

SELEÇÃO DE
ESTRUTURAS DE
COVARIÂNCIA PARA DADOS
COM MEDIDAS REPETIDAS

Francisco Marcelo Monteiro da Rocha

DISSERTAÇÃO APRESENTADA
AO
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA
DA
UNIVERSIDADE DE SÃO PAULO
PARA OBTENÇÃO DO GRAU
DE
MESTRE EM ESTATÍSTICA

Área de Concentração: **Estatística**

Orientador: **Prof. Dr. Julio da Motta Singer**

Este trabalho foi financiado pelo CNPq.

São Paulo, abril de 2004

SELEÇÃO DE ESTRUTURAS DE COVARIÂNCIA PARA DADOS COM MEDIDAS REPETIDAS

Este exemplar corresponde à redação final da dissertação devidamente corrigida e defendida por Francisco Marcelo Monteiro da Rocha e aprovada pela banca examinadora.

Banca examinadora:

- Prof. Dr. Julio da Motta Singer (orientador) - IME-USP
- Prof. Dr. Dalton F. Andrade - UFSC
- Profa. Dra. Solange Andreoni - UNIFESP

São Paulo, abril de 2004.

Agradecimentos

Agradeço a Deus por ter concluído este trabalho.

Agradeço à minha esposa Lane, pela compreensão durante a elaboração deste trabalho, pelas conversas e pela luz inspiradora que ela representa em minha vida.

Agradeço em especial aos meus pais Nunes e Tarcila e à minha irmã Sharliane pelo incentivo e ajuda em todos os momentos. Obrigado aos meus tios João e Ivonilde que me acolheram em sua casa na difícil chegada em São Paulo.

Agradeço também aos profs. João Maurício de Araújo Mota e Dalton Francisco de Andrade pelo incentivo constante e amizade e ao prof. Julio da Motta Singer pela orientação, amizade e paciência.

Resumo

Propomos um conjunto de técnicas gráficas e analíticas para o diagnóstico da matriz de covariâncias intra-unidades amostrais em estudos com medidas repetidas. Essas técnicas compreendem a análise das matrizes de covariâncias e correlações amostrais, dos gráficos de dispersão múltipla, dos gráficos de perfis médios e de perfis individuais, dos gráficos de perfis das linhas das matrizes de correlações amostrais em função da defasagem e do variograma amostral das observações padronizadas. A aplicação dessas técnicas é ilustrada por meio de exemplos práticos.

Abstract

We propose a set of graphical and analytical techniques for the diagnostic of the within-subject covariance matrix in repeated measures studies. They, comprise the analysis of sample covariance and correlation matrices, of multiple dispersion plots, of mean profile and individual profile plots, of the sample correlation matrices profile plots and, and of the sample variogram. We illustrate the use of such techniques with practical examples.

Conteúdo

1	Introdução	1
1.1	Modelo Estatístico	17
1.2	Estruturas de Covariâncias	18
1.3	Inferência Estatística	28
1.4	Teste de Hipótese para Componentes de Variância	33
1.5	Critérios de Informação	35
2	Modelagem da Estrutura de Covariância	36
2.1	Introdução	36
2.2	Técnicas Gerais	37
2.3	Técnicas Específicas	48
2.3.1	Estudos Não Longitudinais	48
2.3.2	Estudos Longitudinais Completos ou Balanceados	50
2.3.3	Estudos Longitudinais Incompletos ou Não Balanceados	55
2.4	Estratégias de Análise	58
3	Exemplos de Modelagem da Estrutura de Covariância	60
3.1	Introdução	60
3.2	Modelagem da Estrutura de Covariância para Estudos Não Longitudinais Completos	61
3.3	Estudos Não Longitudinais Incompletos	64
3.4	Estudos Longitudinais Completos	77

3.4.1	Exemplo 1.1 - Concentração de bilirrubina	77
3.4.2	Exemplo 1.3 - Microdureza de esmalte dentário	88
3.5	Estudos Longitudinais Incompletos	95
4	Conclusões e Sugestões	103
	Referências Bibliográficas	105

Capítulo 1

Introdução

Experimentos com medidas repetidas apresentam uma ou mais variáveis respostas (ou características de interesse) avaliadas repetidamente em uma mesma unidade amostral. As medidas podem ser realizadas em diferentes instantes ou condições de avaliação ou mesmo em circunstâncias semelhantes, por exemplo, quando as observações em cada unidade amostral são obtidas em triplicata.

Existem basicamente duas formas de obtenção de medidas repetidas. A primeira corresponde à atribuição aleatória da ordem com que as medidas sob diferentes condições de avaliação são realizadas na mesma unidade amostral. A segunda ocorre quando a ordem com que as medidas realizadas sob as diferentes condições de avaliação estão dispostas numa escala ordenada, ou seja, são obtidas seqüencialmente e, neste caso, os estudos são denominados longitudinais. A escala utilizada pode ser, por exemplo, o tempo, dosagens crescentes de uma determinada droga ou a distância com relação a uma origem pré-determinada, como em estudos planejados para avaliar a relação entre a intensidade de poluição e a distância de uma ou mais fontes poluidoras. Mais detalhes sobre estudos longitudinais podem ser encontrados em Diggle et al. (2002), Singer e Andrade (2000), Crowder e Hand (1996) ou Jones (1993).

Estudos com medidas repetidas têm a vantagem de requerer um número menor de unidades amostrais relativamente a estudos do tipo transversal (“cross sectional”), que consistem na observação da variável resposta uma única vez em cada unidade amostral.

Além disso, diminuem a variabilidade decorrente de diferenças individuais e permitem avaliar mudanças que ocorrem dentro e entre as unidades amostrais com mais eficiência. As principais desvantagens de estudos com medidas repetidas são duas: a primeira é que a análise é mais complicada por requerer que as correlações das medidas realizadas no mesmo indivíduo sejam modeladas. A segunda desvantagem está relacionada com a presença de dados incompletos ou omissos (“missing data”), o que na prática ocorre com frequência.

Estudos com medidas repetidas são classificados como regulares em relação às condições de avaliação quando o intervalo entre duas medidas consecutivas é constante; em caso contrário, são chamados de irregulares. Quanto ao planejamento, eles podem ser classificados como balanceados em relação às condições de avaliação se as medidas forem obtidas nos mesmos instantes de avaliação em todas as unidades amostrais; em caso contrário, são ditos não balanceados. Em relação às condições de avaliação, esses estudos são completos se não houver observações perdidas (dados omissos) e incompletos, se existirem. Estudos com dados incompletos ocorrem quando, por alguma razão, as observações de uma ou mais unidades de investigação, não são obtidas sob alguma condição de avaliação. Dados incompletos dificultam o processo de estimação, os procedimentos de testes de hipóteses e podem viciar os resultados. Vonesh e Chinchilli (1997, Caps. 1 e 6), entre outros, discutem os efeitos de dados incompletos em estudos com medidas repetidas.

Dados omissos são ditos “omissos completamente ao acaso” (Missing completely at random “MCAR”) quando ocorrem por razões totalmente aleatórias, ou seja, quando a probabilidade de omissão da resposta for independente dos valores observados e não observados. Dados omissos com esta característica não causam problemas sérios à análise estatística e a maioria dos procedimentos padrão de análise são válidos nessas circunstâncias. Dados omissos também podem ser classificados como “omissos ao acaso” (Missing at random “MAR”) quando a probabilidade de a resposta estar faltando depender somente dos valores observados. Em estudos com medidas repetidas, onde os dados omissos são do tipo MAR, procedimentos de estimação que não sejam baseados no método de máxima verossilhança podem produzir inferências inválidas. Dados omissos são considerados “in-

formativos” quando a probabilidade de uma omissão depender de características não observadas da unidade amostral. Em estudos com medidas repetidas, em que os dados omissos são informativos, os procedimentos de análise deverão incluir também a modelagem do processo segundo o qual eles ocorrem a fim de permitir inferências válidas. Mais detalhes sobre inferências na presença de dados omissos podem ser encontradas em Rubin (1976), Little (1988), Diggle (1989), Diggle e Kenward (1994), Little (1995), Vonesh e Chinchilli (1997), Singer e Andrade (2000), Diggle et al. (2002), Little e Rubin (2002) e Davis (2002).

Exemplo 1.1. Concentração de bilirrubina.

Os dados da Tabela 1.1 correspondem a um estudo realizado na Escola Paulista de Medicina (UNIFESP), em que foi medida a concentração de bilirrubina (μ mol/L) em 89 recém-nascidos a termo saudáveis em aleitamento materno durante 1, 2, 3, 4, 5, 6, 8, 10 e 12 dias após o nascimento. O objetivo era explicar a variação da concentração de bilirrubina em função da idade (Draque et al., 2001).

Tabela 1.1: Concentração de bilirrubina (μ mol/L) em recém-nascidos (RN) saudáveis em aleitamento materno.

