{\partial \theta_r} \right] \right]_{\theta_r = \hat{\theta}_r} + \\ + \sum_{i=1}^n \left[\text{tr}(\Sigma_i^{-1}(\hat{\theta}_r) X_i' \Sigma_i^{-1}(\hat{\theta}_r) \dot{\Sigma}_i(\hat{\theta}_r) \Sigma_i^{-1}(\hat{\theta}_r) X_i) \right] = 0, \end{aligned} \quad (1.11)$$

com $\dot{\Sigma}_i(\hat{\theta}_r)$ definido em (1.7).

O estimador de máxima verossimilhança restrita de β é obtido de (1.7) com $\hat{\theta}$ substituído por $\hat{\theta}_r$.

Como descrito no item 1.3.1, pode-se obter a matriz de covariâncias de $\hat{\theta}_r$ através da matriz de informação de Fisher associada à função log-verossimilhança restrita.

Maiores detalhes sobre o assunto podem ser encontrados em Jennrich and Schluchter (1986), Wolfinger, Tobias and Sall (1994) e Singer and Andrade (1999).

1.4. Processos iterativos

Existem situações em que as equações dadas em (1.7) e (1.11) têm solução explícita, como nos casos de dados completos e balanceados em que $\Sigma_i(\theta)$ tem estrutura uniforme ou não estruturada. Porém, em geral, essas equações não têm solução explícita, implicando o uso de processos iterativos para a sua resolução. Dentre os mais conhecidos, encontram-se os algoritmos de Newton-Raphson, o *Scoring* de Fisher e o EM. Detalhes sobre estes três processos iterativos são encontrados em Jennrich and Schluchter (1986).

Basicamente, todos os processos iterativos seguem as etapas descritas abaixo. Suponha que η seja o vetor de parâmetros a ser estimado.

1. Uma estimativa inicial $\eta^{(0)}$ deve ser fornecida para se começar o processo iterativo;
2. Em cada iteração, será obtida uma nova estimativa para η que é calculada com base na estimativa da iteração anterior através da equação: $\eta^{(m+1)} = f(\eta^{(m)})$, onde f é determinada pelo algoritmo utilizado e (m) denota o número da iteração.
3. O processo iterativo é encerrado quando alguma função da diferença entre $\eta^{(m+1)}$ e $\eta^{(m)}$ satisfaz um critério de parada. Por exemplo, quando $\|\eta^{(m+1)} - \eta^{(m)}\| < \delta$, para $\delta > 0$ fixado.

1.5. Comparações de modelos

A modelagem estatística envolve testes de hipóteses sobre os parâmetros de regressão (β) e de covariância (θ). Esses testes caracterizam-se pela comparação de dois modelos com diferentes estruturas para os parâmetros de regressão e mesma estrutura para os parâmetros de covariância, ou vice-versa. São 3 os critérios de comparação de modelos citados neste trabalho: o teste de Wald, o teste da razão de verossimilhanças (TRV) e aquele baseado na estatística de Akaike. Mais informações sobre esses e demais critérios são encontradas em Diggle, Liang and Zeger (1994) e Jones (1993), entre outros.

3 critério de comparação de modelos

1) O critério de Akaike pode ser aplicado para a comparação de quaisquer modelos, exceto aqueles que diferem nos parâmetros de regressão e cujo método de estimação utilizado é o de máxima verossimilhança restrita. Esse critério calcula, para cada um dos dois modelos envolvidos, uma estatística que consiste na diferença entre o logaritmo da verossimilhança estimada e o número de parâmetros de covariância do modelo. O modelo com maior valor dessa estatística pode ser considerado mais adequado.

2) O teste de Wald serve para a comparação de dois modelos onde os parâmetros do primeiro podem ser escritos como uma combinação linear dos parâmetros do segundo. Modelos desse tipo serão chamados, em particular neste trabalho, como combinados. A hipótese de interesse pode, então, ser expressa como $C\beta=0$, com C uma matriz ($c \times p$) de constantes conhecidas. A estatística do teste tem distribuição assintótica qui-quadrado com c graus de liberdade e é dada por:

$$(C\hat{\beta})' (C \text{var}(\hat{\beta})^{-1} C') (C\hat{\beta}).$$

3) O TRV é baseado na estatística dada por $-2\ln(L_0/L_1)$, em que L_0 e L_1 denotam as verossimilhanças associadas aos modelos das hipóteses nula e alternativa, respectivamente. Este teste serve para a comparação de modelos encaixados, isto é, em que o vetor de parâmetros de um deles for um subconjunto do vetor de parâmetros do outro. Observe que este é um caso particular dos modelos combinados, onde os coeficientes da combinação linear são iguais a 0 ou 1. Para a comparação de dois modelos encaixados com relação aos parâmetros de regressão (e iguais relativamente aos parâmetros de covariância), o TRV só é válido quando o método de estimação adotado for o de máxima verossimilhança.³ Sob algumas condições de regularidade, a estatística do teste tem distribuição assintótica qui-quadrado com o número de graus de liberdade dado pela diferença entre o número de parâmetros dos modelos das hipóteses alternativa e nula. Por outro lado, quando os modelos forem encaixados com relação aos parâmetros de covariância (e equivalentes nos de regressão), o TRV pode ser utilizado para a comparação independentemente do método de estimação utilizado. Porém, como os estimadores dos parâmetros de covariância pertencem a um espaço paramétrico restrito (pois devem ser maiores ou iguais a zero), em geral, o TRV não satisfaz as condições usuais de regularidade. Assim, a distribuição de sua estatística é diferente da citada anteriormente. Em particular, para o teste correspondente à hipótese de que um único parâmetro de covariância é

³ Isto porque o logaritmo da verossimilhança restrita não depende dos parâmetros de regressão, inviabilizando o uso do TRV para essa comparação.

nulo, ela será dada por uma combinação 50:50 de duas distribuições qui-quadrado: uma com número de graus de liberdade igual ao número de parâmetros de covariância da hipótese alternativa e outra com o número de graus de liberdade igual ao número de parâmetros de covariância da hipótese nula. Mais detalhes são encontrados em Stram and Lee (1994).

Um esquema com as diferentes estratégias de comparação entre modelos encontra-se na Figura 1.2.

Figura 1.2. Esquema com os métodos estatísticos apropriados para a comparação de dois modelos.

Modelos	Akaike*	Wald	TRV*
Encaixados	X	X	X
Combinados	X	X	
Outros	X		

* Para comparação de modelos diferentes nos parâmetros de regressão, só é válido quando o método de estimação for o de máxima verossimilhança.

Capítulo 2

Descrição do Experimento

No conjunto de dados analisado neste trabalho, as unidades experimentais são mexilhões (moluscos bivalves) coletados no litoral de São Sebastião. Após a coleta, eles foram transportados em caixas de isopor para São Paulo e colocados em tanques equipados com filtros biológicos contendo água do mar do próprio local da coleta (salinidade de 3.4%). Os moluscos permaneceram durante 3 ou 4 dias nesses tanques e, a seguir, alguns foram transportados para dois outros tanques com salinidade de 2.4% e 1.9%, respectivamente. Os animais transferidos permaneceram nestes tanques por aproximadamente duas semanas para aclimação. Assim, a amostra subdividiu-se em três grupos diferenciados pela salinidade de aclimação (3.4%, 2.4% e 1.9%).

Antes de serem utilizados, os moluscos foram submetidos a um teste de reação de fechamento de valvas. Para isso foi aplicada uma determinada pressão nas valvas e o animal era considerado em bom estado se elas se fechassem e permanecessem assim; em caso contrário, ele era desprezado. O nervo conectivo cérebro-visceral foi extraído de cada animal e colocado

numa câmara especial para a estimulação elétrica, que consiste na aplicação de uma diferença de potencial entre o seu nervo e o meio onde está inserido. Como resposta ao estímulo elétrico, mediram-se, em diversas condições (que serão descritas mais adiante), três características de interesse: amplitude (mV), duração (min) e velocidade (m/s) do potencial de ação composto (PAC). Se a amplitude do PAC obtida estivesse abaixo de 6 mV, o animal era desprezado. Este valor foi considerado o limite inferior de indicação do *stress* do animal devido ao processo necessário para se realizarem as medições. Ao final do estudo, os grupos com salinidade de aclimação 1.9%, 2.4% e 3.4% eram constituídos de 18, 18 e 20 mexilhões, respectivamente.

As avaliações feitas em cada molusco seguiram as seguintes etapas, representadas esquematicamente na Figura 2.1:

- (i) Realizou-se o teste de fechamento de valvas;
- (ii) Mediram-se as características de interesse no final do período de aclimação;
- (iii) Submeteu-se o animal a uma salinidade (chamada, daqui em diante, de salinidade teste) diferente de sua salinidade de aclimação por um intervalo de dois minutos e, ao final deste tempo, observaram-se, novamente, as características de interesse (variáveis respostas);
- (iv) O molusco foi recolocado no tanque com a salinidade original (de aclimação) e, após três minutos, fizeram-se as avaliações das três características sob essas condições (consideradas medidas controles associadas à última salinidade teste à qual o molusco foi submetido);
- (v) Realizou-se o teste de fechamento de valvas.

As salinidades testes, aleatorizadas para cada animal, eram quatro das seguintes: 0.7%, 1.4%, 1.9%, 2.4% e 3.4%. Os itens (iii), (iv) e (v) eram repetidos até que se concluíssem todos os quatro testes ou até que os moluscos fossem desprezados por causa do *stress*.

Figura 2.1. Esquema de avaliação dos moluscos para cada grupo.

Esse procedimento foi repetido em dois tipos de meio: hipo e iso osmótico. Para o meio iso osmótico, as salinidades testes foram apenas aquelas menores que a salinidade de aclimação. Todas as observações foram coletadas pela mesma pessoa sob uma temperatura de 20°C.

Tendo em vista a descrição acima, pode-se classificar este estudo como tendo um planejamento com medidas repetidas com um fator intraunidades amostrais quantitativo. O interesse do estudo é comparar o comportamento das respostas eletrofisiológicas dos mexilhões não estressados dos três grupos, em cada um dos meios sob as diferentes salinidades testes. Maiores informações sobre os fenômenos investigados neste estudo podem ser encontrados em Stucchi-Zucchi (1988).

É importante destacar que, neste trabalho, os modelos estatísticos utilizados na análise dos dados serão comparados sem entrar em detalhes sobre as características biológicas do estudo. Assim, os capítulos consecutivos analisam apenas a amplitude do PAC medida no meio hipo osmótico. Os dados estão apresentados na Tabela 2.1.

Tabela 2.1. Amplitude do PAC (mV) no meio hipo osmótico.

a) Grupo com salinidade de aclimação 1.9%

Molusco	Salinidades Testes					Medidas Controles				
	0.7%	1.4%	1.9%	2.4%	3.4%	0.7%	1.4%	1.9%	2.4%	3.4%
20		5.0	8.4				6.4			
21	2.8	7.2	15.0	12.0	13.0	6.0	9.2		12.0	15.0
22	5.3	8.4	15.0	11.0	12.0	7.0	9.8		11.6	15.0
23	4.0	7.0	11.0	9.0	9.0	5.6	8.0		9.4	11.0
55	4.4		6.4	6.4	6.8	5.6			6.4	6.4
56		9.6	12.8	12.4	12.8		11.2		12.4	12.8
57	4.5		10.6	10.0	10.4	10.0			10.4	10.6
58	6.4	11.0	15.0	16.0	16.0	12.0	13.5		15.0	15.0
59	5.6	10.4	11.0	12.8	12.8	10.8	12.4		11.6	11.0
61	7.2	13.0	16.0	15.0	15.0	14.5	14.0		14.5	16.0
63	6.8	11.5	16.5	14.0	15.5	13.0	14.0		13.0	16.5
64	6.0	9.6	10.0	7.2	6.4	8.4	10.0		6.8	6.8
65	4.8	9.6	14.0	9.2	8.8	10.0	11.2		8.4	9.2
66	5.2	8.4	11.6	9.2	8.4	10.0	10.4		8.4	8.4
69			10.8	8.0	8.8				7.4	8.8
70		4.2	10.0				6.0			
81	3.2	5.8	7.8	8.0	6.4	7.0	7.0		7.8	7.5
82	4.0		7.6		6.0	5.2				7.6
média	5.01	8.62	11.64	10.68	10.51	8.94	10.22		10.34	11.10
d.p.	1.31	2.55	3.07	2.92	3.41	2.99	2.70		2.78	3.50

b) Grupo com salinidade de aclimação 2.4%

Molusco	Salinidades Testes					Medidas Controles				
	0.7%	1.4%	1.9%	2.4%	3.4%	0.7%	1.4%	1.9%	2.4%	3.4%
8	3.0			14.0	10.0	5.8				14.0
9	5.0			15.0	14.0	12.0				15.0
10	1.5			9.6	9.2	6.2				9.6
14	2.7	7.2	12.0	15.0	16.0	1.2	11.0	13.0		15.0
15			7.2	11.0	11.0			8.0		11.0
16	2.4	6.0	10.0	12.4	13.2	5.5	8.0	10.8		12.4
35		4.8	9.5	17.0			6.4	12.0		
36	3.2			9.2		6.4				
38	6.0	9.0	8.2	18.0	14.5	9.0	13.0	11.0		16.0
44	3.2	5.0	6.1	6.4	6.4	5.6	6.0	6.6		6.4
45			5.2	9.4	8.0			6.0		8.0
46	4.2	4.6	6.0	12.0	9.6	6.6	6.0	7.0		10.0
48	2.0	4.6	8.4	10.4		5.6	6.6	9.2		
49	7.2	8.8	10.6	13.0	13.0	10.0	10.8	11.2		13.0
52		8.0	8.8	11.2	8.4		11.2	9.6		8.4
53	4.8	8.8	10.2	12.0	10.0	10.8	12.4	12.4		9.6
89	3.2	5.6	10.0	12.0	11.8	6.0	9.2	11.2		12.0
90			4.8	14.0	12.0			6.0		12.0
média	3.72	6.58	8.36	12.31	11.14	6.98	9.15	9.57		11.49
d.p.	1.64	1.81	2.21	2.89	2.68	2.83	2.66	2.45		2.83