RN	Dias após o nascimento								
	1	2	3	4	5	6	8	10	12
1	2.7	0.4	0.0	0.5	0.6	0.0	0.0	0.5	0.8
2	4.5	5.5	3.9	2.7	2.9	2.0	1.5	1.3	1.7
3	7.0	9.2	13.1	12.1	12.3	10.3	8.5	6.2	3.8
4	4.4	6.1	8.8	7.4	6.8	6.2	5.5	4.9	3.3
5	4.0	7.7	7.5	10.3	10.6	9.3	7.3	5.8	6.4
6	6.5	7.4	5.3	4.5	2.9	3.3	3.0	3.2	3.4
7	3.2	2.3	0.9	0.2	0.0	0.0	0.6	0.5	0.0
8	6.2	7.7	5.9	4.1	4.5	2.4	2.0	1.5	0.9
9	3.0	4.5	8.0	8.1	8.6	7.4	6.8	5.5	4.9
10	2.0	2.4	2.9	2.3	1.9	2.1	2.0	1.2	0.1
11	6.2	8.1	10.9	8.0	5.4	6.4	5.0	3.5	3.4
12	4.6	6.3	5.8	4.6	4.6	3.0	3.0	3.0	2.5
13	4.5	7.5	8.0	10.6	10.9	11.0	8.1	7.2	4.3
14	6.6	9.1	7.7	8.4	8.1	7.6	3.9	3.5	1.9
15	2.8	4.3	3.3	2.3	1.0	1.6	1.6	1.5	1.4
16	4.0	5.0	4.2	3.9	2.1	2.2	2.1	2.2	2.1
17	2.8	3.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0
18	6.0	8.1	3.8	2.7	1.7	1.7	1.4	0.0	1.1
19	2.4	4.5	4.7	5.3	4.5	3.5	3.3	3.0	4.0
20	5.4	5.7	3.7	3.0	2.0	0.0	0.0	0.0	0.0
21	3.1	4.3	3.8	4.3	2.8	2.1	1.6	1.3	1.8
22	5.4	6.9	6.9	5.5	6.2	6.2	3.6	3.4	3.8
23	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.5	0.5
24	4.6	4.1	5.1	3.5	4.0	3.4	1.9	2.9	1.3
25	5.9	7.9	10.1	9.8	10.7	11.2	10.6	9.7	9.2
26	5.5	6.4	8.6	8.1	6.4	6.3	6.5	7.0	7.2
27	2.4	5.3	2.4	1.7	1.6	0.3	0.1	0.0	0.0
28	2.2	1.7	1.2	1.3	0.3	1.5	1.0	0.2	0.2
29	2.9	2.5	1.2	1.1	1.3	1.3	1.3	0.0	0.0
30	5.5	7.8	9.2	8.3	8.2	6.5	5.4	5.4	5.6
31	3.5	4.6	4.4	4.5	3.2	3.0	1.7	1.0	0.0
32	4.2	4.5	1.8	0.7	0.2	1.1	0.6	0.5	0.1
33	6.3	7.8	9.9	6.1	5.3	3.0	3.7	3.5	4.0
34	4.6	1.1	0.7	1.8	1.1	1.0	1.0	1.2	1.6
35	3.6	7.3	7.2	6.0	4.8	4.4	3.6	3.4	3.1
36	2.6	3.8	3.5	1.9	1.2	0.7	0.7	0.9	1.4
37	2.9	4.4	7.3	8.3	8.7	7.6	8.0	7.8	6.8
38	5.0	9.2	9.6	12.3	9.8	12.2	8.6	8.2	7.4
39	4.5	7.0	8.7	10.5	10.3	7.4	7.9	7.5	6.8
40	6.6	7.7	10.4	9.3	9.8	11.7	10.8	8.8	9.0
41	4.9	9.7	11.2	11.0	8.6	7.7	5.5	4.8	5.0
42	2.3	2.2	2.0	2.9	2.7	2.0	1.4	1.1	0.2
43	0.6	0.2	0.0	0.2	0.0	0.0	0.0	0.4	0.0
44	3.5	5.2	7.4	9.4	7.2	7.0	4.8	4.0	2.1
45	5.0	8.9	8.8	9.4	10.3	10.4	11.8	11.3	10.6

Tabela 1.1 (Continuação): Concentração de bilirrubina (μ mol/L) em recém-nascidos (RN) saudáveis em aleitamento materno.

RN	Dias após o nascimento									
	1	2	3	4	5	6	8	10	12	
46	7.3	9.7	12.3	10.9	13.7	13.1	12.1	13.6	10.2	
47	7.0	9.0	9.6	10.0	10.3	10.1	8.7	7.9	6.8	
48	6.5	6.6	8.0	4.4	3.4	3.3	1.4	2.0	1.2	
49	5.4	7.3	7.8	8.6	9.1	9.0	9.7	6.3	6.0	
50	3.5	5.0	4.8	4.5	3.5	3.5	3.6	3.0	2.8	
51	0.0	0.8	0.0	0.0	0.2	0.1	0.4	0.5	0.6	
52	1.5	4.8	4.1	5.0	4.1	3.5	2.1	3.5	3.3	
53	3.9	2.9	1.5	1.0	1.5	1.1	0.7	0.5	0.1	
54	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
55	2.8	2.2	1.5	1.2	0.9	0.3	0.5	0.0	0.2	
56	6.2	6.8	4.6	3.5	2.4	3.1	2.7	3.0	3.5	
57	5.9	10.0	8.6	9.3	10.5	8.9	10.0	9.7	7.2	
58	8.8	9.9	7.5	11.0	7.3	5.0	5.1	4.0	2.9	
59	5.4	8.2	8.0	8.2	6.7	4.6	3.2	2.7	4.0	
60	5.1	6.0	2.7	1.0	0.8	0.5	0.4	0.5	0.5	
61	4.1	9.1	10.5	12.8	14.8	11.8	13.5	11.2	7.3	
62	8.0	10.2	9.4	7.6	5.7	2.9	3.0	3.4	2.0	
63	4.0	6.5	7.8	8.6	5.5	6.7	3.8	2.2	1.2	
64	6.3	5.6	5.5	4.6	3.2	0.9	1.4	1.2	1.2	
65	1.4	1.3	0.2	0.0	0.3	0.4	0.6	0.8	1.0	
66	3.6	7.3	7.3	8.0	4.2	5.0	2.8	2.5	2.0	
67	2.6	3.4	1.3	0.5	1.5	1.0	0.2	0.5	1.1	
68	6.7	11.4	14.6	14.1	13.6	9.6	9.0	7.8	3.9	
69	2.6	8.6	9.0	9.8	9.6	8.6	6.2	5.9	5.5	
70	4.0	6.1	5.6	3.7	3.1	3.9	2.3	2.3	2.2	
71	5.5	7.5	8.3	6.6	6.0	4.0	3.0	2.7	3.3	
72	6.3	7.5	9.0	8.3	8.7	9.2	9.2	7.6	7.4	
73	9.1	11.3	14.4	14.3	9.8	10.0	12.6	13.2	12.4	
74	5.9	7.3	10.6	8.7	8.8	7.1	6.9	4.2	4.0	
75	6.7	7.9	9.6	10.6	8.7	9.6	8.1	6.4	5.9	
76	1.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
77	6.7	4.0	2.0	1.2	1.5	1.2	0.3	0.9	0.4	
78	6.4	5.7	2.0	1.5	0.5	0.7	1.2	0.6	0.3	
79	1.2	0.9	0.2	0.6	0.0	0.0	0.0	0.6	0.4	
80	5.1	4.5	4.2	3.6	3.3	2.5	1.7	1.6	1.6	
81	2.5	1.5	2.4	1.1	1.9	2.0	1.6	0.8	0.6	
82	3.6	5.0	4.1	2.6	1.3	2.4	2.0	2.2	2.4	
83	4.3	5.0	6.6	3.1	2.2	2.8	2.5	1.7	1.2	
84	7.5	10.6	12.0	13.5	9.6	11.5	9.2	7.6	7.7	
85	6.5	8.1	11.0	10.7	10.0	9.7	7.9	7.4	5.2	
86	7.1	9.2	10.4	11.7	7.9	8.4	4.7	4.2	3.5	
87	3.6	6.6	9.9	8.8	11.5	12.0	12.0	11.3	9.7	
88	3.6	3.7	2.8	2.0	1.5	0.0	1.2	1.6	0.5	
89	2.6	1.4	1.3	1.0	1.6	0.4	0.0	0.3	0.6	

Exemplo 1.2. Evolução de características cardiovasculares em recém-nascidos pré-termo.

Os dados da Tabela 1.2 foram analisados pelo Centro de Estatística Aplicada do IME/USP da Universidade de São Paulo; um dos objetivos era estimar a curva média de crescimento do diâmetro diastólico da aorta por unidade de peso (AO) em mm/kg para recém-nascidos pré-termo pequenos para a idade gestacional (PIG), no período compreendido entre a 26^a à 40^a semanas pós concepção. Mais detalhes podem ser encontrados em Singer e Andrade. (2000).

Tabela 1.2: Diâmetro diastólico da aorta por unidade de peso (AO) em mm/kg, para recém nascidos pré-termo (RNPT) pequenos para a idade gestacional (PIG).

RNPT	Semanas após o nascimento														
	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
1							7.16	6.81		5.51			4.72		
2								7.08	7.98	7.74			6.50	5.56	
3						7.44	8.31	9.38	10.00	9.20			8.04		
4							7.68		6.56		5.51		4.61		
5									6.50					4.38	
6					7.63	8.63	9.26	7.98		6.64		5.00		4.72	
7					6.64	8.37		8.18	7.57		6.64				
8							7.05	6.27			6.09	5.92	5.73		4.8
9						8.54	8.39					4.92			
10				8.31	7.42		6.17			4.56		3.84			
11				9.76			9.06		7.30		5.35				
12				6.60			6.36	5.80	5.03					3.90	
13									7.80	7.90		7.05		6.04	
14				6.39											
15								5.85	5.87		4.69				
16								5.98		5.27					
17								6.20	5.83		4.77				
18		6.89	8.92		8.51	7.14		6.27		5.26		5.03		4.38	
19				7.82	8.50	8.26			6.97	6.18		5.09			
20									7.31		6.08		5.25		
21						7.05	6.79	6.41		4.94					
22							6.89								
23							8.57	9.42	9.03		8.21		7.52		7.3
24									4.93	5.64	5.70	5.47	5.28		4.6
25			7.27	7.53	6.88		6.05		6.13			4.92		4.49	
26						6.77	6.11	5.95		5.49		4.63			
27					8.20	7.50		6.84	6.15			5.09	4.56		
28				6.57	8.43		7.24	6.45			5.06		3.94		3.6
29	8.87	10.48	9.85	11.00	10.46	10.00	7.35			6.19	5.86				
30				9.71	9.08		8.00	7.66							
31		8.51													
32				10.92	10.75	9.37		7.96			5.76		4.91		

Exemplo 1.3. Microdureza de esmalte dentário.

Em um estudo realizado na Faculdade de Odontologia da Universidade de São Paulo, o principal interesse era avaliar o efeito da irradiação de "laser" na microdureza média do esmalte dentário de dentes expostos a desafios cariogênicos. Três blocos de esmalte dentário foram confeccionados a partir de cada um de 32 terceiros molares humanos extraídos por indicação ortodôntica, totalizando 96 blocos. Esses blocos, após limpos, foram examinados para verificação de defeitos e blocos com defeitos foram substituídos. Em seguida, os blocos sem defeitos foram limpos e subseqüentemente, alocados aleatoriamente a um de três tratamentos de forma balanceada (32 blocos por tratamento), a saber: Controle extra bucal, o Controle in situ e Laser. Os blocos submetidos aos tratamentos Laser e Controle in situ foram colocados em aparelhos intrabuciais devidamente fabricados para cada um de oito voluntários e usados durante 28 dias consecutivos para sofrerem a ação de agentes cariogênicos. Durante esse período, os blocos submetidos ao tratamento Controle extra bucal não sofreram ação de agentes cariogênicos. Após esse período, os blocos dos três tratamentos foram recolhidos e medidas da microdureza em diferentes profundidades (10, 20, 40, 60, 120 e 180 μm) em cada bloco foram observadas em triplicata. Mais detalhes são encontrados em Korytnicki (2002). Os dados estão dispostos na Tabela 1.3.

Tabela 1.3: *Microdureza do esmalte dentário (em KHN) sob o tratamento Controle Extra Bucal.*

Voluntário	Dente	Profundidade (μm)						
		10	20	40	60	80	120	180
1	1	162.33	135.67	166.67	207.00	290.00	289.67	266.67
1	2	193.67	168.33	178.00	196.33	275.00	317.33	330.67
1	3	159.33	165.33	165.33	255.67	295.33	308.67	319.00
1	4	189.00	175.67	162.00	218.00	293.67	273.33	287.33
2	5	164.33	202.33	199.33	241.00	287.67	333.33	301.67
2	6	175.33	191.33	151.33	234.00	285.67	317.00	307.33
2	7	178.67	193.00	172.00	264.33	298.00	310.67	310.67
2	8	170.00	192.67	159.67	216.33	317.67	301.67	311.33
3	9	145.33	179.00	188.33	219.67	245.67	263.33	274.00
3	10	167.33	171.00	166.00	200.33	259.67	283.00	286.33
3	11	159.67	193.33	187.33	226.67	266.00	310.67	282.00
3	12	160.33	163.67	155.67	241.67	252.67	277.00	305.00
4	13	142.00	213.33	199.33	203.33	300.00	298.00	300.00
4	14	115.67	155.33	192.00	251.33	290.33	280.67	309.67
4	15	148.00	193.00	190.33	197.00	294.00	264.67	284.67
4	16	125.67	181.00	172.00	182.33	274.00	307.33	284.00
5	17	163.00	190.33	178.00	221.33	279.67	283.33	290.33
5	18	177.00	164.33	160.67	208.33	302.67	309.00	292.33
5	19	195.00	191.00	185.00	245.33	274.00	288.67	293.67
5	20	193.33	151.67	176.00	234.33	261.67	293.67	299.33
6	21	154.33	189.67	180.00	244.67	267.33	294.67	296.33
6	22	177.00	161.00	161.00	240.67	274.33	278.33	280.33
6	23	160.33	160.33	188.00	214.00	282.67	289.67	296.33
6	24	169.67	174.33	162.67	218.00	287.00	281.33	295.67
7	25	177.67	154.00	190.67	237.33	288.00	270.67	286.33
7	26	152.33	130.33	138.00	241.33	310.67	292.67	299.67
7	27	172.67	165.67	174.33	211.67	259.00	304.00	300.67
7	28	165.00	139.33	164.67	218.00	268.67	290.67	281.67
8	29	167.33	171.67	205.33	261.67	296.33	294.33	296.33
8	30	159.00	159.33	197.00	246.33	277.33	302.67	303.33
8	31	135.33	173.00	158.67	250.67	292.67	289.33	300.33
8	32	170.33	153.67	148.67	243.33	302.33	285.33	286.33

Tabela 1.3 (Continuação): *Microdureza do esmalte dentário (em KHN) sob o tratamento Controle in Situ.*

Voluntário	Dente	Profundidade (μm)						
		10	20	40	60	80	120	180
1	1	271.00	297.67	301.00	302.33	304.67	313.00	314.67
1	2	288.00	295.67	302.67	308.00	309.33	310.00	302.00
1	3	289.67	296.33	319.33	318.00	302.67	316.00	324.00
1	4	280.00	311.33	305.67	311.33	308.00	296.67	313.00
2	5	289.00	301.67	304.33	311.33	309.67	317.00	322.00
2	6	325.00	305.00	316.00	309.67	308.00	322.00	312.00
2	7	282.33	273.33	297.00	305.67	305.00	309.67	318.67
2	8	289.00	288.33	292.00	290.33	299.67	298.33	325.67
3	9	287.00	292.33	296.33	314.00	311.33	299.00	337.67
3	10	287.00	284.67	305.00	305.33	305.67	302.67	306.00
3	11	276.33	308.33	312.33	304.67	300.67	310.33	302.33
3	12	296.33	296.67	304.33	313.00	323.00	316.67	313.00
4	13	286.00	286.00	299.33	302.67	326.33	310.33	309.33
4	14	294.00	294.33	294.67	305.33	319.00	304.33	309.33
4	15	275.00	298.00	313.33	307.00	306.33	319.67	320.33
4	16	279.67	292.33	307.67	309.67	307.00	309.67	315.33
5	17	281.67	306.33	305.00	300.67	307.33	298.33	300.00
5	18	292.67	296.33	309.00	314.33	314.33	313.00	311.00
5	19	292.67	287.67	292.67	320.67	302.33	316.00	315.67
5	20	282.67	308.33	314.33	306.67	305.33	318.33	314.00
6	21	282.33	290.67	302.67	304.33	314.33	315.33	319.67
6	22	277.00	321.33	301.67	316.00	302.33	312.33	317.00
6	23	276.00	287.33	307.67	312.33	301.00	298.00	310.00
6	24	267.33	299.67	302.00	313.33	312.00	309.00	313.33
7	25	280.67	313.67	290.00	299.00	312.67	310.33	313.00
7	26	269.67	297.67	302.33	295.67	300.67	295.00	301.00
7	27	264.67	292.33	304.00	316.00	310.67	291.67	310.00
7	28	270.33	302.00	306.67	303.00	326.00	301.33	320.67
8	29	283.67	313.67	296.33	309.00	308.33	322.67	322.00
8	30	279.33	303.33	295.00	311.67	317.33	310.67	319.67
8	31	286.67	302.67	277.67	300.33	302.00	313.00	314.67
8	32	287.67	299.00	306.33	307.67	317.00	317.67	310.00

Tabela 1.3 (Continuação): *Microdureza do esmalte dentário (em KHN) sob o tratamento Laser.*

Voluntário	Dente	Profundidade (μm)						
		10	20	40	60	80	120	180
1	1	239.00	304.00	282.67	300.00	308.67	330.67	318.67
1	2	275.33	331.67	310.00	268.00	272.33	316.33	367.67
1	3	196.67	301.67	296.67	286.00	256.33	318.33	318.00
1	4	251.67	339.33	285.00	299.67	317.00	297.67	299.00
2	5	265.67	335.33	292.67	308.67	284.33	294.33	311.33
2	6	268.67	337.67	300.00	290.33	311.67	311.33	309.67
2	7	246.00	307.67	323.67	270.67	278.33	319.00	310.67
2	8	266.33	347.00	326.67	298.33	290.67	300.33	339.67
3	9	226.33	319.67	277.00	282.00	283.33	274.00	286.00
3	10	216.00	336.67	283.67	282.33	252.67	299.00	255.67
3	11	216.00	362.33	290.67	252.67	238.67	304.33	304.67
3	12	243.33	373.33	301.67	294.00	236.67	256.00	262.33
4	13	246.67	355.00	316.33	283.67	313.00	324.67	301.67
4	14	258.33	321.67	315.33	327.67	336.33	328.67	272.67
4	15	250.33	302.00	296.00	321.67	333.00	316.67	274.67
4	16	232.33	334.67	290.33	265.67	277.67	301.00	300.00
5	17	227.33	338.00	278.33	266.00	312.00	283.67	280.67
5	18	249.33	352.67	322.67	287.00	301.33	315.67	307.33
5	19	247.67	330.33	303.67	293.67	260.00	315.33	301.00
5	20	255.00	341.33	279.67	285.67	299.33	294.33	305.00
6	21	239.33	350.00	331.00	320.33	309.00	354.67	331.33
6	22	251.00	301.33	322.33	291.33	365.00	328.67	344.33
6	23	197.00	360.00	325.67	259.00	284.33	281.00	296.00
6	24	221.33	312.00	321.33	297.67	311.33	319.00	341.00
7	25	224.67	332.33	311.00	312.33	345.67	346.00	338.67
7	26	232.00	281.00	268.67	284.33	270.67	337.00	305.67
7	27	193.00	339.00	286.33	290.33	282.33	303.33	350.00
7	28	240.67	376.00	300.67	281.33	300.33	297.67	276.33
8	29	242.67	381.33	295.33	306.00	312.33	339.00	313.67
8	30	248.33	350.33	306.33	322.33	344.33	351.00	328.33
8	31	237.67	322.67	313.00	281.67	303.67	326.67	355.67
8	32	265.67	328.33	323.67	335.33	309.67	325.33	353.67

Exemplo 1.4. Estudo do Potencial de Ação Composto (PAC) em mexilhões.

Neste estudo, as unidades amostrais são mexilhões (moluscos bivalves) coletados no litoral de São Sebastião. Após a coleta, os mexilhões foram colocados em tanques

com água do mar do próprio local de coleta (salinidade 3.4%), onde permaneceram por 3 ou 4 dias. Após esse período, alguns foram transferidos para tanques com salinidade 2.4% e outros para tanques com salinidade 1.9%, onde permaneceram por aproximadamente duas semanas para aclimação. Desta forma, 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 fechamento de valvas; se após o teste elas permanecessem fechadas, o molusco era considerado bom para o estudo; em caso contrário, era desprezado. Em seguida, o nervo conectivo cérebro visceral foi extraído de cada animal e colocado numa câmara especial para a estimulação elétrica. Como resposta ao estímulo elétrico foram medidas três características de interesse, das quais a amplitude (mV) do PAC (potencial de ação composto) constitui o foco deste exemplo. Se a amplitude do PAC obtida estivesse abaixo de 6 mV, o animal era desprezado. Este valor foi considerado como o limite inferior de indicação de *stress* devido ao processo necessário para a realização das medições. Ao final do período de aclimação, a amplitude do PAC era observada e, em seguida, cada animal era submetido aleatoriamente a uma salinidade teste diferente de sua salinidade de aclimação por dois minutos. As salinidades testes foram 0.7%, 1.4%, 1.9%, 2.4% e 3.4%. No final deste período, a variável de interesse era novamente medida e o molusco retornava ao seu tanque de origem (tanque com a salinidade de aclimação), onde permanecia durante três minutos e, após esse período, observavam-se novamente as características de interesse. As medidas obtidas sob essas condições foram consideradas medidas controles associadas à última salinidade teste à qual o molusco foi submetido. Novamente a medida da variável de interesse foi obtida e o processo foi repetido para as demais salinidades teste. O objetivo desse estudo era comparar o comportamento das respostas eletrofisiológicas dos mexilhões não estressados dos três grupos sob as diferentes salinidades testes. Mais detalhes sobre o fenômeno investigado podem ser encontrados em Stucchi-Zucchi (1988) e em Cúri (1999). Os dados estão dispostos na Tabela 1.4.