Tabela 2.1. (continuação).

c) Grupo com salinidade de aclimação 3.4%

Molusco	Salinidades Testes					Medidas Controles				
	0.7%	1.4%	1.9%	2.4%	3.4%	0.7%	1.4%	1.9%	2.4%	3.4%
4	4.4	9.2		11.0	13.0	6.6	10.0		11.0	
11	5.2			14.0	17.0	12.4			16.0	
12	3.0	6.6	9.6	12.0	13.2	8.4	10.0	12.8	13.2	
13			4.8	13.0	17.8			6.0	16.0	
17			8.6	15.0	18.0			12.0	17.0	
18	4.4	8.6	10.5	13.0	16.0	10.0	11.0	13.5	15.0	
19			9.6	10.8	15.0			10.6	13.6	
24			8.0	4.4	7.6			9.2	6.0	
25		5.6		11.0	13.0		7.0		12.0	
26	2.8	4.4	5.2	7.2	13.5	8.0	7.8	10.0	10.8	
27		5.0			18.0		6.6			
28		4.4	7.0		11.2		6.6	8.4		
29	4.6	6.6			9.8	8.0	9.8			
30		5.2	11.0	14.0	17.0		8.2	14.0	17.0	
31	6.5	8.5	12.5		17.0	14.0	14.0	16.0		
32	5.0	5.6	7.8		10.6	10.8	8.4	9.6		
33	8.5	8.0	8.4	9.2	14.0	13.0	11.0	11.0	11.0	
34	5.2	7.8	9.2	11.4	12.0	12.0	13.2	13.2	13.2	
72				7.6	16.0				11.6	
73	9.0	13.0	6.8	7.4	18.0	16.0	17.5	8.8	8.0	
média	5.33	7.04	8.50	10.73	14.39	10.84	10.08	11.08	12.76	
d.p.	1.98	2.36	2.14	3.02	3.08	2.93	3.13	2.66	3.20	

Capítulo 3

Modelagem dos Parâmetros de Regressão sob Independência dos Erros

Este capítulo irá concentrar-se no ajuste de um modelo para a relação entre a amplitude média do PAC e as salinidades testes, supondo todas as observações independentes. Primeiramente será feita uma análise descritiva para explorar qual a forma dessa relação. Em seguida, o modelo linear geral (1.1), com $\Sigma_{ij} = \sigma^2 \mathbf{I}_{t_{ij}}$, será ajustado aos dados, também com propósitos exploratórios.

3.1. Análise descritiva da relação entre as variáveis resposta e independentes

Visando conhecer o comportamento da amplitude média do PAC de acordo com as salinidades testes, construiu-se o gráfico de perfis médios apresentado na Figura 3.1. Pode-se

notar que, para cada grupo, a amplitude média do PAC parece decrescer levemente para salinidades testes maiores ou iguais à salinidade de aclimação, e decrescer acentuadamente para salinidades testes menores que a salinidade de aclimação. Por isso, resolveu-se estudar o comportamento da amplitude média do PAC em função das diferenças entre salinidades testes e de aclimação. Embora o objetivo seja ajustar uma curva de crescimento (em que a escala ordenada ao longo das quais se fazem as observações é a diferença de salinidade e não o tempo), o estudo não é longitudinal, dada a aleatorização da ordem de observação das respostas sob as diferentes salinidades testes.

Figura 3.1. Perfis médios da amplitude do PAC (mV) no meio hipo osmótico em função das salinidades teste.

O gráfico de perfis médios da amplitude do PAC em função das diferenças de salinidades está apresentado na Figura 3.2. Esse gráfico sugere que a amplitude do PAC pode ser descrita através de modelos de regressão segmentada (Draper and Smith, 1981), com mudança de coeficiente angular no ponto onde a diferença de salinidade é nula. Mais especificamente, podem-se ajustar duas semi-retas concorrentes para cada um dos grupos com salinidades de aclimação 1.9% e 2.4% (uma para diferenças de salinidade negativas e outra para diferenças de salinidade positivas) e uma semi-reta para o grupo com aclimação de 3.4%, pois para ele não foram consideradas diferenças positivas de salinidade.

Figura 3.2. Perfis médios da amplitude do PAC (mV) no meio hipo osmótico em função das diferenças de salinidades.

Para avaliar a necessidade de uso das medidas controles como covariáveis nos modelos sob investigação, construíram-se gráficos de dispersão da amplitude do PAC obtida em cada salinidade teste contra aquela do controle correspondente (Figura 3.3)¹. Esses gráficos sugerem a existência de uma possível associação positiva e linear entre estas duas variáveis. Observando os coeficientes angulares das retas ajustadas em cada salinidade (Figura 3.3), nota-se que sua inclinação parece ser maior para salinidades testes mais próximas das salinidades de aclimação. Desta forma, as medidas controles podem ser usadas no modelo como covariáveis que variam com as salinidades testes (Patel, 1986). Pelo fato de não se terem medidas controles da amplitude do PAC correspondentes às medidas realizadas sob as salinidades de aclimação, supôs-se, neste caso, que os valores controles eram iguais aos valores obtidos na salinidade de aclimação.

Um artifício muito usado nas análises de covariância consiste na inclusão de covariáveis centradas. Isto é bastante interessante pois implica a comparação das respostas médias dos tratamentos para um valor da covariável igual à sua média amostral. Quando a média da covariável não é a mesma para todos os tratamentos, pode-se centrá-la diferentemente em cada tratamento. Os valores médios das medidas controles dos grupos com salinidade de aclimação 1.9%, 2.4% e 3.4% são, respectivamente, iguais a 10.53, 10.13 e 12.09. Através de uma análise de variância rejeitou-se a igualdade desses valores ($p=0.0001$). Isto sugere que uma maneira de

¹ Essencialmente, as medidas controles correspondem a medidas de envelhecimento (*ageing*) dos moluscos e a idéia que fundamenta sua inclusão no modelo é uma tentativa de ajustar as conclusões a esse processo. As análises supõem que as respostas obtidas nessas situações não são afetadas pelos tratamentos.

incluir esta covariável no modelo seria centrando-a na média de cada grupo.

Figura 3.3. Gráficos de dispersão da amplitude do PAC (mV) em função da respectiva medida controle em cada salinidade teste no meio hipo osmótico. Os valores entre parênteses correspondem ao coeficiente angular e ao coeficiente de determinação da reta de regressão ajustada.

a) Grupo com salinidade de aclimação 1.9%

Figura 3.3. (continuação).

b) Grupo com salinidade de aclimação 2.4%

c) Grupo com salinidade de aclimação 3.4%

3.2. Modelagem dos parâmetros de regressão considerando as observações independentes

Inicialmente, considerou-se um modelo da forma (1.1) saturado em relação aos parâmetros de regressão com observações independentes e homocedásticas.

Sejam n_j o número de unidades amostrais do j -ésimo grupo ($n_1=20$ e $n_2=n_3=18$), $i=1, \dots, n_j$, $j=1,2,3$ e t_{ij} o número de salinidades testes a que o i -ésimo molusco do j -ésimo grupo foi submetido (no máximo igual a 5). Definindo $\mathbf{W}_{ij} = \mathbf{C}_{ij} - \bar{\mathbf{C}}_j \mathbf{1}_{t_{ij}}$, com $\mathbf{C}_{ij}$ denotando o vetor com dimensão $(t_{ij} \times 1)$ das medidas controles do i -ésimo molusco do j -ésimo grupo e $\bar{\mathbf{C}}_j = \sum_{i=1}^{n_j} \sum_{k=1}^{t_{ij}} \mathbf{C}_{ijk} / \sum_{i=1}^{n_j} t_{ij}$ a média amostral das medidas controles dos animais do j -ésimo grupo, um modelo saturado pode ser expresso como

$$\mathbf{y}_{ij} = \mathbf{X}_{ij} \boldsymbol{\beta}_j + \boldsymbol{\varepsilon}_{ij}, \quad (3.1)$$

em que $\mathbf{y}_{ij} = (y_{ij1} \ y_{ij2} \ \dots \ y_{ijt_{ij}})^t$ denota o vetor das respostas da amplitude do PAC do i -ésimo molusco do j -ésimo grupo, com dimensão $(t_{ij} \times 1)$,

$$\mathbf{X}_{ij} = \begin{pmatrix} \mathbf{I}_{t_{ij}} & \text{diag}(\mathbf{W}_{ij}^t) \end{pmatrix}$$

representa a matriz de especificação do modelo com dimensão $(t_{ij} \times 2 t_{ij})$,

$$\boldsymbol{\beta}_j = (\mu_{j1} \ \mu_{j2} \ \dots \ \mu_{jt_{ij}} \ \xi_{j1} \ \xi_{j2} \ \dots \ \xi_{jt_{ij}})^t$$

representa o vetor dos parâmetros de regressão com dimensão $(2t_{ij} \times 1)$, μ_{jk} é a amplitude do PAC média do j -ésimo grupo na k -ésima salinidade teste ajustada pela medida controle (isto é, para um valor da medida controle igual à média dessa medida no grupo considerado), ξ_{jk} é o coeficiente de ajuste da amplitude média pela medida controle do j -ésimo grupo na k -ésima salinidade teste (este coeficiente será chamado daqui em diante de coeficiente da medida controle) e $\boldsymbol{\varepsilon}_{ij} = (\varepsilon_{ij1} \ \varepsilon_{ij2} \ \dots \ \varepsilon_{ijt_{ij}})^t$ é um vetor de erros aleatórios para o qual se assume uma distribuição normal t_{ij} -variada com vetor de médias $\mathbf{0}$ e matriz de covariâncias $\Sigma_{ij} = \sigma^2 \mathbf{I}_{t_{ij}}$.

Para o ajuste do modelo foi utilizado o método de máxima verossimilhança restrita. O valor do logaritmo da verossimilhança restrita maximizada, $l_r(\hat{\sigma}^2)$, é -288.70. A estatística de

Akaike tem valor igual a -289.70. As estimativas dos parâmetros de regressão e de dispersão correspondentes estão apresentadas na Tabela 3.1.

Tabela 3.1. Estimativas de máxima verossimilhança restrita dos parâmetros do modelo saturado (3.1) e seus respectivos erros padrão.