Tabela 1.4: Amplitude média do PAC (mV) no meio hipo osmótico para o grupo com salinidade de aclimação a 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

Tabela 1.4 (Continuação): Amplitude média do PAC (mV) no meio hipo osmótico para o grupo com salinidade de aclimação a 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

Tabela 1.4 (Continuação): Amplitude média do PAC (mV) no meio hipo osmótico para o grupo com salinidade de aclimação a 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	

Os Exemplos 1.1, 1.2 e 1.3 correspondem a estudos longitudinais. O Exemplo 1.1 é um estudo com um fator intra-unidades amostrais (tempo), irregular, balanceado em relação às condições de avaliação e completo. O Exemplo 1.2 apresenta um fator intra-unidades amostrais (tempo), é um estudo regular, balanceado e incompleto. O Exemplo 1.3 apresenta dois fatores intra-unidades amostrais, profundidade com sete níveis e tratamento com três níveis, além de dois fatores entre unidades amostrais, a saber: dente com 32 níveis e voluntário com oito níveis. É um estudo irregular, balanceado em relação às condições de avaliação e completo. O Exemplo 1.4 corresponde a um estudo com medidas repetidas não longitudinais, com um fator intra-unidades amostrais (salinidade teste com 4 níveis); um fator entre unidades amostrais (salinidade de aclimação) com três níveis e uma covariável dependente do tempo (salinidade controle) com quatro níveis.

É um estudo irregular, não balanceado e incompleto.

A análise de dados com medidas repetidas envolve essencialmente modelos de ANOVA, ANCOVA e Regressão. Os modelos de ANOVA e ANCOVA são adequados quando o fator intra-unidades amostrais corresponde a uma variável categorizada, como no caso do Exemplo 1.3. Quando o fator intra-unidades amostrais corresponde a uma variável contínua, modelos de regressão são mais apropriados, como ocorre nos Exemplos 1.1 e 1.2. O Exemplo 1.4 permite análises sob os dois enfoques.

A característica que diferencia os estudos com medidas repetidas dos casos clássicos é a possível correlação entre as observações intra-unidades amostrais. Modelos para a análise de estudos com medidas repetidas devem apresentar estruturas de covariância que incorporem essas correlações. Impor uma estrutura para os elementos da matriz de covariância pode ser fundamental na análise, especialmente em situações com dados omissos ("missing data") ou em situações com dados desbalanceados em relação ao tempo ou ainda quando o número de condições de avaliação é elevado. Porque, por exemplo, em estudos com medidas repetidas com um número de unidades amostrais pequeno em relação ao número de condições de avaliação e com dados omissos a estimação de modelos com uma estrutura de covariâncias não estruturada não seria possível. Mais detalhes em Singer e Adrade (2002).

Vários pesquisadores têm estudado estratégias de análise em que os parâmetros de regressão e os de covariância são modelados separadamente ou conjuntamente, dentre os quais se destacam Diggle (1988), Wolfinger (1993 e 1996), Diggle et al. (2002) e Singer e Adrade (2002).

Modelos mistos são bastante utilizados para a análise de estudos com medidas repetidas por permitirem a modelagem de matrizes de covariância com um número pequeno de parâmetros. Eles comportam conjuntos de dados incompletos e não balanceados e podem ser avaliados por intermédio de técnicas de verossimilhança e verossimilhança restrita. A escolha da estrutura de localização é similar àquela utilizada na análise usual de modelos lineares, mas a escolha da estrutura de covariância pode apresentar dificuldades. Sua importância está associada à obtenção de estimadores mais precisos dos parâmetros

de localização e, por sua vez, à realização de inferências eficientes.

O objetivo deste trabalho é avaliar métodos para auxiliar a escolha de estruturas de covariância em estudos com medidas repetidas. Assumiremos que a variável resposta tem distribuição normal e que os dados omissos são do tipo MCAR. Nas seções seguintes deste capítulo apresentamos um resumo da teoria dos modelos lineares mistos. No Capítulo 2 é descrita a estratégia para a modelagem da estrutura de covariância além de métodos analíticos e gráficos envolvidos em sua seleção. No Capítulo 3 são discutidos a seleção da estrutura de covariância para os exemplos dessa seção e os aspectos computacionais envolvidos. Por fim, no Capítulo 4, são dadas sugestões para pesquisas futuras.

1.1 Modelo Estatístico

A forma geral do modelo misto é

$$\mathbf{y}_i = \mathbf{X}_i\boldsymbol{\beta} + \mathbf{Z}_i\mathbf{b}_i + \mathbf{e}_i, \quad i = 1, \dots, N, \quad (1.1)$$

em que \mathbf{y}_i ($n_i \times 1$) é o vetor de observações da i -ésima unidade amostral, $\boldsymbol{\beta}$ ($p \times 1$) é o vetor de parâmetros populacionais desconhecidos (*também chamados de efeitos fixos ou parâmetros de localização*), \mathbf{X}_i ($n_i \times p$) é a matriz de constantes conhecidas que corresponde à especificação para os efeitos fixos, \mathbf{b}_i ($q \times 1$) é um vetor de efeitos aleatórios, que reflete o comportamento individual da i -ésima unidade amostral, \mathbf{Z}_i ($n_i \times q$) é uma matriz de constantes conhecidas correspondente à especificação para os efeitos aleatórios e \mathbf{e}_i ($n_i \times 1$) é um vetor de erros aleatórios, que representa a variabilidade residual das observações realizadas na i -ésima unidade amostral. Assumimos independência entre os vetores \mathbf{e}_i , independência entre os vetores \mathbf{b}_i e independência entre \mathbf{b}_i e \mathbf{e}_i . Além disso, assumimos que $\mathbf{b}_i \sim NID_q(\mathbf{0}, \mathbf{G})$, em que \mathbf{G} ($q \times q$) é uma matriz simétrica definida positiva, $\mathbf{e}_i \sim NID_{n_i}(\mathbf{0}, \mathbf{R}_i)$, com \mathbf{R}_i ($n_i \times n_i$) denotando uma matriz simétrica definida positiva. Portanto, o vetor de respostas associado à i -ésima unidade amostral tem distribuição normal multivariada com vetor de médias e matriz de covariâncias dados, respectivamente,

por

$$\begin{aligned} E(\mathbf{y}_i) &= \mathbf{X}_i\boldsymbol{\beta} \\ \text{Var}(\mathbf{y}_i) &= \mathbf{V}_i = \mathbf{Z}_i\mathbf{G}\mathbf{Z}_i^t + \mathbf{R}_i. \end{aligned} \quad (1.2)$$

As matrizes \mathbf{R}_i e \mathbf{G} são funções de parâmetros desconhecidos, ou seja, $\mathbf{G} = \mathbf{G}(\boldsymbol{\theta}_1)$ e $\mathbf{R}_i = \mathbf{R}_i(\boldsymbol{\theta}_2)$, de modo que a matriz de covariâncias para a i -ésima unidade amostral, $\mathbf{V}_i = \mathbf{V}_i(\boldsymbol{\theta})$, dependerá de $\boldsymbol{\theta} = (\boldsymbol{\theta}_1^t, \boldsymbol{\theta}_2^t)^t$.

Este modelo também pode ser considerado como um *modelo linear em dois estágios*. De acordo com Laird e Ware (1982), no primeiro estágio, consideramos fixos, ou seja, conhecidos os efeitos aleatórios \mathbf{b}_i , de forma que

$$\mathbf{y}_i | \mathbf{b}_i \sim N(\mathbf{X}_i\boldsymbol{\beta} + \mathbf{Z}_i\mathbf{b}_i, \mathbf{R}_i). \quad (1.3)$$

No segundo estágio, assumimos que os vetores \mathbf{b}_i são independentes com distribuição $N_q(\mathbf{0}, \sigma^2\mathbf{G})$. Portanto, o modelo marginal é dado por

$$\mathbf{y}_i \sim N(\mathbf{X}_i\boldsymbol{\beta}, \mathbf{Z}_i\mathbf{G}\mathbf{Z}_i^t + \mathbf{R}_i), \quad (1.4)$$

com erro marginal dado por $\boldsymbol{\varepsilon}_i = \mathbf{y}_i - \mathbf{X}_i\boldsymbol{\beta} = \mathbf{e}_i + \mathbf{Z}_i\mathbf{b}_i$.

Quando $\mathbf{R}_i = \sigma^2\mathbf{I}_{n_i}$ com \mathbf{I}_{n_i} denotando a matriz identidade de dimensão n_i , o modelo é chamado de *modelo de independência condicional*, indicando que as n_i observações associadas à i -ésima unidade amostral são independentes condicionalmente a \mathbf{b}_i .

Mais detalhes sobre modelos mistos para estudos com medidas repetidas ou longitudinais podem ser encontrados em Crowder e Hand (1990), Jones (1993), Crowder e Hand (1996), Verbeke e Molenberghs (1997), Vonesh e Chinchilli (1997), Singer e Andrade (2000), Diggle et al. (2002) e Davis (2002) entre outros.

1.2 Estruturas de Covariâncias

O modelo para matriz de covariâncias \mathbf{V}_i deve depender da maneira pela qual as observações foram obtidas e do conhecimento sobre o mecanismo gerador das observações.

Diggle (1988) e Diggle et al. (2002) comentam que a matriz de covariâncias deve ser suficientemente flexível para incluir no mínimo três fontes diferentes de variação aleatória, nomeadamente: i) a variação devida a efeitos aleatórios, quando as unidades de investigação formam uma amostra aleatória da população com a característica de interesse; ii) a variação que pode ser explicada por correlação serial, em que se esperam observações próximas mais fortemente correlacionadas que observações mais distantes e iii) a variação devida a erros de medição.

No contexto dos modelos mistos, a covariância entre as observações obtidas em uma mesma unidade amostral poderá ser modelada indiretamente por meio dos efeitos aleatórios, \mathbf{b}_i , que representa a variabilidade entre as unidades amostrais, diretamente através da matriz \mathbf{R}_i , que representa a covariância das observações intra-unidades amostrais como nos modelos lineares usuais, ou ainda, como uma combinação, de \mathbf{b}_i e \mathbf{R}_i , como sugerem Vonesh e Chinchilli (1997), permitindo deste modo incluir na estrutura de covariância as três fontes de variação: os efeitos aleatórios e a variação devida à medição, na matriz \mathbf{G} e a correlação serial na matriz \mathbf{R}_i .

No modelo linear misto com independência condicional, $\mathbf{R}_i = \sigma^2 \mathbf{I}_{n_i}$ e \mathbf{G} é uma matriz de componentes de variância como sugerido em Searle et al. (1971 e 1992). Modelos desse tipo são utilizados, por exemplo, em estudos com planejamentos em blocos aleatórios ou do tipo split-plot.

Das possíveis estruturas de covariâncias disponíveis na literatura estatística, algumas estão descritas abaixo, por intermédio de exemplos com $n_i = 4$.

1. Estrutura Uniforme

$$\mathbf{V}_i(\boldsymbol{\theta}) = \begin{bmatrix} \sigma^2 + \tau & \tau & \tau & \tau \\ \tau & \sigma^2 + \tau & \tau & \tau \\ \tau & \tau & \sigma^2 + \tau & \tau \\ \tau & \tau & \tau & \sigma^2 + \tau \end{bmatrix}. \quad (1.5)$$

Segundo esse modelo, as observações realizadas na mesma unidade amostral têm tanto variâncias quanto covariâncias homogêneas. Portanto, a estrutura é função

de apenas dois parâmetros agrupados no vetor $\boldsymbol{\theta} = (\sigma^2, \tau)^t$ e pode ser modelada de duas maneiras. A primeira considera $\mathbf{G} = \mathbf{0}$ e $\mathbf{R}_i = \sigma^2 \mathbf{I}_{n_i} + \tau \mathbf{J}_{n_i}$ com \mathbf{J}_{n_i} denotando uma matriz de dimensão $(n_i \times n_i)$ com todos os elementos iguais a 1. Na segunda, tomamos $\mathbf{G} = \tau$ com $\tau > 0$, $\mathbf{Z}_i = \mathbf{1}_{n_i}$ é um vetor de dimensão $(n_i \times 1)$ com todos os elementos iguais a 1 e $\mathbf{R}_i = \sigma^2 \mathbf{I}_{n_i}$. A primeira formulação é mais geral, pois não impõe à restrição de que $\tau \geq 0$.

2. Estrutura AR(1)

$$\mathbf{V}_i(\boldsymbol{\theta}) = \sigma^2 \begin{bmatrix} 1 & \phi & \phi^2 & \phi^3 \\ 0 & 1 & \phi & \phi^2 \\ \phi^2 & \phi & 1 & \phi \\ \phi^3 & \phi^2 & \phi & 1 \end{bmatrix}. \quad (1.6)$$

Sob esta estrutura, $\mathbf{G} = \mathbf{0}$ e \mathbf{R}_i é uma matriz que é gerada por um modelo em que os erros aleatórios das medidas realizadas no i -ésimo indivíduo têm a seguinte característica: $e_{ij} = \phi e_{i,j-1} + \delta_{ij}$, $\delta_{ij} \sim NIID(0, \tau^2)$ e são não correlacionados com e_{il} , $l = 1, \dots, j-1$. Esta estrutura é função de $\boldsymbol{\theta} = (\phi, \sigma^2)^t$, com $\sigma^2 = \tau^2/(1 - \phi^2)$ e $|\phi| < 1$ para garantir a estacionariedade. Mais detalhes podem ser encontrados em Chi e Reinsel (1989), Jones (1990), Jones e Boadi-Boateng (1991), Hamilton (1994), Littell et al (1996) e Verbeke e Molenberghs (1997).

3. Estrutura ARMA(1,1)

$$\mathbf{V}_i(\boldsymbol{\theta}) = \sigma^2 \begin{bmatrix} 1 & \gamma & \gamma\phi & \gamma\phi^2 \\ \gamma & 1 & \gamma & \gamma\phi \\ \gamma\phi & \gamma & 1 & \gamma \\ \gamma\phi^2 & \gamma\phi & \gamma & 1 \end{bmatrix}. \quad (1.7)$$

Sob esta estrutura, $\mathbf{G} = \mathbf{0}$ e \mathbf{R}_i é uma matriz que pode ser gerada por um modelo em que os erros aleatórios associados ao i -ésimo indivíduo têm a seguinte característica: $e_{ij} = \phi e_{i,j-1} + \delta_{ij} + \theta \delta_{i,j-1}$, $\delta_{ij} \sim NIID(0, \sigma^2)$, ou seja, segue um processo auto-regressivo de médias móveis; usando a notação usual, ϕ é o parâmetro

autoregressivo, θ modela a média móvel e σ^2 representa a variância residual. Portanto, o vetor de parâmetros para essa estrutura é $\theta = (\phi, \gamma, \sigma^2)^t$. A covariância entre o erro aleatório relativo ao i -ésimo indivíduo entre os instantes t e $t+h$ é $\gamma\phi^{h-1}$, com

$$\gamma = \frac{(1 - \phi\theta)(\phi - \theta)}{1 - \phi^2}; \quad (1.8)$$

para garantir estacionariedade, assumimos $|\phi| < 1$. Mais detalhes sobre modelos auto-regressivos de médias móveis podem ser encontrados em Fuller (1976) ou Hamilton (1994) e para aplicações deste modelo em estudos com medidas repetidas em Rochon e Helms (1989), Littell et al. (1996) e Verbeke e Molenberghs (1997).

4. Estrutura Ante-Dependência de ordem 1

$$\mathbf{V}_i(\theta) = \begin{bmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho_1 & \sigma_1\sigma_3\rho_1\rho_2 & \sigma_1\sigma_4\rho_1\rho_2\rho_3 \\ \sigma_1\sigma_2\rho_1 & \sigma_2^2 & \sigma_2\sigma_3\rho_2 & \sigma_2\sigma_4\rho_2\rho_3 \\ \sigma_1\sigma_3\rho_1\rho_2 & \sigma_2\sigma_3\rho_2 & \sigma_3^2 & \sigma_3\sigma_4\rho_3 \\ \sigma_1\sigma_4\rho_1\rho_2\rho_3 & \sigma_2\sigma_4\rho_2\rho_3 & \sigma_3\sigma_4\rho_3 & \sigma_4^2 \end{bmatrix}. \quad (1.9)$$

Sob esta estrutura, $\mathbf{G} = \mathbf{0}$ e \mathbf{R}_i é uma matriz que pode ser gerada pelo seguinte modelo para o erro aleatório das medidas realizadas no i -ésimo indivíduo: $e_{ij} = \phi_j e_{i,j-1} + \delta_{ij}$, com $\delta_{ij} \sim N(0, \tau_j^2)$, não correlacionado com e_{il} , $l = 1, \dots, j-1$ e $|\phi_j| < 1$. A variância em cada instante j é dada por $\sigma_j^2 = \sum_{l=1}^{j-1} (\prod_{k=l+1}^j \phi_k \tau_l^2) + \tau_j^2$. A correlação entre as medidas realizadas nos instantes $j-1$ e j para $j > 1$ é dada por $\rho_{j-1} = \phi_j \tau_{j-1} / \tau_j$, o que implica $\phi_j = \rho_{j-1} \tau_j / \tau_{j-1}$. Esta estrutura permite que as variâncias sejam diferentes e é aplicável em estudos longitudinais em que as condições de avaliação não são igualmente espaçadas, apresentam heterogeneidade de variância e correlação serial. O elemento (i, j) da matriz de covariância para $j > i$ é dado por

$$\sigma_i \sigma_j \prod_{k=i}^{j-1} \rho_k \quad (1.10)$$

em que $i, j = 1, \dots, n_i$ e $|\rho_k| < 1$. Essa matriz é função do vetor de parâmetros $\theta = (\sigma_1^2, \dots, \sigma_{n_i}^2, \rho_1, \dots, \rho_{n_i-1})^t$. O preço da flexibilidade deste modelo é o acréscimo no

número de parâmetros (da ordem $2n_i - 1$). Mais detalhes sobre este modelo podem ser encontrados em Gabriel (1962), Kenward (1987), Wolfinger (1996), Timothy et al. (1997) e Zimmerman e Núñez (2001).

5. Estrutura Toeplitz

$$\mathbf{V}_i(\boldsymbol{\theta}) = \begin{bmatrix} \sigma^2 & \sigma_1 & \sigma_2 & \sigma_3 \\ \sigma_1 & \sigma^2 & \sigma_1 & \sigma_2 \\ \sigma_2 & \sigma_1 & \sigma^2 & \sigma_1 \\ \sigma_3 & \sigma_2 & \sigma_1 & \sigma^2 \end{bmatrix}. \quad (1.11)$$

Sob esta estrutura, $\mathbf{G} = \mathbf{0}$ e \mathbf{R}_i é a matriz de covariância de um processo de médias móveis de ordem $q = n_i - 1$ (neste exemplo $q = 3$) gerada por um modelo em que o erro aleatório das medidas realizadas no i -ésimo indivíduo tem a seguinte característica: $e_{ij} = \delta_{ij} + \theta_1\delta_{i(j-1)} + \dots + \theta_{n_i-1}\delta_{i(j-n_i+1)}$, $\delta_{ij} \sim NIID(0, \tau^2)$ e são não correlacionados com e_{il} , $l = 1, \dots, j - 1$. A variância é dada por $\sigma^2 = (1 + \theta_1 + \dots + \theta_{n_i-1})$ e as covariâncias são $\sigma_j = (\theta_j + \theta_{j+1}\theta_1 + \theta_{j+2}\theta_2 + \dots + \theta_{n_i}\theta_{n_i-1})$. Esta estrutura é função de $\boldsymbol{\theta} = (\sigma^2, \sigma_1, \dots, \sigma_{n_i-1})^t$; para o exemplo, $\boldsymbol{\theta} = (\sigma^2, \sigma_1, \sigma_2, \sigma_3)^t$. Mais detalhes sobre esta estrutura podem ser encontrados em Littell et al. (1996), Wolfinger (1996) e Verbeke e Molenberghs (1997).

6. Estrutura Espacial ou Markov

$$\mathbf{V}(\boldsymbol{\theta})_i = \sigma^2 \begin{bmatrix} 1 & \rho^{d_{12}} & \rho^{d_{13}} & \rho^{d_{14}} \\ \rho^{d_{21}} & 1 & \rho^{d_{23}} & \rho^{d_{24}} \\ \rho^{d_{31}} & \rho^{d_{32}} & 1 & \rho^{d_{34}} \\ \rho^{d_{41}} & \rho^{d_{42}} & \rho^{d_{43}} & 1 \end{bmatrix}. \quad (1.12)$$

Sob esta estrutura, $\mathbf{G} = \mathbf{0}$ e \mathbf{R}_i é uma matriz de covariância espacial. Esta estrutura apresenta variância homogênea, como as estruturas autoregressivas, e é indicada quando as condições de avaliação sob as quais as observações não são igualmente espaçadas. O índice d_{ij} denota a distância entre os instantes de observação i e j , $d_{ij} = |i - j|$. Esta estrutura é invariante relativamente a transformações lineares

na escala das condições de avaliação e é função de $\theta = (\sigma^2, \rho)^t$. Mais detalhes podem ser encontrados em Crowder e Hand (1990), Crowder e Hand (1996), Littell et al. (1996) e Verbeke e Molenberghs (1997).