Salinidade de Aclimação	Salinidade Teste	Amplitudes Médias Ajustadas pelas Medidas Controles			Coeficientes das Medidas Controles		
		Parâmetro	Estimativa	EP	Parâmetro	Estimativa	EP
1.9 %	0.7 %	μ_{11}	5.59	0.25	ξ_{11}	0.36	0.08
	1.4 %	μ_{12}	8.91	0.22	ξ_{12}	0.92	0.08
	1.9 %	μ_{13}	10.53	0.20	ξ_{13}	1.00	0.06
	2.4 %	μ_{14}	10.88	0.21	ξ_{14}	1.03	0.08
	3.4 %	μ_{15}	9.98	0.21	ξ_{15}	0.92	0.06
2.4 %	0.7 %	μ_{21}	5.03	0.35	ξ_{21}	0.42	0.08
	1.4 %	μ_{22}	7.22	0.26	ξ_{22}	0.65	0.10
	1.9 %	μ_{23}	8.84	0.22	ξ_{23}	0.86	0.09
	2.4 %	μ_{24}	10.13	0.24	ξ_{24}	1.00	0.07
	3.4 %	μ_{25}	9.97	0.23	ξ_{25}	0.86	0.08
3.4 %	0.7 %	μ_{31}	6.04	0.27	ξ_{31}	0.57	0.09
	1.4 %	μ_{32}	8.42	0.26	ξ_{32}	0.69	0.07
	1.9 %	μ_{33}	9.23	0.23	ξ_{33}	0.73	0.08
	2.4 %	μ_{34}	10.14	0.22	ξ_{34}	0.89	0.07
	3.4 %	μ_{35}	12.09	0.23	ξ_{35}	1.00	0.06
Parâmetro de covariância							
			σ^2	0.66			

EP: erro padrão.

Em todos os casos, os parâmetros são significativamente diferentes de zero ($p < 0.0001$).

A simplificação do modelo saturado (3.1) será feita através da estruturação dos parâmetros de regressão em duas etapas: a primeira consiste na estruturação dos coeficientes das medidas controles e a segunda consiste na estruturação das amplitudes médias ajustadas pelas medidas controles.

3.2.1. Redução dos coeficientes das medidas controles

Numa tentativa de redução do número de coeficientes das medidas controles foram propostos dois modelos alternativos: no primeiro, os parâmetros da covariável são iguais para todas as salinidades testes, independentemente de grupo ($\xi_{jk} = \xi$ em (3.1)) e, no segundo, esses parâmetros são iguais para todas as salinidades testes dentro de cada grupo ($\xi_{jk} = \xi_j$ em (3.1)). Nenhum desses modelos se ajusta aos dados ($p < 0.0001$), indicando que as medidas controles devem realmente ser usadas como covariáveis que mudam com as salinidades (covariáveis dependentes do tempo).

Os coeficientes das medidas controles correspondentes ao modelo saturado estão representados graficamente na Figura 3.4. Observando os gráficos dessa figura juntamente com os valores ajustados da Tabela 3.1, nota-se que os coeficientes de ajuste da amplitude do PAC média pelas medidas controles tendem a ser maiores para as salinidades próximas à salinidade de aclimação, ou seja, quanto maior o módulo da diferença entre as salinidades teste e de aclimação, menor o valor desse coeficiente. Essa correlação negativa entre o valor do coeficiente de ajuste da amplitude do PAC média pelas medidas controles e o módulo da diferença de salinidade não parece ser igual entre os três grupos (Tabela 3.2). Isso pode ser evidenciado observando-se que para a diferença de salinidade -2.7 o coeficiente da medida controle estimado pelo modelo para o grupo 3.4% é de 0.57; para a diferença de -1.7, esse coeficiente estimado para o grupo 2.4% é igual a 0.42 e para a diferença -1.2, o coeficiente estimado para o grupo 1.9% é igual a 0.36.

Figura 3.4. Relação entre a amplitude do PAC e a respectiva medida controle sob o modelo saturado (3.1).

a) Grupo com salinidade de aclimação 1.9%

b) Grupo com salinidade de aclimação 2.4%

c) Grupo com salinidade de aclimação 3.4%

Tabela 3.2. Estimativas dos coeficientes de ajuste da amplitude do PAC média pela medida controle sob o modelo saturado (3.1).

Diferença Entre as Salinidades Teste e de Aclimação (%)	Salinidades de Aclimação		
	1.9%	2.4%	3.4%
-2.7			0.57
-2.0			0.69
-1.7		0.42	
-1.5			0.73
-1.2	0.36		
-1.0		0.65	0.89
-0.5	0.92	0.86	
0	1	1	1
0.5	1.03		
1.0		0.86	
1.5	0.92		

Para incorporar essas conclusões no modelo, considere y_{ij} e W_{ij} definidos em (3.1), Δ_{ijk} o valor da diferença entre a k-ésima salinidade teste e a salinidade de aclimação do i-ésimo molusco do j-ésimo grupo e $I(u)$ a função indicadora de u ($=0$, se u for falso e $=1$, se u for verdadeiro). Um modelo que supõe uma relação inversa entre magnitude da diferença de salinidade e valor do coeficiente da medida controle é da forma (3.1) em que

$$X_{ij} = \begin{pmatrix} I_{t_{ij}} & W_{ij} & \text{diag}(W_{ij}\Delta_{ij}^-) & \text{diag}(W_{ij}\Delta_{ij}^+) \end{pmatrix} \text{ e} \quad (3.2)$$

$$\beta_j = \begin{pmatrix} \mu_{j1} & \mu_{j2} & \cdots & \mu_{jt_{ij}} & \alpha_j & \delta_j^- & \delta_j^+ \end{pmatrix}^t,$$

em que

$$\Delta_{ij}^- = \begin{pmatrix} \Delta_{ij1} \times I(\Delta_{ij1} \leq 0) & \cdots & \Delta_{ijt_{ij}} \times I(\Delta_{ijt_{ij}} \leq 0) \end{pmatrix},$$

$$\Delta_{ij}^+ = \begin{pmatrix} \Delta_{ij1} \times I(\Delta_{ij1} > 0) & \cdots & \Delta_{ijt_{ij}} \times I(\Delta_{ijt_{ij}} > 0) \end{pmatrix},$$

α_j é o coeficiente angular correspondente à medida controle para a salinidade de aclimação do j-ésimo grupo, δ_j^- corresponde à variação no coeficiente α_j por diminuição de 1% na salinidade teste nos casos em que ela é menor que a salinidade de aclimação no j-ésimo grupo e δ_j^+ corresponde à variação no coeficiente α_j por aumento de 1% na salinidade teste nos casos em que ela é maior que a salinidade de aclimação no j-ésimo grupo.² Para facilitar

² O grupo 3.4% não tem este parâmetro, pois não há salinidades testes maiores que esta.

a compreensão, segue um exemplo da matriz $\mathbf{X}_{ij}$ para uma unidade experimental do grupo com aclimatação de 1.9% e com observações coletadas em todas as salinidades testes:

$$\mathbf{X}_{i1} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & w_{i11} & -1.2 \times w_{i11} & 0 \\ 0 & 1 & 0 & 0 & 0 & w_{i12} & -0.5 \times w_{i12} & 0 \\ 0 & 0 & 1 & 0 & 0 & w_{i13} & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & w_{i14} & 0 & 0.5 \times w_{i14} \\ 0 & 0 & 0 & 0 & 1 & w_{i15} & 0 & 1.5 \times w_{i15} \end{pmatrix},$$

em que w_{ijk} denota o elemento da k -ésima linha do vetor das medidas controles centradas $\mathbf{W}_{ij}$.

Tabela 3.3. Estimativas de máxima verossimilhança restrita dos parâmetros do modelo (3.2) e seus respectivos erros padrão.

Salinidade de Aclimação	Salinidade Teste	Parâmetro	Estimativa	Erro Padrão	p
1.9 %	0.7 %	μ_{11}	5.67	0.24	0.0001
	1.4 %	μ_{12}	8.86	0.22	0.0001
	1.9 %	μ_{13}	10.47	0.20	0.0001
	2.4 %	μ_{14}	10.87	0.21	0.0001
	3.4 %	μ_{15}	9.98	0.21	0.0001
2.4 %	0.7 %	μ_{21}	5.04	0.32	0.0001
	1.4 %	μ_{22}	7.23	0.25	0.0001
	1.9 %	μ_{23}	8.82	0.22	0.0001
	2.4 %	μ_{24}	10.11	0.23	0.0001
	3.4 %	μ_{25}	9.97	0.23	0.0001
3.4 %	0.7 %	μ_{31}	6.04	0.26	0.0001
	1.4 %	μ_{32}	8.41	0.23	0.0001
	1.9 %	μ_{33}	9.27	0.22	0.0001
	2.4 %	μ_{34}	10.17	0.21	0.0001
	3.4 %	μ_{35}	12.07	0.22	0.0001

Tabela 3.3. (continuação).

Salinidade de Aclimação	Salinidade Teste	Parâmetro	Estimativa	Erro Padrão	p
1.9 %	0.7 e 1.4 %	δ_1^-	0.54	0.08	0.0001
	1.9 %	α_1	1.06	0.05	0.0001
	2.4 e 3.4%	δ_1^+	-0.09	0.06	0.0997
2.4 %	0.7; 1.4 e 1.9 %	δ_2^-	0.35	0.06	0.0001
	2.4 %	α_2	1.01	0.06	0.0001
	3.4%	δ_2^+	-0.15	0.10	0.1318
3.4 %	0.7; 1.4; 1.9 e 2.4 %	δ_3^-	0.16	0.03	0.0001
	3.4 %	α_3	1.01	0.05	0.0001
Parâmetro de covariância					
			σ^2	0.65	

p: nível descritivo associado ao teste da hipótese de o parâmetro ser nulo.

Quando comparado com o modelo saturado (3.1), o modelo reduzido (3.2) apresenta ajuste adequado tanto pelo teste de Wald ($p=0.7215$) quanto pelo critério de Akaike que, neste modelo, é igual a -284.07 (-289.70 , no caso anterior). Observando as estimativas dos parâmetros do modelo (3.2) na Tabela 3.3, há indícios de que δ_1^+ e δ_2^+ são nulos ($p = 0.0997$ e $p = 0.1318$, respectivamente) e que os coeficientes angulares α_1 , α_2 e α_3 podem ser considerados iguais ($p = 0.7737$).

Esses resultados sugerem uma nova redução no modelo com as seguintes restrições: para salinidades testes maiores ou iguais à de aclimação, o coeficiente de correção da amplitude média pela medida controle é constante e independente da salinidade de aclimação; para diferenças de salinidade negativas, esse coeficiente de correção diminui proporcionalmente a essa diferença. Essas considerações são incorporadas num modelo reduzido com a mesma forma dada em (3.1) em que

$$\mathbf{X}_{ij} = \begin{pmatrix} \mathbf{I}_{t_{ij}} & \mathbf{W}_{ij} & \text{diag}(\mathbf{W}_{ij} \Delta_{ij}^-) \end{pmatrix} \quad (3.3)$$

e

$$\beta_j = \left(\mu_{j1} \quad \mu_{j2} \quad \cdots \quad \mu_{jt_{ij}} \quad \alpha \quad \delta_j^- \right)^t,$$

em que α é o coeficiente angular correspondente à medida controle quando a salinidade teste é maior ou igual à de aclimação, independentemente de grupo e os demais elementos têm as mesmas definições apresentadas anteriormente.

Pelo teste de Wald, não há evidências contrárias à redução do modelo saturado (3.1) para o modelo (3.3) acima descrito ($p=0.8903$).

3.2.2. Redução dos coeficientes das amplitudes médias do PAC ajustadas pelo controle

Conservando a estrutura proposta em (3.3) para os parâmetros correspondentes às medidas controles, os modelos apresentados a seguir concentram-se na estruturação das amplitudes médias do PAC (ajustadas pelo controle).

Figura 3.5. Amplitude média ajustada pela medida controle estimada sob o modelo saturado (3.1).

A Figura 3.5 apresenta os perfis médios da amplitude ajustada pela medida controle sob o modelo saturado (3.1) em função das diferenças de salinidades. Confirmando o que já se havia destacado, a resposta média ajustada também sugere um modelo em que diferenças de salinidade negativas são representadas por um segmento de reta e diferenças de salinidades positivas, por outro, com coeficientes angulares diferentes para cada grupo. Um modelo que impõe, para cada grupo, a condição de estas retas serem concorrentes no ponto onde a diferença de salinidade é nula é da forma dada em (3.1) com

$$\begin{aligned}
 \mathbf{X}_{ij} &= (\mathbf{1}_{i,j} \quad \Delta_{ij}^- \quad \Delta_{ij}^+ \quad \mathbf{W}_{ij} \quad \text{diag}(\mathbf{W}_{ij} \Delta_{ij}^-)), \\
 \beta_1 &= (\phi_1 \quad \psi_1^- \quad \psi_1^+ \quad \alpha \quad \delta_1^-)^t, \\
 \beta_2 &= (\phi_2 \quad \psi_2^- \quad \psi_2^+ \quad \alpha \quad \delta_2^-)^t \\
 \text{e } \beta_3 &= (\phi_3 \quad \psi_3^- \quad \alpha \quad \delta_3^-)^t,
 \end{aligned} \tag{3.4}$$

sendo ϕ_j a amplitude média corrigida pelo controle para salinidade teste igual à de aclimação no j-ésimo grupo, ψ_j^- a taxa de variação de ϕ_j por diminuição de 1% na salinidade teste nos casos em que ela é menor que a salinidade de aclimação no j-ésimo grupo e ψ_j^+ a taxa de variação de ϕ_j por aumento de 1% na salinidade teste nos casos em que ela é maior que a salinidade de aclimação no j-ésimo grupo. Os demais elementos foram definidos anteriormente. A Tabela 3.4 traz as estimativas dos parâmetros desse modelo e seus erros padrão.