7. Estrutura Uniforme Heterogênea

$$\mathbf{V}_i(\theta) = \begin{bmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho & \sigma_1\sigma_3\rho & \sigma_1\sigma_4\rho \\ \sigma_2\sigma_1\rho & \sigma_2^2 & \sigma_2\sigma_3\rho & \sigma_2\sigma_4\rho \\ \sigma_3\sigma_1\rho & \sigma_3\sigma_2\rho & \sigma_3^2 & \sigma_3\sigma_4\rho \\ \sigma_4\sigma_1\rho & \sigma_4\sigma_2\rho & \sigma_4\sigma_3\rho & \sigma_4^2 \end{bmatrix}. \quad (1.13)$$

Sob esta estrutura, $\mathbf{G} = \mathbf{0}$ e \mathbf{R}_i é a matriz de covariâncias gerada quando os erros aleatórios associados ao i -ésimo indivíduo na j -ésima condição de avaliação têm distribuição $e_{ij} \sim N(0; \sigma_j^2)$. Esta estrutura é função do vetor de parâmetros $\theta = (\sigma_1^2, \sigma_2^2, \dots, \sigma_n^2, \rho)^t$; [no nosso exemplo, $\theta = (\sigma_1^2, \sigma_2^2, \sigma_3^2, \sigma_4^2, \rho)^t$] em que ρ representa a correlação entre as medições. Mais detalhes podem ser encontrados em Wolfinger (1993), Littell et al. (1996), Wolfinger (1996) e Verbeke e Molenberghs (1997).

8. Estrutura ARH(1)

$$\mathbf{V}(\theta)_i = \begin{bmatrix} \sigma_1^2 & \sigma_1\sigma_2\phi & \sigma_1\sigma_3\phi^2 & \sigma_1\sigma_4\phi^3 \\ \sigma_2\sigma_1\phi & \sigma_2^2 & \sigma_2\sigma_3\phi & \sigma_2\sigma_4\phi^2 \\ \sigma_3\sigma_1\phi^2 & \sigma_3\sigma_2\phi & \sigma_3^2 & \sigma_3\sigma_4\phi \\ \sigma_4\sigma_1\phi^3 & \sigma_4\sigma_2\phi^2 & \sigma_4\sigma_3\phi & \sigma_4^2 \end{bmatrix}. \quad (1.14)$$

Sob esta estrutura, $\mathbf{G} = \mathbf{0}$ e \mathbf{R}_i é a matriz que é gerada por um modelo em que os erros aleatórios associados ao i -ésimo indivíduo são modelados por $e_{ij} = \phi e_{i,j-1} + \delta_{ij}$, em que $\delta_{ij} \sim N(0, \tau_j^2)$ são independentes e não correlacionados com e_{il} , $l = 1, \dots, j-1$. A variância no instante j é dada por $\sigma_j^2 = \phi^{2(j-1)}\tau_1^2 + \phi^{2(j-2)}\tau_2^2 + \dots + \phi^2\tau_{j-1}^2 + \tau_j^2$ e a covariância entre as observações realizadas nos instantes i e j é dada por $\sigma_i\sigma_j\phi^{|i-j|}$ com $|\phi| < 1$. Esta estrutura corresponde a uma generalização da estrutura AR(1) e é função do vetor de parâmetros, $\theta = (\sigma_1^2, \sigma_2^2, \dots, \sigma_j^2, \phi)^t$. Mais detalhes podem ser encontrados Wolfinger (1993), Littell et al. (1996), Wolfinger (1996) e Verbeke e Molenberghs (1997).

9. Estrutura Toeplitz Heterogênea

$$\mathbf{V}(\boldsymbol{\theta})_i = \begin{bmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho_1 & \sigma_1\sigma_3\rho_2 & \sigma_1\sigma_4\rho_3 \\ \sigma_2\sigma_1\rho_1 & \sigma_2^2 & \sigma_2\sigma_3\rho_1 & \sigma_2\sigma_4\rho_2 \\ \sigma_3\sigma_1\rho_2 & \sigma_3\sigma_2\rho_1 & \sigma_3^2 & \sigma_3\sigma_4\rho_1 \\ \sigma_4\sigma_1\rho_3 & \sigma_4\sigma_2\rho_2 & \sigma_4\sigma_3\rho_1 & \sigma_4^2 \end{bmatrix}. \quad (1.15)$$

Sob esta estrutura, $\mathbf{G} = \mathbf{0}$ e \mathbf{R}_i é a matriz de covariância gerada por um modelo em que os erros aleatórios das medidas realizadas no i -ésimo indivíduo têm a forma $e_{ij} = \delta_{ij} + \theta\delta_{ij-1}$, e $\delta_{ij} \sim NI(0, \tau_j^2)$. A variância no instante j é dada por $\sigma_j^2 = \tau_j^2 + \theta^2\tau_{j-1}^2 + \theta^4\tau_{j-2}^2 + \dots + \theta^{2j-2}\tau_1^2$ e as covariâncias são da forma $\sigma_i\sigma_j\rho_{|i-j|}$ com $|\rho_i| < 1$. É função do vetor de parâmetros $\boldsymbol{\theta} = (\sigma_1^2, \dots, \sigma_j^2, \rho_1, \dots, \rho_{j-1})^t$. Mais detalhes podem ser encontrados em Wolfinger (1993), Littell et al. (1996), Wolfinger (1996) e Verbeke e Molenberghs (1997).

10. Modelo de efeitos aleatórios

Esse modelo é similar ao apresentado em (1.1), sendo $\mathbf{R}_i = \sigma^2\mathbf{I}_i$. Considerando, por exemplo, efeitos aleatórios para os coeficientes lineares e angulares, a matriz de covariância será dada por

$$\mathbf{V}_i = \mathbf{Z}_i \begin{bmatrix} \sigma_0^2 & \sigma_{01} \\ \sigma_{01} & \sigma_1^2 \end{bmatrix} \mathbf{Z}_i^t + \sigma^2\mathbf{I}_i. \quad (1.16)$$

11. Não Estruturada (NE)

Neste caso nenhuma restrição é imposta aos parâmetros. Para n_i condições de avaliação, o número de parâmetros é $n_i(n_i + 1)/2$.

$$\begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} & \sigma_{14} \\ \sigma_{12} & \sigma_2^2 & \sigma_{23} & \sigma_{24} \\ \sigma_{13} & \sigma_{23} & \sigma_3^2 & \sigma_{34} \\ \sigma_{14} & \sigma_{24} & \sigma_{34} & \sigma_4^2 \end{bmatrix}. \quad (1.17)$$

Uma classe de estruturas bastante geral é a linear apresentada em Andrade e Helms (1984) ou em Singer e Andrade (2000), por exemplo. Os elementos dessa classe podem ser expressos como

$$\mathbf{V}_i(\boldsymbol{\theta}) = \sum_{h=1}^d \phi_h \mathbf{F}_{hi} \quad (1.18)$$

em que \mathbf{F}_{hi} são matrizes convenientes conhecidas e ϕ_h são parâmetros desconhecidos. Este modelo para a estrutura de covariâncias é adequado quando os parâmetros da matriz de covariâncias são aditivos, como é o caso das estruturas Uniforme e Toeplitz. Se os parâmetros forem multiplicativos como ocorre, por exemplo, na estrutura AR(1), que possui apenas dois parâmetros, precisaremos de quatro parâmetros para escrevê-la como um modelo linear, duplicando o número de parâmetros a serem estimados. O modelo Toeplitz poderá ser escrito nessa forma com $\phi_1 = \sigma^2, \phi_2 = \sigma_1, \phi_3 = \sigma_2, \phi_4 = \sigma_3$ e as matrizes \mathbf{F}_{hi} dadas por

$$\mathbf{F}_{1i} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}; \quad \mathbf{F}_{2i} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}; \quad (1.19)$$

$$\mathbf{F}_{3i} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}; \quad \mathbf{F}_{4i} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}.$$

Galecki (1994) propõe uma classe de estruturas de covariâncias baseada no produto de Kronecker, para estudos com medidas repetidas com mais de um fator intra-unidades amostrais, como no exemplo 1.3. Dentre as estruturas disponíveis temos

1. Não estruturado Kronecker Auto-regressivo

$$\mathbf{R}(\boldsymbol{\theta}) = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_2^2 & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & \sigma_3^2 \end{bmatrix} \otimes \begin{bmatrix} 1 & \phi & \phi^2 & \phi^3 & \phi^4 & \phi^5 & \phi^6 \\ \phi & 1 & \phi & \phi^2 & \phi^3 & \phi^4 & \phi^5 \\ \phi & \phi^2 & 1 & \phi & \phi^2 & \phi^3 & \phi^4 \\ \phi & \phi^2 & \phi^3 & 1 & \phi & \phi^2 & \phi^3 \\ \phi & \phi^2 & \phi^3 & \phi^4 & 1 & \phi & \phi^2 \\ \phi & \phi^2 & \phi^3 & \phi^4 & \phi^5 & 1 & \phi \\ \phi & \phi^2 & \phi^3 & \phi^4 & \phi^5 & \phi^6 & 1 \end{bmatrix} \quad (1.20)$$

2. Não estruturado Kronecker Uniforme

$$\mathbf{R}(\boldsymbol{\theta}) = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_2^2 & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & \sigma_3^2 \end{bmatrix} \otimes \begin{bmatrix} 1 & \tau & \tau & \tau & \tau & \tau & \tau \\ \tau & 1 & \tau & \tau & \tau & \tau & \tau \\ \tau & \tau & 1 & \tau & \tau & \tau & \tau \\ \tau & \tau & \tau & 1 & \tau & \tau & \tau \\ \tau & \tau & \tau & \tau & 1 & \tau & \tau \\ \tau & \tau & \tau & \tau & \tau & 1 & \tau \\ \tau & \tau & \tau & \tau & \tau & \tau & 1 \end{bmatrix}. \quad (1.21)$$

Wolfinger (1993) descreve uma metodologia, baseada no teste da razão de verossimilhanças para comparar modelos com estruturas de covariância encaixadas e nos critérios de informação de Akaike (AIC_R) e de Schwarz (BIC_R) para comparar modelos em que as estruturas de covariância não são necessariamente encaixadas. A principal desvantagem do procedimento é a necessidade do ajuste de um número elevado de estruturas, além disso ele não em consideração o tipo de planejamento.

Uma outra estratégia de análise, sugerida em Grady e Helms (1995), para seleção de matrizes de covariância em estudos incompletos consiste em considerar inicialmente um modelo com menor número de parâmetros para os efeitos fixos em relação ao saturado e uma matriz de covariância não estruturada. Após o ajuste desse modelo, esses autores sugerem a construção de um gráfico com os valores dos elementos das linhas da matriz de covariâncias intra-unidades amostrais ou da matriz de correlações correspondente versus a

defasagem (“lag”) entre as condições de avaliação. Gráficos que apresentarem covariâncias e correlações decrescentes com as distâncias entre as observações evidenciam a presença de correlação serial nas observações da mesma unidade amostral. Por intermédio desses gráficos, poderemos selecionar como candidatas para modelar a correlação serial intra unidades amostrais estruturas de covariância que incorporem esse padrão de correlação, como modelos Toeplitz, auto-regressivo média móvel (ARMA), auto-regressivos (AR). Grady e Helms (1995) também propõem ajustes de estruturas de covariância obtidas por meio da inclusão de efeitos aleatórios e de estruturas de covariâncias que envolvem a combinação de efeitos aleatórios e uma estrutura auto-regressiva para o erro. Para comparação desses modelos sugerem o teste da razão de verossimilhanças como proposto por Jennrich e Schluchter (1996).

Diggle (1988) e Diggle et al. (2002), por outro lado, propõem ajustar um modelo saturado para os efeitos fixos com a finalidade de evitar a introdução de correlação espúria e uma estrutura de covariância intra-unidades amostrais que seja suficientemente flexível para incorporar, quando necessário, efeitos aleatórios, erro de medição e correlação serial.

Littell et al. (2000) sugerem o ajuste de um modelo inicial saturado tanto para os efeitos fixos quanto para a estrutura de covariância e o ajuste subsequente de estruturas covariância mais parcimoniosas tais como uniforme, uniforme heterogênea, auto-regressiva etc. Eles sugerem que a comparação dessas estruturas de covariância seja feita por meio dos critérios de informação AIC_R e BIC_R .

A estratégia proposta por Grady e Helms (1995) permite o ajuste de um modelo saturado para a estrutura de covariância intra-unidades amostrais, o que possibilita visualizar a presença de heterogeneidade de variâncias e da correlação serial de forma direta. A desvantagem é que estudos com medidas repetidas, principalmente os longitudinais, freqüentemente apresentam dados omissos, o que impossibilita o ajuste de modelos com matrizes de covariância intra-unidades amostrais não estruturadas que a estratégia requer. A estratégia proposta por Diggle (1988), como observou Cúri (1999), não é adequada na presença de uma variável independente contínua, pois, neste caso, não existe um modelo saturado para os parâmetros de regressão, uma vez que este deve incorporar a forma da

relação entre as variáveis independentes e a variável resposta. Além destas desvantagens, a determinação do número de efeitos aleatórios que farão parte do modelo não é abordada e a escolha desses efeitos é fundamental para a explicação da variabilidade dos dados.

Neste trabalho, sugerimos uma estratégia baseada na proposta de Diggle et al. (2002), mas que considera: i) o tipo de planejamento envolvido no estudo, ii) a análise dos gráficos de perfis médios e individuais, iii) a análise da matriz de covariâncias intra-unidades amostrais quando é possível calculá-la, iv) a análise do gráfico de perfis das linhas da matriz de covariâncias amostral para a identificação de possíveis efeitos aleatórios, v) a análise do variograma amostral para identificação de correlação serial, vi) testes de hipóteses sobre os componentes de variância para verificar a inclusão ou não de efeitos aleatórios e vii) os critérios de informação AIC_R e BIC_R como ferramenta auxiliar na identificação da estrutura de covariâncias intra-unidades amostrais que melhor represente a variabilidade dos dados.

1.3 Inferência Estatística

Para a construção da função de verossimilhança ou da verossimilhança restrita consideraremos o modelo para todas as unidades amostrais, a saber:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{b} + \mathbf{e}, \quad (1.22)$$

em que $\mathbf{Y} = [y_1^t, \dots, y_N^t]^t$ ($\sum n_i \times 1$) é o vetor de respostas das unidades amostrais, $\mathbf{X} = [\mathbf{X}_1^t, \dots, \mathbf{X}_N^t]$ ($\sum n_i \times p$) é a matriz de especificação dos efeitos fixos, $\mathbf{Z} = \text{diag}[\mathbf{Z}_1^t, \dots, \mathbf{Z}_N^t]$ ($\sum n_i \times Nq$), é a matriz de especificação dos efeitos aleatórios. O vetor $\mathbf{b} = [b_1^t, \dots, b_N^t]^t$ ($Nq \times 1$) engloba os efeitos aleatórios, e $\mathbf{e} = [e_1^t, \dots, e_N^t]^t$ ($\sum n_i \times 1$) é o vetor de erros aleatórios. Portanto, sob esta formulação, $\mathbf{b} \sim N_{Nq}(\mathbf{0}, \boldsymbol{\Psi}(\boldsymbol{\theta}_1))$ em que $\boldsymbol{\Psi}(\boldsymbol{\theta}_1) = \mathbf{I}_N \otimes \mathbf{G}(\boldsymbol{\theta}_1)$, e \mathbf{b} é independente de $\mathbf{e} \sim N_{\sum n_i}(\mathbf{0}, \mathbf{R})$ com $\mathbf{R} = \text{diag}[\mathbf{R}_1^t, \dots, \mathbf{R}_N^t]$ e $\mathbf{R}_i = \mathbf{R}_i(\boldsymbol{\theta}_2)$. Conseqüentemente, $\mathbf{Y} \sim N_{\sum n_i}(\mathbf{X}\boldsymbol{\beta}, \mathbf{V}(\boldsymbol{\theta}))$, em que $\mathbf{V}(\boldsymbol{\theta}) = \mathbf{Z}\boldsymbol{\Psi}(\boldsymbol{\theta}_1)\mathbf{Z}^t + \mathbf{R}(\boldsymbol{\theta}_2)$.

O logaritmo da verossimilhança marginal dos dados é

$$l(\boldsymbol{\beta}, \boldsymbol{\theta}) = -\frac{N}{2} \ln 2\pi - \frac{1}{2} \ln |\mathbf{V}(\boldsymbol{\theta})| - \frac{1}{2} \boldsymbol{\varepsilon}^t \mathbf{V}(\boldsymbol{\theta})^{-1} \boldsymbol{\varepsilon}. \quad (1.23)$$

A metodologia de máxima verossimilhança gera estimadores não viciados para os efeitos fixos mas estimadores viciados para os parâmetros da matriz de covariância intra-unidades amostrais por não levar em consideração a perda de graus de liberdade na estimação dos efeitos fixos. A fim de reduzir o viés desses estimadores, muitos autores recomendam o uso do método de estimação por máxima verossimilhança restrita (MVR). Este método foi proposto por Patterson e Thompson (1971) para estimar componentes de variância e consiste em maximizar a verossimilhança de uma transformação linear do tipo $\mathbf{Y}^* = \mathbf{U}^t \mathbf{Y}$ em que \mathbf{U} tem dimensão $(\sum n_i \times (\sum n_i - p))$, e é tal que $E(\mathbf{Y}^*) = \mathbf{0}$, ou seja, $\mathbf{U}^t \mathbf{X} = \mathbf{0}$. Assim, a verossimilhança obtida a partir da transformação linear não depende dos efeitos fixos β . Uma maneira de se obter a verossimilhança restrita é considerar $\mathbf{U} = \mathbf{I} - \mathbf{X}(\mathbf{X}^t \mathbf{X})^{-1} \mathbf{X}^t$, que gera os resíduos do ajuste obtido por mínimos quadrados ordinários. Daí a denominação de verossimilhança residual empregada por alguns autores. Para os dados transformados, temos

$$\begin{aligned} E(\mathbf{Y}^*) &= \mathbf{0} \\ \text{Var}(\mathbf{Y}^*) &= \mathbf{U} \mathbf{V}(\boldsymbol{\theta}) \mathbf{U}^t, \end{aligned} \quad (1.24)$$

e além disso, $\mathbf{Y}^* \sim N_{\sum n_i - p}(\mathbf{0}, \mathbf{U}^t \mathbf{V}(\boldsymbol{\theta}) \mathbf{U})$.

O logaritmo da verossimilhança restrita marginal dos dados transformados é

$$l_R(\boldsymbol{\theta}) = -\frac{1}{2} \ln |\mathbf{V}(\boldsymbol{\theta})| - \frac{1}{2} \ln |\mathbf{X}^t \mathbf{V}(\boldsymbol{\theta})^{-1} \mathbf{X}| - \frac{N-p}{2} \boldsymbol{\varepsilon}^t \mathbf{V}(\boldsymbol{\theta})^{-1} \boldsymbol{\varepsilon} - \frac{N-p}{2} \ln(2\pi), \quad (1.25)$$

cuja maximização gera os estimadores de $\hat{\boldsymbol{\theta}}^t = (\hat{\sigma}^2, \hat{\boldsymbol{\theta}}_1^t, \hat{\boldsymbol{\theta}}_2^t)^t$. Em particular, o estimador de

$$\hat{\sigma}^2 = \frac{1}{N-p} \hat{\boldsymbol{\varepsilon}}^t [\mathbf{Z} \boldsymbol{\Psi}(\hat{\boldsymbol{\theta}}_1) \mathbf{Z}^t + \mathbf{R}(\hat{\boldsymbol{\theta}}_2)]^{-1} \hat{\boldsymbol{\varepsilon}}. \quad (1.26)$$

Os preditores para os efeitos aleatórios, $\hat{\mathbf{b}}_i$, são importantes em nosso estudo, pois, a partir deles, podemos obter os resíduos condicionais, $\hat{\boldsymbol{\varepsilon}} = \mathbf{Y} - \mathbf{X} \hat{\boldsymbol{\beta}} - \mathbf{Z} \hat{\mathbf{b}}$, que são úteis na escolha da estrutura de covariância intra-unidades amostrais. Henderson (1950, 1975, 1984) obtém conjuntamente os preditores dos efeitos aleatórios e os estimadores para os efeitos fixos por meio da função de distribuição conjunta dos efeitos aleatórios \mathbf{b} e das observações \mathbf{Y} .