Tabela 3.4. Estimativas de máxima verossimilhança restrita dos parâmetros do modelo (3.4) e seus respectivos erros padrão.

Salinidade de Aclimação	Salinidade Teste	Parâmetro	Estimativa	Erro Padrão	p
1.9 %	0.7 e 1.4 %	ψ_1^-	4.20	0.25	0.0001
	1.9 %	ϕ_1	10.80	0.16	0.0001
	2.4 e 3.4 %	ψ_1^+	-0.50	0.18	0.0058
2.4 %	0.7; 1.4 e 1.9 %	ψ_2^-	3.02	0.21	0.0001
	2.4 %	ϕ_2	10.24	0.17	0.0001
	3.4 %	ψ_2^+	-0.42	0.27	0.1207
3.4 %	0.7; 1.4; 1.9 e 2.4 %	ψ_3^-	2.13	0.11	0.0001
	3.4 %	ϕ_3	12.29	0.17	0.0001
1.9 %	0.7 e 1.4 %	δ_1^-	0.45	0.06	0.0001
2.4 %	0.7; 1.4 e 1.9 %	δ_2^-	0.31	0.04	0.0001
3.4 %	0.7; 1.4; 1.9 e 2.4 %	δ_3^-	0.15	0.02	0.0001
Todos	Todas	α	0.97	0.02	0.0001
Parâmetro de covariância					
			σ^2	0.68	

p: nível descritivo associado ao teste da hipótese de o parâmetro ser nulo.

A redução do modelo saturado para o modelo dado em (3.4) pode ser considerada aceitável pelo teste de Wald ($p=0.3968$). Uma redução mais drástica incorpora a igualdade dos coeficientes angulares para os segmentos de reta correspondentes a diferenças de salinidades positivas. Nesse caso, o modelo é expresso na forma (3.1) com

$$\begin{aligned} \mathbf{X}_{ij} &= (\mathbf{1}_{i_j} \quad \Delta_{ij}^- \quad \Delta_{ij}^+ \quad \mathbf{W}_{ij} \quad \text{diag}(\mathbf{W}_{ij} \Delta_{ij}^-)), \\ \beta_1 &= (\phi_1 \quad \psi_1^- \quad \psi^+ \quad \alpha \quad \delta_1^-)^t, \\ \beta_2 &= (\phi_2 \quad \psi_2^- \quad \psi^+ \quad \alpha \quad \delta_2^-)^t \end{aligned} \quad (3.5)$$

e

$$\beta_3 = (\phi_3 \quad \psi_3^- \quad \alpha \quad \delta_3^-)^t,$$

em que ψ^+ é a variação no coeficiente ϕ_j por aumento de 1% na salinidade teste nos casos em que ela é maior que a salinidade de aclimação, independentemente de grupo. Os demais elementos obedecem às definições dos modelos anteriores.

Aqui, a redução do modelo saturado para o modelo (3.5) também não é contrariada pelo teste de Wald ($p=0.4554$). As estimativas dos parâmetros correspondentes estão apresentadas na Tabela 3.5 e indicam que a amplitude média ajustada tende a diminuir linearmente conforme a salinidade teste se distancia da de aclimação. Para diferenças de salinidades positivas, a “intensidade de queda” da amplitude média do PAC é a mesma tanto para o grupo com salinidade de aclimação igual a 1.9% quanto para aquele com salinidade de aclimação igual a 2.4%. Por outro lado, para variações de salinidade negativas, esta amplitude média também diminui, mas de maneira distinta em cada um dos três grupos. Quanto menor a salinidade de aclimação, maior é esta “intensidade de queda”.

Tabela 3.5. Estimativas de máxima verossimilhança restrita dos parâmetros do modelo (3.5) e seus respectivos erros padrão.

Salinidade de Aclimação	Salinidade Teste	Parâmetro	Estimativa	Erro Padrão	p
1.9 %	0.7 e 1.4 %	ψ_1^-	4.18	0.24	0.0001
	1.9 %	ϕ_1	10.79	0.14	0.0001
2.4 %	0.7; 1.4 e 1.9 %	ψ_2^-	3.03	0.20	0.0001
	2.4 %	ϕ_2	10.26	0.15	0.0001
1.9% e 2.4%		ψ^+	-0.48	0.15	0.0016
3.4 %	0.7; 1.4; 1.9 e 2.4 %	ψ_3^-	2.13	0.11	0.0001
	3.4 %	ϕ_3	12.29	0.17	0.0001

Tabela 3.5. (continuação).

Salinidade de Aclimação	Salinidade Teste	Parâmetro	Estimativa	Erro Padrão	p
1.9 %	0.7 e 1.4 %	δ_1^-	0.45	0.06	0.0001
2.4 %	0.7; 1.4 e 1.9 %	δ_2^-	0.32	0.04	0.0001
3.4 %	0.7; 1.4; 1.9 e 2.4 %	δ_3^-	0.15	0.02	0.0001
Todos	Todas	α	0.97	0.02	0.0001
Parâmetro de covariância					
			σ^2	0.68	

p: nível descritivo associado ao teste da hipótese de o parâmetro ser nulo.

A Figura 3.6 ilustra os perfis médios da amplitude ajustada pelas medidas controles sob o modelo (3.5) em função das diferenças de salinidade.

Figura 3.6. Amplitude média ajustada pelas medidas controles estimada sob o modelo (3.5).

Essa análise exploratória permite a identificação de um modelo plausível para os parâmetros de regressão, sem a preocupação com a estrutura de covariância intraunidades amostrais. Uma reanálise dos dados sob um modelo que engloba os dois aspectos será desenvolvida no próximo capítulo possibilitando uma comparação das conclusões das duas formas de análise.

Capítulo 4

Modelagem Simultânea dos Parâmetros de Regressão e de Covariância

Este capítulo irá concentrar-se na estruturação tanto dos parâmetros de regressão quanto na matriz de covariâncias intraunidades amostrais Σ_{ij} . O modelo linear geral (1.1), como no Capítulo 3, será usado para a análise dos dados, porém sem a suposição de que $\Sigma_{ij} = \sigma^2 \mathbf{I}_{t_{ij}}$. Estruturas apropriadas que permitam incorporar, não só uma possível dependência entre as observações feitas em um mesmo molusco, como também sua heterocedasticidade, serão consideradas. Particularmente, neste trabalho, essa modelagem será feita por meio de modelos de efeitos aleatórios e autorregressivos.

4.1. Estratégias de ajuste dos modelos

Vários autores têm estudado o problema da estratégia de análise em situações em que a modelagem dos parâmetros de regressão e dos parâmetros de covariância pode ser feita separada ou simultaneamente.

Grady and Helms (1995), por exemplo, sugerem que os parâmetros de regressão sejam modelados sob uma matriz de covariância não estruturada, ou seja, ajustando o melhor modelo reduzido nos parâmetros de regressão e saturado nos parâmetros de covariância. A modelagem da matriz de covariância, conservando a estrutura adotada para os parâmetros de regressão, deve ser feita a seguir. Diggle, Liang and Zeger (1994), por sua vez, sugerem estruturar a matriz de covariância sob um modelo saturado nos parâmetros de regressão e, em seguida, modelar esses parâmetros usando a estrutura de dependência escolhida para os dados.

Ambas as estratégias têm vantagens e desvantagens. Na primeira, o problema aparece quando os dados são incompletos, pois as observações podem ser tão esparsas que a matriz de covariância pode não ser estimável ou gerar estimativas não confiáveis. A segunda, por sua vez, não se aplica na presença de uma variável independente contínua, porque, nesses casos, normalmente não existe a possibilidade de ajustar-se um modelo saturado aos parâmetros de regressão, ou seja, os modelos já devem incorporar a forma da relação entre as variáveis independente (contínua) e resposta, seja ela linear, quadrática ou outra qualquer.

No caso dos dados analisados neste trabalho, a estratégia de análise foi diferente das propostas por esses autores. Estudou-se, inicialmente, um modelo saturado tanto nos parâmetros de regressão quanto nos de covariância, considerando uma relação linear entre as medidas controles e a resposta da amplitude do PAC. Os parâmetros de regressão deste modelo saturado foram definidos da mesma maneira que em (3.1), ou seja, uma amplitude média ajustada pelo controle e um coeficiente de ajuste da amplitude média do PAC pelo controle para cada combinação de grupo e salinidade teste¹. A tática adotada para os parâmetros de covariância constituiu na avaliação de diferentes estruturas para a matriz de covariância sob o modelo saturado nos parâmetros de regressão. Adotando a estrutura para Σ_{ij} que melhor se ajustou aos dados, procurou-se, então, modelar os parâmetros de regressão. Note que esta estratégia não é a mesma sugerida por Diggle, Liang and Zeger (1994), uma vez

¹ Esta parametrização determina que a relação entre a amplitude do PAC e medidas controles é linear, com diferentes coeficientes para cada combinação de grupo e salinidade teste. Pela Figura 3.3, esta suposição, que já foi adotada na análise do capítulo anterior, parece ser bem razoável.

que o modelo aqui chamado de saturado impõe que a relação entre as medidas controles e a resposta da amplitude do PAC é linear.

4.2. Análise descritiva da estrutura de covariância amostral

A proposta de um modelo para a estrutura de covariância intraunidades amostrais, em geral, é baseada numa análise descritiva que explora como os valores obtidos numa mesma unidade experimental se relacionam uns com os outros. Para remover os possíveis efeitos de diferença de salinidade e das medidas controles sobre os valores da amplitude do PAC, o estudo descritivo da estrutura de Σ_{ij} foi baseado nos resíduos gerados pelo ajuste do modelo (3.1).²

A análise descritiva da estrutura de covariância foi baseada nas matrizes de covariância e correlação amostrais e no variograma (Diggle, Liang and Zeger (1994)).

4.2.1 Análise descritiva baseada nas matrizes de covariância e correlação

A Tabela 4.1 apresenta as covariâncias e as correlações amostrais. Esses valores foram calculados apenas com os moluscos aferidos em ambas as salinidades (o que indica que essas estimativas não foram calculadas com o mesmo número de observações).

² É importante lembrar que, para salinidades testes iguais à de aclimação, o resíduo é igual a zero, pois foi feita a suposição de que a resposta e o controle da amplitude do PAC eram iguais nessas salinidades. Por essa razão, estas salinidades não foram consideradas no estudo para a estruturação da matriz Σ_{ij} .

Tabela 4.1. Matrizes de covariância (região hachurada) e correlação (região em branco) dos resíduos.

a) Grupo com salinidade de aclimação 1.9%

Salinidade Teste	0.7%	1.4%	2.4%	3.4%
0.7%	0.53	0.19	0.00	-0.05
1.4%	0.42	0.30	-0.10	-0.24
2.4%	0.00	-0.31	0.31	0.43
3.4%	-0.05	-0.32	0.67	1.37

b) Grupo com salinidade de aclimação 2.4%

Salinidade Teste	0.7%	1.4%	1.9%	3.4%
0.7%	1.32	0.26	-0.02	0.32
1.4%	0.36	0.31	0.05	-0.08
1.9%	-0.02	0.13	0.42	0.23
3.4%	0.21	-0.17	0.54	1.24

c) Grupo com salinidade de aclimação 3.4%

Salinidade Teste	0.7%	1.4%	1.9%	2.4%
0.7%	1.14	0.80	0.30	0.50
1.4%	0.67	0.96	0.29	0.84
1.9%	0.30	0.39	0.84	0.23
2.4%	0.40	0.71	0.29	1.08

Para facilitar a interpretação do comportamento das covariâncias e correlações em função das salinidades teste, construíram-se gráficos que relacionam os valores dessas covariâncias e correlações com diferença de salinidade (Figura 4.1), seguindo a sugestão de Grady and Helms (1995). Cada linha dos gráficos apresentados à esquerda na Figura 4.1 representa as covariâncias entre os valores dos resíduos de uma determinada salinidade teste e de todas as salinidades teste maiores ou iguais a ela. Analogamente, podem-se interpretar os gráficos das correlações apresentados à direita na Figura 4.1. Observe que as covariâncias relacionadas com diferenças nulas de salinidade representam as variâncias dos resíduos em cada salinidade teste.

Figura 4.1. Covariâncias e correlações dos resíduos em função da diferença de salinidade.

a) Grupo com salinidade de aclimação 1.9%

b) Grupo com salinidade de aclimação 2.4%

c) Grupo com salinidade de aclimação 3.4%

Com base na Tabela 4.1 e na Figura 4.1, pode-se observar que:

- A variância dos resíduos tende a crescer com o aumento da magnitude da diferença de salinidade;
- Nos grupos com aclimação de 1.9% e 3.4%, as correlações mais altas são aquelas entre os resíduos obtidos em diferenças de salinidades de mesmo sinal, ou seja, as salinidades testes representadas pelo mesmo segmento de reta no modelo dado por (3.5);
- No grupo com salinidade de aclimação de 2.4%, as correlações não apresentam uma relação muito clara com as diferenças de salinidades.

Uma outra alternativa que pode facilitar a identificação de algum padrão no comportamento das covariâncias é o gráfico dos perfis das colunas da matriz de covariância dos resíduos (Suyama and Singer, 1996). Nesses gráficos, perfis de covariâncias com padrão comum, podem sugerir a inclusão de termos de efeitos aleatórios em um modelo de efeitos aleatórios com independência condicional. Por exemplo, um comportamento quadrático dos perfis sugere um modelo de efeitos aleatórios com independência condicional, que relaciona o valor esperado da amplitude do PAC e diferença de salinidade através de um polinômio de grau 2, com 3 efeitos aleatórios: um no intercepto, outro no coeficiente linear e outro no coeficiente quadrático. Para os dados sob análise, esses gráficos estão apresentados na Figura 4.2 e não se observa um padrão no comportamento dos perfis de covariâncias.

Figura 4.2. Gráfico dos perfis das colunas da matriz de covariâncias dos resíduos.

a) Grupo com salinidade de aclimação 1.9%

b) Grupo com salinidade de aclimação 2.4%

c) Grupo com salinidade de aclimação 3.4%

4.2.2 Análise descritiva baseada no variograma

A utilização do variograma para se estudar o comportamento da estrutura de covariância entre observações obtidas em uma mesma unidade amostral é uma alternativa interessante nos casos de dados incompletos e desbalanceados, onde nem sempre o cálculo das covariâncias e correlações amostrais é possível. No contexto de medidas repetidas, o variograma amostral é um gráfico que relaciona a distância entre dois resíduos (r_{ip} e r_{iq}) obtidos nos tempos z_{ip} e z_{iq} associados a uma mesma unidade experimental,

$$v_{ipq} = \frac{1}{2}(r_{ip} - r_{iq})^2,$$

e o correspondente intervalo de tempo (ou, especificamente no conjunto de dados deste trabalho, diferença de salinidade), chamado de defasagem,

$$u_{ipq} = z_{ip} - z_{iq},$$

$p > q$, $i=1, \dots, n$. O aumento dos valores v_{ipq} para valores crescentes da defasagem u_{ipq} indica que a correlação decresce com o tempo.

Quando mais de uma unidade experimental for observada nos mesmos momentos de avaliação p e q , existirá mais de um valor de v para a mesma defasagem u . Alguns autores sugerem que se considere a média de todos os valores de v correspondentes a cada u . Porém, na presença de algum valor aberrante de v , a média pode não caracterizar bem o comportamento dos dados. Nesses casos, a representação de todos os pontos, sem resumí-los através da média, e a inclusão de uma reta de regressão que relacione v e u podem ser mais ilustrativos. Para o problema em estudo, os variogramas estão apresentados na Figura 4.3. Note que os pontos discrepantes realmente influenciam muito o valor médio de v para cada defasagem, podendo resultar numa má interpretação do variograma (gráficos apresentados à direita na Figura 4.3). Observando o gráfico com os pontos individuais (apresentados à esquerda na Figura 4.3), tem-se alguma evidência de que a correlação decresce com o aumento das diferenças de salinidade.

Pelo fato de existirem poucas defasagens, isto é, poucas diferenças de salinidade, pode-se dizer, neste exemplo, que o variograma não parece ser uma boa ferramenta para a análise da dependência dos dados. Neste caso, a modelagem da estrutura de dependência dos dados deve ser avaliada principalmente a partir de outras ferramentas, como as covariâncias e correlações amostrais.

Figura 4.3. Variograma dos resíduos em função da defasagem (diferença de salinidades):

• pontos individuais, — reta de regressão, — média, - - - média excluindo os pontos apontados pelas flechas.

a) Grupo com salinidade de aclimação 1.9%

b) Grupo com salinidade de aclimação 2.4%

c) Grupo com salinidade de aclimação 3.4%

4.3. Ajuste da estrutura de Σ_{ij} para o modelo saturado nos parâmetros de regressão

Uma das opções para a estrutura de Σ_{ij} é o modelo de efeitos aleatórios. O gráfico dos perfis individuais da amplitude do PAC em função da diferença de salinidade pode auxiliar a decisão. Pelo gráfico, apresentado na Figura 4.4, observa-se que os perfis individuais da amplitude do PAC têm a mesma relação entre a amplitude do PAC e diferença de salinidade. Isso sugere que um modelo de efeitos aleatórios pode representar uma boa alternativa de análise (Diggle, Liang and Zeger, 1994).

Figura 4.4. Perfis individuais (—) e médios (—) da amplitude do PAC (mV) em função da diferença de salinidade.

a) Grupo com salinidade de aclimação 1.9%

Figura 4.4. (continuação).

b) Grupo com salinidade de aclimação 2.4%

c) Grupo com salinidade de aclimação 3.4%

Com base nesses gráficos e nas tendências observadas da estrutura de covariância dos resíduos na Seção 4.2.1, foi proposto um modelo de efeitos aleatórios saturado para os parâmetros de regressão como definido em (3.1) e com a seguinte estrutura para a matriz de covariância:

$$\Sigma_{ij} = \mathbf{Z}_{ij} \mathbf{G}_j \mathbf{Z}_{ij}^t + \mathbf{R}_{ij}, \quad (4.1)$$

para $i=1, \dots, n_j$ ($n_1=20$ e $n_2=n_3=18$); $j=1,2,3$,

$$\mathbf{Z}_{i1} = \begin{pmatrix} -1.2 & 0 \\ -0.5 & 0 \\ 0 & 0 \\ 0 & 0.5 \\ 0 & 1.5 \end{pmatrix}, \quad \mathbf{Z}_{i2} = \begin{pmatrix} -1.7 & 0 \\ -1 & 0 \\ -0.5 & 0 \\ 0 & 0 \\ 0 & 1 \end{pmatrix}, \quad \mathbf{Z}_{i3} = \begin{pmatrix} -2.7 \\ -2 \\ -1.5 \\ -1 \\ 0 \end{pmatrix},$$

$$\mathbf{G}_1 = \begin{pmatrix} \varphi_{11}^2 & 0 \\ 0 & \varphi_{21}^2 \end{pmatrix}, \mathbf{G}_2 = \begin{pmatrix} \varphi_{12}^2 & 0 \\ 0 & \varphi_{22}^2 \end{pmatrix}, \mathbf{G}_3 = (\varphi_{13}^2) \text{ e } \mathbf{R}_{ij} = \sigma^2 \mathbf{I}.$$

Esse modelo induz a seguinte estrutura de covariância intraunidades amostrais:

$$\begin{aligned} \Sigma_{i1} &= \begin{pmatrix} 1.4\varphi_{11}^2 + \sigma^2 & 0.6\varphi_{11}^2 & 0 & 0 & 0 \\ 0.6\varphi_{11}^2 & 0.3\varphi_{11}^2 + \sigma^2 & 0 & 0 & 0 \\ 0 & 0 & \sigma^2 & 0 & 0 \\ 0 & 0 & 0 & 0.3\varphi_{21}^2 + \sigma^2 & 0.8\varphi_{21}^2 \\ 0 & 0 & 0 & 0.8\varphi_{21}^2 & 2.3\varphi_{21}^2 + \sigma^2 \end{pmatrix}, \\ \Sigma_{i2} &= \begin{pmatrix} 2.9\varphi_{12}^2 + \sigma^2 & 1.7\varphi_{12}^2 & 0.9\varphi_{12}^2 & 0 & 0 \\ 1.7\varphi_{12}^2 & \varphi_{12}^2 + \sigma^2 & 0.5\varphi_{12}^2 & 0 & 0 \\ 0.9\varphi_{12}^2 & 0.5\varphi_{12}^2 & 0.3\varphi_{12}^2 + \sigma^2 & 0 & 0 \\ 0 & 0 & 0 & \sigma^2 & 0 \\ 0 & 0 & 0 & 0 & \varphi_{22}^2 + \sigma^2 \end{pmatrix} \text{ e} \\ \Sigma_{i3} &= \begin{pmatrix} 7.3\varphi_{13}^2 + \sigma^2 & 5.4\varphi_{13}^2 & 4.1\varphi_{13}^2 & 2.7\varphi_{13}^2 & 0 \\ 5.4\varphi_{13}^2 & 4.0\varphi_{13}^2 + \sigma^2 & 3.0\varphi_{13}^2 & 2.0\varphi_{13}^2 & 0 \\ 4.1\varphi_{13}^2 & 3.0\varphi_{13}^2 & 2.3\varphi_{13}^2 + \sigma^2 & 1.5\varphi_{13}^2 & 0 \\ 2.7\varphi_{13}^2 & 2.0\varphi_{13}^2 & 1.5\varphi_{13}^2 & 2.3\varphi_{13}^2 + \sigma^2 & 0 \\ 0 & 0 & 0 & 0 & \sigma^2 \end{pmatrix} \end{aligned}$$

O modelo (4.1) contempla dois efeitos aleatórios para cada um dos grupos com salinidades de aclimação 1.9% e 2.4% (um para diferenças de salinidades negativas e outro para diferenças de salinidades positivas) e um efeito aleatório para o grupo com salinidade de aclimação 3.4%³. Conseqüentemente, tem-se uma matriz de covariância intraunidades amostrais ajustada para cada grupo. Além disso, esse modelo assume que i) as observações obtidas em salinidades testes maiores que a de aclimação não estão correlacionadas com aquelas obtidas em salinidades testes menores que a de aclimação, ii) as medidas observadas na salinidade de aclimação não estão correlacionadas com as demais, iii) a variância da amplitude do PAC aumenta com o aumento da magnitude da diferença de salinidade, e iv) a correlação diminui com o aumento da magnitude da diferença de salinidade.

Dado que a análise gráfica das covariâncias intraunidades amostrais não apresenta evidências fortes de nenhum modelo em particular, outras alternativas para Σ_{ij} também foram propostas com finalidade de comparação. Entre eles:

³ Note que a matriz de especificação dos efeitos aleatórios tem uma coluna com os valores das diferenças de salinidades negativas e outra com os valores das diferenças de salinidades positivas.

A) Uma estrutura uniforme com diferentes parâmetros de covariância para cada grupo, ou seja, Σ_{ij} como definido em (4.1) sendo

$$\mathbf{Z}_{i1} = \mathbf{Z}_{i2} = \mathbf{Z}_{i3} = \mathbf{1}_5, \mathbf{G}_j = (\rho_j) \text{ e } \mathbf{R}_{ij} = \sigma^2 \mathbf{I}. \quad (4.2)$$

Este modelo uniforme supõe que, para cada um dos 3 grupos, a variância da amplitude do PAC é igual nas 5 salinidades testes e que as correlações entre as amplitudes do PAC obtidas em diferentes salinidades testes também é constante.

B) Uma estrutura que supõe independência e heterocedasticidade entre as respostas observadas num mesmo molusco sobre diferentes salinidades teste, ou seja,

$$\Sigma_{ij} = \begin{pmatrix} \sigma_1^2 & 0 & \dots & 0 \\ 0 & \sigma_2^2 & \dots & 0 \\ \vdots & \dots & \ddots & \vdots \\ 0 & \dots & 0 & \sigma_{t_{ij}}^2 \end{pmatrix}. \quad (4.3)$$

Sob este modelo, as estruturas para as matrizes de covariâncias intraunidades experimentais são iguais entre os 3 grupos, com diferentes variâncias da amplitude do PAC em cada salinidade teste. Além disso, assume que as observações feitas num mesmo molusco não estão correlacionadas.

C) Uma estrutura de correlação exponencial com diferentes parâmetros de covariância para cada grupo, ou seja,

$$\Sigma_{ij} = \begin{pmatrix} \sigma_j^2 & \rho_j^{z_{ij2}-z_{ij1}} & \dots & \rho_j^{z_{ijt_{ij}}-z_{ij1}} \\ \rho_j^{z_{ij2}-z_{ij1}} & \sigma_j^2 & \dots & \rho_j^{z_{ijt_{ij}}-z_{ij2}} \\ \vdots & \dots & \ddots & \vdots \\ \rho_j^{z_{ijt_{ij}}-z_{ij1}} & \dots & \rho_j^{z_{ijt_{ij}}-z_{ij4}} & \sigma_j^2 \end{pmatrix}, \quad (4.4)$$

em que z_{ijk} denota a salinidade teste do i -ésimo molusco do j -ésimo grupo na k -ésima condição de avaliação.

Este modelo também supõe que, para cada um dos 3 grupos, as variâncias da amplitude do PAC são iguais para as diferentes salinidades testes e que as correlações entre as amplitudes do PAC obtidas em diferentes salinidades testes diminuem com o aumento da magnitude dessas diferenças.

A Tabela 4.2 apresenta as estatísticas de Akaike e de -2 vezes o logaritmo da verossimilhança restrita para a avaliação do ajuste dos modelos acima apresentados.

Tabela 4.2. Comparação dos modelos saturados (4.1), (4.2), (4.3) e (4.4).

Modelo	Nº Parâmetros de Covariância	Critério de Akaike	$-2 l_r$
(4.1) <i>efeitos aleatórios</i>	6	-261.97	509.94
(4.2) <i>uniforme</i>	4	-287.20	566.39
(4.3) <i>independência e heterocedasticidade</i>	5	-289.26	568.52
(4.4) <i>correlação exponencial</i>	6	-284.78	557.56

$-2 l_r$: -2 vezes o logaritmo da verossimilhança restrita.

Pelo critério de Akaike, o modelo de efeitos aleatórios dado em (4.1) é o que se ajusta melhor aos dados. Pode-se observar o mesmo comparando-se os valores estimados das matrizes de covariâncias intraunidades experimentais sob cada um desses modelos, apresentados nas Tabelas 4.3 a 4.6, com os valores das covariâncias amostrais, apresentados na Tabela 4.1.

Tabela 4.3. Matrizes de covariância intraunidades experimentais estimadas sob o modelo (4.1).

a) Grupo com salinidade de aclimação 1.9%

Salinidade Teste	0.7%	1.4%	1.9%	2.4%	3.4%
0.7%	0.59	0.14	0	0	0
1.4%	0.14	0.32	0	0	0
1.9%	0	0	0.26	0	0
2.4%	0	0	0	0.40	0.41
3.4%	0	0	0	0.41	1.50

b) Grupo com salinidade de aclimação 2.4%

Salinidade Teste	0.7%	1.4%	1.9%	2.4%	3.4%
0.7%	1.08	0.48	0.24	0	0
1.4%	0.48	0.54	0.14	0	0
1.9%	0.24	0.14	0.33	0	0
2.4%	0	0	0	0.26	0
3.4%	0	0	0	0	1.34

c) Grupo com salinidade de aclimação 3.4%

Salinidade Teste	0.7%	1.4%	1.9%	2.4%	3.4%
0.7%	1.60	0.99	0.74	0.49	0
1.4%	0.99	0.99	0.55	0.37	0
1.9%	0.74	0.55	0.67	0.27	0
2.4%	0.49	0.37	0.27	0.44	0
3.4%	0	0	0	0	0.26

Tabela 4.4. Matrizes de covariância intraunidades experimentais estimadas sob o modelo (4.2).

a) Grupo com salinidade de aclimação 1.9%

Salinidade Teste	0.7%	1.4%	1.9%	2.4%	3.4%
0.7%	0.57	0.02	0.02	0.02	0.02
1.4%	0.02	0.57	0.02	0.02	0.02
1.9%	0.02	0.02	0.57	0.02	0.02
2.4%	0.02	0.02	0.02	0.57	0.02
3.4%	0.02	0.02	0.02	0.02	0.57

b) Grupo com salinidade de aclimação 2.4%

Salinidade Teste	0.7%	1.4%	1.9%	2.4%	3.4%
0.7%	0.64	0.10	0.10	0.10	0.10
1.4%	0.10	0.64	0.10	0.10	0.10
1.9%	0.10	0.10	0.64	0.10	0.10
2.4%	0.10	0.10	0.10	0.64	0.10
3.4%	0.10	0.10	0.10	0.10	0.64

c) Grupo com salinidade de aclimação 3.4%

Salinidade Teste	0.7%	1.4%	1.9%	2.4%	3.4%
0.7%	0.76	0.21	0.21	0.21	0.21
1.4%	0.21	0.76	0.21	0.21	0.21
1.9%	0.21	0.21	0.76	0.21	0.21
2.4%	0.21	0.21	0.21	0.76	0.21
3.4%	0.21	0.21	0.21	0.21	0.76

Tabela 4.5. Matriz de covariância intraunidades experimentais estimada sob o modelo (4.3).

Salinidade Teste	0.7%	1.4%	1.9%	2.4%	3.4%
0.7%	0.77	0	0	0	0
1.4%	0	0.45	0	0	0
1.9%	0	0	0.59	0	0
2.4%	0	0	0	0.51	0
3.4%	0	0	0	0	1.38

Tabela 4.6. Matrizes de covariância intraunidades experimentais estimadas sob o modelo (4.4).

a) Grupo com salinidade de aclimação 1.9%

Salinidade Teste	0.7%	1.4%	1.9%	2.4%	3.4%
0.7%	0.63	0.36	0.24	0.16	0.07
1.4%	0.36	0.63	0.42	0.28	0.13
1.9%	0.24	0.42	0.63	0.42	0.19
2.4%	0.16	0.28	0.42	0.63	0.28
3.4%	0.07	0.13	0.19	0.28	0.63

b) Grupo com salinidade de aclimação 2.4%

Salinidade Teste	0.7%	1.4%	1.9%	2.4%	3.4%
0.7%	0.68	0	0	0	0
1.4%	0	0.68	0	0	0
1.9%	0	0	0.68	0	0
2.4%	0	0	0	0.68	0
3.4%	0	0	0	0	0.68

c) Grupo com salinidade de aclimação 3.4%

Salinidade Teste	0.7%	1.4%	1.9%	2.4%	3.4%
0.7%	0.75	0.26	0.12	0.06	0.01
1.4%	0.26	0.75	0.35	0.16	0.04
1.9%	0.12	0.35	0.75	0.35	0.08
2.4%	0.06	0.16	0.35	0.75	0.16
3.4%	0.01	0.04	0.08	0.16	0.75

As estimativas dos parâmetros de regressão e de covariância sob o modelo (4.1) encontram-se na Tabela 4.7.

Tabela 4.7. Estimativas de máxima verossimilhança restrita e respectivos erros padrão dos parâmetros do modelo (4.1) sob uma estrutura saturada para os parâmetros de regressão.

Salinidade Aclimação	Salinidade Teste	Amplitude Média Ajustada Pela Medida Controle			Coeficientes das Medidas Controles		
		Parâmetro	Estimativa	EP	Parâmetro	Estimativa	EP
1.9 %	0.7 %	μ_{11}	5.58	0.23	ξ_{11}	0.37	0.07
	1.4 %	μ_{12}	8.90	0.15	ξ_{12}	0.90	0.06
	1.9 %	μ_{13}	10.53	0.13	ξ_{13}	1.00	0.04
	2.4 %	μ_{14}	10.85	0.16	ξ_{14}	1.01	0.06
	3.4 %	μ_{15}	9.99	0.31	ξ_{15}	0.91	0.09
2.4 %	0.7 %	μ_{21}	4.90	0.39	ξ_{21}	0.38	0.09
	1.4 %	μ_{22}	7.00	0.22	ξ_{22}	0.57	0.07
	1.9 %	μ_{23}	8.78	0.16	ξ_{23}	0.84	0.06
	2.4 %	μ_{24}	10.13	0.15	ξ_{24}	1.00	0.04
	3.4 %	μ_{25}	9.97	0.33	ξ_{25}	0.86	0.11
3.4 %	0.7 %	μ_{31}	6.11	0.35	ξ_{31}	0.62	0.08
	1.4 %	μ_{32}	8.32	0.27	ξ_{32}	0.69	0.06
	1.9 %	μ_{33}	9.34	0.21	ξ_{33}	0.78	0.06
	2.4 %	μ_{34}	10.12	0.17	ξ_{34}	0.93	0.05
	3.4 %	μ_{35}	12.09	0.14	ξ_{35}	1.00	0.04
Parâmetros de covariância							
1.9 %	0.7 e 1.4 %	φ_{11}^2	0.23				
	2.4 e 3.4 %	φ_{21}^2	0.55				
2.4 %	0.7 a 1.9 %	φ_{12}^2	0.28				
	3.4 %	φ_{22}^2	1.08				
3.4 %	0.7 a 2.4 %	φ_{13}^2	0.18				
		σ^2	0.26				

EP: erro padrão.

4.4. Modelagem dos parâmetros de regressão sob a estrutura (4.1) para a matriz de covariância intraunidades amostrais

Avaliou-se a possibilidade de ajustar modelos com estrutura reduzida para os parâmetros de regressão, conservando o padrão definido em (4.1) para a matriz de covariância intraunidades amostrais. A estratégia adotada para a redução dos parâmetros de regressão do modelo (4.1) foi a mesma utilizada no capítulo anterior.

4.4.1. Redução dos coeficientes das medidas controles

Uma estrutura alternativa para os parâmetros associados às medidas controles está dada em (3.3). Assume-se que, para todas as diferenças de salinidades positivas, o coeficiente da medida controle é constante e que, para diferenças negativas de salinidades, esse coeficiente diminui proporcionalmente a essa diferença, segundo o grupo. Esse modelo é dado por

$$y_{ij} = X_{ij}\beta_j + Z_{ij}\gamma_{ij} + \varepsilon_{ij} \quad (4.5)$$

sendo X_{ij} e β_j definidos em (3.3), $\gamma_{ij} = (\gamma_{ij}^- \quad \gamma_{ij}^+)^t$ denotando o vetor de efeitos aleatórios de dimensão (2×1) para $j=1, 2$ e de dimensão (1×1) para $j=3$ e $\varepsilon_{ij} = (\varepsilon_{ij1} \quad \varepsilon_{ij2} \quad \dots \quad \varepsilon_{ijt_{ij}})^t$ representando o vetor de erros aleatórios. Assume-se que γ_{ij} e ε_{ij} são independentes e normalmente distribuídos com vetores de médias $\mathbf{0}$ e matrizes de covariâncias G_j e $\sigma^2 I$ e que Z_{ij} e G_j são dadas em (4.1).

O teste de Wald para avaliar a redução do modelo saturado (4.1) para o modelo (4.5) apresenta um nível descritivo $p=0.6457$, sugerindo que a estrutura dos coeficientes das medidas controles adotada para o modelo de independência também se ajusta satisfatoriamente quando se incorpora uma estrutura de correlação não nula entre as observações intraunidades amostrais.

4.4.2. Redução dos coeficientes das amplitudes médias do PAC ajustadas pelo controle

A Figura 4.5 representa os valores da amplitude média do PAC ajustada pela medida controle estimados sob o modelo (4.1). Este gráfico é bastante parecido com o da amplitude média ajustada estimada no modelo saturado do capítulo anterior (Figura 3.5). Assim, uma possível estrutura para esses parâmetros é dada em (3.5).

Figura 4.5. Amplitude média ajustada pela medida controle estimada no modelo saturado (4.1).

O modelo correspondente é da forma dada em (4.5) com:

$$X_{ij} \text{ e } \beta_j \text{ definidos em (3.5)} \quad (4.6)$$

e

$$\Sigma_{ij} \text{ definido em (4.1).}$$

As estimativas dos parâmetros desse modelo encontram-se na Tabela 4.8.

Tabela 4.8. Estimativas de máxima verossimilhança restrita e respectivos erros padrão dos parâmetros do modelo (4.6).

Salinidade de Aclimação	Salinidade Teste	Parâmetro	Estimativa	Erro Padrão	p
1.9 %	0.7 e 1.4 %	ψ_1^-	4.19	0.23	0.0001
	1.9 %	ϕ_1	10.79	0.10	0.0001
2.4 %	0.7; 1.4 e 1.9 %	ψ_2^-	3.11	0.21	0.0001
	2.4 %	ϕ_2	10.22	0.11	0.0001
1.9% e 2.4%		ψ^+	-0.47	0.18	0.0090
3.4 %	0.7; 1.4; 1.9 e 2.4 %	ψ_3^-	2.03	0.13	0.0001
	3.4 %	ϕ_3	12.22	0.12	0.0001
1.9 %	0.7 e 1.4 %	δ_1^-	0.45	0.06	0.0001
2.4 %	0.7; 1.4 e 1.9 %	δ_2^-	0.35	0.04	0.0001
3.4 %	0.7; 1.4; 1.9 e 2.4 %	δ_3^-	0.12	0.02	0.0001
Todos	Todas	α	0.98	0.02	0.0001
Parâmetros de covariância					
1.9 %	0.7 e 1.4 %	ϕ_{11}^2	0.28		
	2.4 e 3.4 %	ϕ_{21}^2	0.51		
2.4 %	0.7 a 1.9 %	ϕ_{12}^2	0.25		
	3.4 %	ϕ_{22}^2	1.01		
3.4 %	0.7 a 2.4 %	ϕ_{13}^2	0.20		
		σ^2	0.30		

p: nível descritivo associado ao teste da hipótese de o parâmetro ser nulo.

A redução do modelo (4.1) para o modelo (4.6) não se justifica em termos de um teste de Wald ($p=0.0075$), ou seja, a estrutura que relaciona a amplitude do PAC média com diferença de salinidade através de duas semi-retas, uma para diferenças negativas de salinidade e outra para positivas, não parece ser adequada sob um modelo de efeitos aleatórios com o padrão de dependência intraunidades amostrais dado em (4.1).

Buscando verificar se a amplitude média corrigida pela medida controle pode relacionar-se com a diferença de salinidade através de um polinômio com grau maior que 1, propôs-se um modelo da forma (4.5) com:

$$\mathbf{X}_{ij} = \left(\mathbf{1}_{t_{ij}} \quad \Delta_{ij} \quad \Delta_{ij}^2 \quad \Delta_{ij}^3 \quad \mathbf{W}_{ij} \quad \text{diag}(\mathbf{W}_{ij} \Delta_{ij}^-) \right), \quad (4.7)$$

$$\beta_j = \left(\beta_{0j} \quad \beta_{1j} \quad \beta_{2j} \quad \beta_{3j} \quad \alpha \quad \delta_j^- \right)^t$$

e Σ_{ij} definido em (4.1).

Sendo:

$$\Delta_{ij} = (\Delta_{ij1} \quad \dots \quad \Delta_{ijt_{ij}})^t, \quad \Delta_{ij}^2 = (\Delta_{ij1}^2 \quad \dots \quad \Delta_{ijt_{ij}}^2)^t, \quad \Delta_{ij}^3 = (\Delta_{ij1}^3 \quad \dots \quad \Delta_{ijt_{ij}}^3)^t,$$

Δ_{ijk} o valor da diferença entre a k-ésima salinidade teste e a salinidade de aclimação do i-ésimo molusco do j-ésimo grupo e $\beta_{0j}, \beta_{1j}, \beta_{2j}, \beta_{3j}$, respectivamente, o intercepto, coeficientes linear, quadrático e cúbico da relação entre a amplitude média corrigida pelas medidas controles e diferença de salinidade.

O modelo parece ajustar-se bem aos dados, conforme se pode comprovar através do teste de Wald ($p=0.6720$). O valor da estatística de Akaike para esse modelo é -255.58 (em contraponto a -261.97, para o modelo (4.1)) e as estimativas dos parâmetros correspondentes encontram-se na Tabela 4.9.

Tabela 4.9. Estimativas de máxima verossimilhança restrita e respectivos erros padrão dos parâmetros do modelo (4.7).

Salinidade de Aclimação	Salinidade Teste	Parâmetro	Estimativa	Erro Padrão	p
1.9 %	Todas	β_{01}	10.38	0.09	0.0001
	0.7; 1.4; 2.4 e 3.4 %	β_{11}	1.95	0.25	0.0001
	0.7; 1.4; 2.4 e 3.4 %	β_{21}	-1.58	0.13	0.0001
	0.7; 1.4; 2.4 e 3.4 %	β_{31}	0.08	0.14	0.5781
2.4 %	Todas	β_{02}	10.09	0.12	0.0001
	0.7; 1.4; 1.9 e 3.4 %	β_{12}	1.94	0.24	0.0001
	0.7; 1.4; 1.9 e 3.4 %	β_{22}	-1.69	0.22	0.0001
	0.7; 1.4; 1.9 e 3.4 %	β_{32}	-0.59	0.16	0.0004
3.4 %	Todas	β_{03}	12.05	0.12	0.0001
	0.7; 1.4; 1.9 e 2.4 %	β_{13}	2.64	0.59	0.0001
	0.7; 1.4; 1.9 e 2.4 %	β_{23}	1.07	0.59	0.0721
	0.7; 1.4; 1.9 e 2.4 %	β_{33}	0.33	0.14	0.0219
1.9 %	0.7 e 1.4 %	δ_1^-	0.49	0.07	0.0001
2.4 %	0.7; 1.4 e 1.9 %	δ_2^-	0.39	0.05	0.0001
3.4 %	0.7; 1.4; 1.9 e 2.4 %	δ_3^-	0.14	0.02	0.0001
Todos	Todas	α	1.02	0.02	0.0001

p: nível descritivo associado ao teste da hipótese de o parâmetro ser nulo.

Tabela 4.9. (continuação).

Salinidade de Aclimação	Salinidade Teste	Parâmetro	Estimativa
Parâmetros de covariância			
1.9 %	0.7 e 1.4 %	φ_{11}^2	0.34
	2.4 e 3.4 %	φ_{21}^2	0.56
2.4 %	0.7; a 1.9 %	φ_{12}^2	0.27
	3.4 %	φ_{22}^2	1.17
3.4 %	0.7; a 2.4 %	φ_{13}^2	0.18
		σ^2	0.26

p: nível descritivo associado ao teste da hipótese de o parâmetro ser nulo.

Observando as estimativas dos parâmetros do modelo (4.7) na Tabela 4.9, pode-se ainda tentar uma redução adicional, excluindo-se o coeficiente cúbico do grupo com aclimação 1.9 % (β_{31}) e o quadrático do grupo com aclimação 3.4 % (β_{23}), $p = 0.5781$ e $p = 0.0721$, respectivamente. Esses resultados sugerem um modelo da forma (4.5) com:

$$\begin{aligned}
 \mathbf{X}_{i1} &= (\mathbf{1} \quad \Delta_{i1} \quad \Delta_{i1}^2 \quad \mathbf{W}_{i1} \quad \text{diag}(\mathbf{W}_{i1} \Delta_{i1}^-)), \\
 \mathbf{X}_{i2} &= (\mathbf{1} \quad \Delta_{i2} \quad \Delta_{i2}^2 \quad \Delta_{i2}^3 \quad \mathbf{W}_{i2} \quad \text{diag}(\mathbf{W}_{i2} \Delta_{i2}^-)), \\
 \mathbf{X}_{i3} &= (\mathbf{1} \quad \Delta_{i3} \quad \Delta_{i3}^3 \quad \mathbf{W}_{i3} \quad \text{diag}(\mathbf{W}_{i3} \Delta_{i3}^-)), \\
 \beta_1 &= (\beta_0 \quad \beta_{11} \quad \beta_{21} \quad \alpha \quad \delta_1^-)^t, \\
 \beta_2 &= (\beta_0 \quad \beta_{12} \quad \beta_{22} \quad \beta_{32} \quad \alpha \quad \delta_2^-)^t, \\
 \beta_3 &= (\beta_0 \quad \beta_{13} \quad \beta_{33} \quad \alpha \quad \delta_3^-)^t
 \end{aligned} \tag{4.8}$$

e

Σ_{ij} como definido em (4.1).

Aceitando o ajuste do modelo (4.8), quando comparado com o modelo (4.1) através do teste de Wald ($p=0.4963$), e não havendo indicações de outras possíveis reduções, decidiu-se adotá-lo como modelo final. As estimativas de seus parâmetros estão apresentadas na Tabela 4.10.

Tabela 4.10. Estimativas de máxima verossimilhança restrita e respectivos erros padrão dos parâmetros do modelo (4.8).

Salinidade de Aclimação	Salinidade Teste	Parâmetro	Estimativa	Erro Padrão	p
1.9 %	Todas	β_{01}	10.37	0.09	0.0001
	0.7; 1.4; 2.4 e 3.4 %	β_{11}	2.06	0.15	0.0001
	0.7; 1.4; 2.4 e 3.4 %	β_{21}	-1.54	0.11	0.0001
2.4 %	Todas	β_{02}	10.09	0.12	0.0001
	0.7; 1.4; 1.9 e 3.4 %	β_{12}	1.94	0.24	0.0001
	0.7; 1.4; 1.9 e 3.4 %	β_{22}	-1.69	0.22	0.0001
	0.7; 1.4; 1.9 e 3.4 %	β_{32}	-0.59	0.16	0.0004
3.4 %	Todas	β_{03}	12.01	0.12	0.0001
	0.7; 1.4; 1.9 e 2.4 %	β_{13}	1.61	0.17	0.0001
	0.7; 1.4; 1.9 e 2.4 %	β_{33}	0.08	0.02	0.0004
1.9 %	0.7 e 1.4 %	δ_1^-	0.49	0.06	0.0001
2.4 %	0.7; 1.4 e 1.9 %	δ_2^-	0.39	0.05	0.0001
3.4 %	0.7; 1.4; 1.9 e 2.4 %	δ_3^-	0.15	0.02	0.0001
Todos	$\geq$ aclimação	α	1.02	0.02	0.0001
Parâmetros de covariância					
1.9 %	0.7 e 1.4 %	φ_{11}^2	0.32		
	2.4 e 3.4 %	φ_{21}^2	0.56		
2.4 %	0.7; a 1.9 %	φ_{12}^2	0.26		
	3.4 %	φ_{22}^2	1.18		
3.4 %	0.7; a 2.4 %	φ_{13}^2	0.19		
		σ^2	0.26		

p: nível descritivo do teste da hipótese de o parâmetro ser nulo.

Com base nas estimativas do modelo (4.8), pode-se dizer que a amplitude média corrigida pela medida controle não se relaciona com diferença de salinidade igualmente nos 3 grupos. Apesar disso, de uma maneira geral, observa-se que a amplitude média do PAC decresce com o aumento da magnitude da diferença de salinidade. Para os grupos com salinidade de aclimação 2.4 % e 3.4 % esta relação é cúbica, enquanto que no grupo 1.9 % ela é quadrática. A Figura 4.6 ilustra a relação entre a amplitude média do PAC e diferença de salinidade ajustada no modelo (4.8).

Figura 4.6. Amplitude média ajustada pelas medidas controles estimada no modelo saturado (4.8).

Uma discussão mais detalhada sobre a adequação da estratégia de análise dos dados apresentada neste capítulo, em contraposição com a apresentada anteriormente, será feita no capítulo 5.

Capítulo 5

Conclusão

O objetivo deste capítulo é contrastar os modelos que consideram independência entre as observações, ajustados no capítulo 3, com os modelos de efeitos aleatórios, ajustados no capítulo 4.

Inicialmente, compararam-se os modelos saturados relativamente aos parâmetros de regressão. As estimativas (e erros padrão) dos parâmetros de ambos os modelos estão apresentadas nas Tabelas 3.1 e 4.7. ^{p. 25} ^{p. 52} As diferenças entre os valores das estimativas dos parâmetros dos dois modelos são muito pequenas, sugerindo que a modelagem para os parâmetros de regressão parece ser robusta em relação à estrutura da matriz de covariâncias intraunidades amostrais. De um modo geral, os erros padrão dos estimadores dos parâmetros de regressão, sob o modelo de independência, são maiores que os correspondentes erros padrão sob o modelo de efeitos aleatórios (dos 30 parâmetros, 23 apresentam erros padrão menores sob o modelo de efeitos aleatórios). Além disso, o traço da matriz de covariâncias dos

estimadores dos parâmetros de regressão é igual a 0.9729 para o modelo de independência e 0.9293 para o de efeitos aleatórios, representando uma diferença de 5%. Pelo critério de Akaike, o modelo saturado de efeitos aleatórios se ajusta melhor aos dados do que o de independência (-261.97 e -288.70, respectivamente).

A estrutura final adotada para os parâmetros de regressão para o modelo de independência, dada em (3.5), difere daquela apresentada em (4.8) para o modelo de efeitos aleatórios. Em ambos, a amplitude média do PAC decresce com o aumento da magnitude da diferença de salinidade. Porém, no modelo com independência entre as observações intraunidades amostrais, este decréscimo se dá de maneira linear. Ao passo que no modelo de efeitos aleatórios, esta relação é expressa por um polinômio de grau 3 (ou 2, no grupo com salinidade de aclimação igual a 1.9%). A Figura 5.1 representa as amplitudes médias do PAC ajustadas nesses dois modelos. Uma das razões para essa diferença pode estar no fato de que os erros padrão dos estimadores sob modelo de efeitos aleatórios são menores que sob modelo de independência, rejeitando, neste caso, o ajuste do modelo que relaciona linearmente a amplitude média do PAC e diferença de salinidade. A análise gráfica dos resíduos desses modelos está apresentada nas Figuras 5.2 e 5.3. Observe que, para ambos os modelos, os resíduos parecem estar aleatoriamente distribuídos em torno do zero e parecem provir de uma distribuição normal. Por outro lado, no modelo de efeitos aleatórios esses resíduos têm amplitude de variação menor do que no modelo de independência, indicando um melhor ajuste dos dados no primeiro caso.

Figura 5.1. Amplitude média ajustada pelas medidas controles estimada nos modelos (3.5) e (4.8).

Figura 5.2. Análise gráfica dos resíduos do modelo (3.5).

a) Histograma do resíduo do modelo (3.5).

b) Gráfico de probabilidade normal do resíduo do modelo (3.5).

c) Gráfico de dispersão do resíduo do modelo (3.5) em função da amplitude média do PAC ajustada.

Figura 5.3. Análise gráfica dos resíduos do modelo (4.8).

a) Histograma do resíduo do modelo (4.8).

b) Gráfico de probabilidade normal do resíduo do modelo (4.8).

c) Gráfico de dispersão do resíduo do modelo (4.8) em função da amplitude média do PAC ajustada.

É importante ressaltar que, mesmo que a estrutura reduzida dos parâmetros de regressão dada em (3.5) não tenha se ajustado para o modelo de efeitos aleatórios, os erros padrão das estimativas dos parâmetros de regressão diminuem quando comparados aos do modelo de independência (Tabelas 3.5 e 4.8). Neste caso, o traço da matriz de covariâncias dos estimadores dos parâmetros de regressão é igual a 0.2092 para o modelo de independência e 0.1888 para o de efeitos aleatórios, representando uma diferença de 11%.

Em resumo, o modelo de efeitos aleatórios parece apresentar um ajuste mais fino do que o de independência, gerando estimativas mais precisas e sugerindo a adoção de um modelo mais complexo. Lembrando que os modelos finais, (3.5) e (4.8), obtidos sob os dois enfoques permitem concluir que a amplitude média do PAC decresce com o aumento da magnitude da diferença de salinidade, a escolha da forma de modelagem depende de considerações biológicas sobre o fenômeno em estudo, que fogem ao objetivo deste trabalho. Se o objetivo for apenas explorar o comportamento da amplitude média do PAC em função da diferença de salinidade, pode-se optar pela estrutura mais simples. Se, por outro lado, o objetivo for prever a amplitude média do PAC em função da salinidade, a estrutura mais complexa pode ser mais conveniente.

Os cálculos estatísticos deste trabalho foram realizados no procedimento MIXED do SAS (SAS Institute Inc., 1996), versão 6.12. Maiores explicações sobre a estrutura do arquivo de dados e sobre o programa do SAS utilizados para a análise estatística aqui realizada encontram-se em apêndice.

Apêndice

Aspectos Computacionais

Este apêndice contém o programa e a forma de entrada dos dados para o ajuste do modelo (4.1) utilizados no procedimento MIXED do software SAS. Os demais modelos considerados neste trabalho podem ser ajustados a partir de pequenas modificações deste exemplo.

- **Arquivo de Dados**

Todas as observações da variável resposta devem estar contidas em uma mesma coluna, independentemente de terem sido coletadas em um mesmo molusco ou não. A planilha dos dados deve também conter outras colunas que identificam o molusco de onde cada observação foi obtida, a condição de avaliação e o valor da covariável associada a cada observação. A Tabela A.1 exemplifica a estrutura do arquivo de dados utilizado para a análise estatística elaborada neste trabalho. Observe que as observações perdidas (*missing values*)

estão representadas no arquivo de dados por um ponto, que é o símbolo utilizado pelo programa SAS para tais situações. Além disso, todas as condições de avaliação estão especificadas na mesma ordem para todos os moluscos, mesmo nos casos onde não houve observações.

Tabela A.1. Estrutura do arquivo de dados.

Rótulo da variável no arquivo tipo SAS				
MOLUSCO	GRUPO	SALTEST	AMPL	CONT
Molusco	Salinidade de Aclimação (Grupo)	Salinidade Teste (Condição de Avaliação)	Amplitude do PAC (Resposta)	Medida Controle (Covariável)
4	3.4	3.4	13.0	.
4	3.4	2.4	11.0	11.0
4	3.4	1.9	.	.
4	3.4	1.4	9.2	10.0
4	3.4	0.7	4.4	6.6
11	3.4	3.4	17.0	.
11	3.4	2.4	14.0	16.0
11	3.4	1.9	.	.
11	3.4	1.4	.	.
11	3.4	0.7	5.2	12.4
etc				
8	2.4	3.4	10.0	14.0
8	2.4	2.4	14.0	.
8	2.4	1.9	.	.
8	2.4	1.4	.	.
8	2.4	0.7	3.0	5.8
9	2.4	3.4	14.0	15.0
9	2.4	2.4	15.0	.
9	2.4	1.9	.	.
9	2.4	1.4	.	.
9	2.4	0.7	5.0	12.0
etc				
20	1.9	3.4	.	.
20	1.9	2.4	.	.
20	1.9	1.9	8.4	.
20	1.9	1.4	5.0	6.4
20	1.9	0.7	.	.
21	1.9	3.4	13.0	15.0
21	1.9	2.4	12.0	12.0
21	1.9	1.9	15.0	.
21	1.9	1.4	7.2	9.2
21	1.9	0.7	2.8	6.0
etc				

• **Programa SAS para o modelo (4.1)**

Programa	Comentários
data DADOS;	O comando “data” especifica qual o arquivo de dados que será analisado.
DIFSAL=GRUPO-SALTEST;	Cria a variável diferença entre as salinidades de aclimação e as salinidades testes.
if DIFSAL=0 then CONTH=AMPL;	Impõe que, na salinidade de aclimação, o valor da medida controle da amplitude do PAC é igual à resposta.
if DIFSAL<0 then SINAL='<=0'; else SINAL='>=0';	Cria a variável SINAL que identifica quais são as diferenças de salinidades positivas e negativas.
if GRUPO=1.9 then COV=CONTH-10.53; if GRUPO=2.4 then COV=CONTH-10.13; if GRUPO=3.4 then COV=CONTH-12.09;	Centra a covariável, separadamente para cada grupo. A covariável centrada recebe o nome de COV.
proc MIXED method=REML;	Especifica o método de estimação MVR (máxima verossimilhança restrita) utilizado no ajuste do modelo.
class MOLUSCO GRUPO SALTEST SINAL;	Define que as variáveis MOLUSCO, GRUPO, SALTEST e SINAL são qualitativas.
model AMPL = GRUPO*SALTEST COV*GRUPO*SALTEST / noint s;	Especifica o modelo: AMPL como variável resposta; gera 15 parâmetros de regressão a partir da combinação das variáveis GRUPO e SALTEST (representados pelos μ 's em (3.1)); e gera mais 15 parâmetros de regressão a partir da combinação das variáveis COV, GRUPO e SALTEST (representados pelos ξ 's em (3.1)). Observe que a variável COV entra no modelo como quantitativa, pois ela não foi incluída no comando “class”. A opção “noint” impõe que o modelo não terá um intercepto. A opção “s” mostra os valores estimados e os erros padrão do parâmetros de regressão do modelo.

Programa	Comentários
<pre>random DIF SAL* SINAL* GRUPO / type=un(1) subject=MOLUSCO g v=1,20,38;</pre>	<p>Especifica os efeitos aleatórios do modelo: cria 5 efeitos aleatórios para diferenças de salinidades, uma efeito para cada uma das combinações das categorias das variáveis SINAL e GRUPO. A opção “type=un(1)” especifica que os efeitos aleatórios do modelo têm variâncias diferentes e são independentes (define as matrizes G’s em (4.1)). A opção “subject=MOLUSCO” produz uma estrutura diagonal em blocos para as matrizes G’s e Z’s em (4.1), gerando um bloco para cada MOLUSCO. A opção “g” imprime o bloco da matriz G correspondente ao primeiro MOLUSCO do arquivo SAS. Já a opção “v=1,20,38” imprime os blocos das matrizes Σ’s correspondentes ao 1º, ao 20º e ao 38º MOLUSCO do arquivo SAS.</p>
<pre>repeated / type=simple subject=MOLUSCO r;</pre>	<p>Este comando especifica a matriz R_{ij} do modelo; no caso, “type=simple” corresponde a $R_{ij} = \sigma^2 I_i$. A opção “subject=MOLUSCO” produz uma estrutura diagonal em blocos para a matriz R em (4.1), gerando um bloco para cada MOLUSCO. A opção “r” imprime o bloco da matriz R correspondente ao primeiro MOLUSCO do arquivo SAS.</p>
<pre>run;</pre>	

Bibliografia

Box, G. E. P. and Jenkins, G. M. (1976). **Time Series Analysis: Forecasting and Control** (revised edition). San Francisco: Holden-Day.

Diggle, P. J., Liang, K. and Zeger, S. L. (1994). **Analysis of Longitudinal Data**. Oxford: Oxford University Press.

Draper, N. R. and Smith, H. (1981). **Applied Regression Analysis**. (2nd edition). New York: John Wiley & Sons.

Frimm, C. C., Pereira, V. F. A., Rodrigues, A. C. T., Medeiros, C. C. J., Moraes, A. V. and Krieger, E. M. (1998). Left ventricular dysfunction and coronary reserve impairment in hypertension: a combined transthoracic and transesophageal study. **European Heart Journal**, **19 (Supplement)**, 419.

- Grady J. J. and Helms, R. W. (1995). Model selection techniques for the covariance matrix for incomplete longitudinal data. **Statistics in Medicine**, **14**, 1397-1416.
- Jennrich, I. J. and Schluchter, M. D. (1986). Unbalanced repeated measures models with structured covariance matrices. **Biometrics**, **42**, 805-820.
- Jones, R. H. (1993). **Longitudinal Data With Serial Correlation: A State-Space Approach**. London: Chapman & Hall.
- o Laird, N. M. and Ware, J. H. (1982). Random effects models for longitudinal data. **Biometrics**, **38**, 963-974.
- Lindsey, J. K. (1993). **Models For Repeated Measurements**. Oxford: Oxford University Press.
- Patel, H. I. (1986). Analysis of repeated measures designs with changing covariates in clinical trials. **Biometrika**, **73**, 707-715.
- Rubin, D. B. (1976). Inference and missing data. **Biometrika**, **63**, 581-592.
- SAS Institute Inc. (1996). **SAS/STAT[®] Software: Changes and Enhancements through Release 6.11**. Cary: SAS Institute Inc.
- Sen, P. K. and Singer, J. M. (1993). **Large Sample Methods in Statistics: An Introduction with Applications**. New York: Chapman and Hall.
- Singer, J. M. and Andrade, D. F. (1999). Analysis of longitudinal data. In **Handbook of Statistics, Volume 17: Bio-Environmental and Public Health Statistics**. C. R. Rao and P. K. Sen (eds). Amsterdam: North Holland.

Singer, J. M. e Mauro, E. S. R. (1987). **Estudo de Respostas Eletrofisiológicas do Nervo Conectivo Cérebro-Visceral de *Perna perna* Exposto a Diferentes Condições de “Stress” Osmótico**. Relatório de Análise Estatística. São Paulo: Centro de Estatística Aplicada, Departamento de Estatística, Instituto de Matemática e Estatística, Universidade de São Paulo.

Stram, D. O. and Lee, J. W. (1994). Variance components testing in the longitudinal mixed effects model. **Biometrics**, **50**, 1171-1177.

Stucchi-Zucchi, A. (1988). **Estudo de Algumas Características do Potencial de Ação Composto Evocado do Conectivo Cérebro-Visceral de *Perna perna* L. 1758 (Mollusca, Bivalva) em Diferentes Condições de Estresse Osmótico**. Dissertação de Mestrado. São Paulo: Departamento de Fisiologia Geral, Instituto de Biociências, Universidade de São Paulo.

Suyama, E. and Singer, J.M. (1996). **Identification of Random Effects Models for Longitudinal Data**. Technical Report. São Paulo: Departamento de Estatística, Instituto de Matemática e Estatística, Universidade de São Paulo.

Wolfinger, R., Tobias, R. and Sall, J. (1994). Computing Gaussian likelihoods and their derivatives for general linear mixed models. **SIAM J. of Sci. Comput.**, **15** n° 6, 1294-1310.