A função densidade conjunta dos efeitos aleatórios e do vetor de observações pode ser expressa como

$$f(\mathbf{Y}, \mathbf{b}) = f(\mathbf{Y}|\mathbf{b})f(\mathbf{b}), \quad (1.27)$$

em que $f(\mathbf{Y}|\mathbf{b})$ e $f(\mathbf{b})$ são as funções densidade de $\mathbf{Y}|\mathbf{b}$ e \mathbf{b} respectivamente. Como $\mathbf{Y}|\mathbf{b} \sim N_{\sum n_i}(\mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{b}, \mathbf{R})$ e $\mathbf{b} \sim N_N(\mathbf{0}, \boldsymbol{\Psi})$, a função densidade conjunta é

$$f(\mathbf{Y}, \mathbf{b}) = C \exp\left[-\frac{1}{2}(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{Z}\mathbf{b})^t \mathbf{R}^{-1}(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{Z}\mathbf{b})\right] \exp\left[-\frac{1}{2}\mathbf{b}^t \mathbf{G}^{-1}\mathbf{b}\right]. \quad (1.28)$$

com $C = (2\pi)^{-\frac{1}{2}(N+q)} |\mathbf{R}|^{-\frac{1}{2}} |\mathbf{G}|^{-\frac{1}{2}}$.

Os estimadores dos efeitos fixos e dos preditores para os efeitos aleatórios são encontrados maximizando essa função em relação a $\boldsymbol{\beta}$ e a \mathbf{b} , o que pode ser feito resolvendo o seguinte sistema de equações proposto por Henderson (1975):

$$\begin{bmatrix} \mathbf{X}^t \hat{\mathbf{R}}^{-1} \mathbf{X} & \mathbf{X}^t \hat{\mathbf{R}}^{-1} \mathbf{Z} \\ \mathbf{Z}^t \hat{\mathbf{R}}^{-1} \mathbf{X} & \mathbf{Z}^t \hat{\mathbf{R}}^{-1} \mathbf{Z} + \hat{\mathbf{G}}^{-1} \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{\beta}} \\ \hat{\mathbf{b}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}^t \hat{\mathbf{R}}^{-1} \mathbf{Y} \\ \mathbf{Z}^t \hat{\mathbf{R}}^{-1} \mathbf{Y} \end{bmatrix}, \quad (1.29)$$

em que $\hat{\mathbf{R}}$ e $\hat{\mathbf{G}}$ são as estimativas das matrizes \mathbf{R} e \mathbf{G} . Resolvendo o sistema, obtemos os seguintes estimadores para os efeitos fixos e preditores para os efeitos aleatórios (detalhes poderam ser encontrados em Henderson (1975) e em Searle et al. (1992)):

$$\boldsymbol{\beta}(\hat{\boldsymbol{\theta}}) = \left(\sum_{i=1}^N \mathbf{X}_i^t \mathbf{V}_i(\hat{\boldsymbol{\theta}})^{-1} \mathbf{X}_i \right)^{-1} \sum_{i=1}^N \mathbf{X}_i^t \mathbf{V}_i(\hat{\boldsymbol{\theta}})^{-1} \mathbf{y}_i \quad (1.30)$$

$$\mathbf{b}_i(\hat{\boldsymbol{\theta}}) = \mathbf{G}(\hat{\boldsymbol{\theta}}) \mathbf{Z}_i^t \mathbf{V}_i(\hat{\boldsymbol{\theta}})^{-1} (\mathbf{y}_i - \mathbf{X}_i \boldsymbol{\beta}(\hat{\boldsymbol{\theta}})). \quad (1.31)$$

Os estimadores dos efeitos fixos coincidem com os de mínimos quadrados generalizados e são os melhores estimadores lineares não viciados (Best Linear Unbiased Estimators, BLUE); os preditores para os efeitos aleatórios são os melhores preditores lineares não viciados (Best Linear Unbiased Predictors, BLUP) e também coincidem com aqueles obtidos por métodos de Bayes empíricos (Robinson (1991)). Mais detalhes sobre a teoria Bayesiana para esse tipo de modelos lineares podem ser encontrados em Lindley e Smith (1972) e Smith (1973); detalhes sobre os BLUP, em Robinson (1991). Durante o processo de maximização do logaritmo da verossimilhança restrita, a $\hat{\mathbf{G}}$ poderá tornar-se

singular. Neste caso, Henderson (1984) propõe modificar o sistema de equações, usando a decomposição de Cholesky de $\hat{\mathbf{G}} = \hat{\mathbf{L}}\hat{\mathbf{L}}^t$, em que, $\hat{\mathbf{L}}$ é uma matriz triangular inferior de dimensão $(Nq \times Nq)$, obtendo o seguinte sistema:

$$\begin{bmatrix} \mathbf{X}^t\hat{\mathbf{R}}^{-1}\mathbf{X} & \mathbf{X}^t\hat{\mathbf{R}}^{-1}\mathbf{Z}\hat{\mathbf{L}} \\ \hat{\mathbf{L}}^t\mathbf{Z}^t\hat{\mathbf{R}}^{-1}\mathbf{X} & \hat{\mathbf{L}}^t\mathbf{Z}^t\hat{\mathbf{R}}^{-1}\mathbf{Z}\hat{\mathbf{L}} + \mathbf{I} \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{\beta}} \\ \hat{\mathbf{b}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}^t\hat{\mathbf{R}}^{-1}\mathbf{Y} \\ \hat{\mathbf{L}}^t\mathbf{Z}^t\hat{\mathbf{R}}^{-1}\mathbf{Y} \end{bmatrix}. \quad (1.32)$$

As matrizes de covariâncias assintóticas de $\hat{\boldsymbol{\beta}}$ e $\hat{\mathbf{b}}$, podem ser obtidas por meio da teoria de grandes amostras e são apresentadas em Laird e Ware (1982). Elas são casos especiais dos modelos gerais obtidos por Harville (1976) e podem ser expressas como

$$V(\boldsymbol{\beta}(\hat{\boldsymbol{\theta}})) = \left(\sum_{i=1}^N \mathbf{X}_i^t \mathbf{V}_i(\boldsymbol{\theta})^{-1} \mathbf{X}_i \right)^{-1} \quad (1.33)$$

$$V(\hat{\mathbf{b}}_i(\hat{\boldsymbol{\theta}})) = \mathbf{G}(\boldsymbol{\theta}) \mathbf{Z}_i^t (\mathbf{V}_i(\boldsymbol{\theta})^{-1} - \mathbf{V}_i(\boldsymbol{\theta})^{-1} \mathbf{X}_i V(\hat{\boldsymbol{\beta}}(\boldsymbol{\theta})) \mathbf{X}_i^t \mathbf{V}_i(\boldsymbol{\theta})^{-1}) \mathbf{Z}_i \mathbf{G}(\boldsymbol{\theta}), \quad (1.34)$$

estimativas dessas matrizes podem ser obtidas substituindo-se $\boldsymbol{\theta}$ por sua estimativa $\hat{\boldsymbol{\theta}}$.

Como a expressão para a variância assintótica dos preditores dos efeitos aleatórios $\hat{\mathbf{b}}_i$ não leva em consideração a variabilidade em torno de \mathbf{b}_i , Laird e Ware (1982) propõem que a variância assintótica dos preditores aleatórios seja calculada por

$$V[\mathbf{b}_i(\hat{\boldsymbol{\theta}}) - \mathbf{b}_i] = \mathbf{G}(\boldsymbol{\theta}) - V(\mathbf{b}_i(\hat{\boldsymbol{\theta}})). \quad (1.35)$$

Mais detalhes sobre as propriedades assintóticas dos estimadores para os efeitos fixos e aleatórios podem ser encontrados em Vonesh e Chinchilli (1997).

Uma vez estimados os parâmetros da matriz de covariâncias intra-unidades amostrais através da maximização da verossimilhança restrita marginal (1.14), as estimativas dos efeitos fixos $\hat{\boldsymbol{\beta}}$ e as predições dos efeitos aleatórios $\hat{\mathbf{b}}_i$ serão obtidas resolvendo-se o sistema de equações (1.15).

O estimador do vetor $\boldsymbol{\theta}$ tem uma forma fechada para dados completos e balanceados e uma estrutura de covariância de independência condicional para o resíduo condicional (ver Graybill (1976), Laird et al. (1987) ou Singer e Andrade (2000)). Como em geral, estaremos trabalhando com dados omissos ou não balanceados, métodos iterativos são necessários para maximizar o logaritmo da verossimilhança restrita marginal.

Dentre os métodos de otimização podemos citar o método de Newton-Raphson, o método Scoring de Fisher, o método Simplex proposto por Nelder e Mead (1965), o método EM proposto por Laird e Ware (1982) ou o método baseado em espaço de estados usando filtro de Kalman proposto por Jones (1993).

O processo de Newton-Raphson com as modificações sugeridas por Jennrich e Schluchter (1986), Lindstrom e Bates (1988) e Wolfinger et al. (1994) é melhor que os demais em relação ao tempo total para a convergência e por este motivo será utilizado neste trabalho.

O algoritmo descrito em Wolfinger et al. (1994) para a otimização do logaritmo da verossimilhança ou do logaritmo da verossimilhança restrita para modelos mistos, tem a vantagem de poder ser aplicado a modelos com qualquer estrutura de covariância tanto para os efeitos aleatórios quanto para o erro residual. Ele faz uso da decomposição de Cholesky, do operador "sweep" e da transformação W (W-transformation (ver Goodnight e Hemmerle (1979))) para torná-lo mais estável e rápido. Detalhes teóricos sobre os métodos utilizados e suas aplicações são encontrados em Laird e Ware (1982), Jennrich e Schluter (1986), Laird et al. (1987), Lindstrom e Bates (1988), Andreoni (1989), Jones (1993) e Wolfinger et al. (1994), Lange (1999) e Singer e Andrade (2000).

O algoritmo poderá ser implementado da seguinte maneira: dados um valor inicial conveniente θ^0 , a função score $\mathbf{U}(\theta) = \partial l_R(\theta)/\partial \theta$ e a matriz hessiana associada, $\mathbf{H}(\theta) = \partial^2 l_R(\theta)/\partial \theta \partial \theta^t$, consideramos

$$\theta^{k+1} = \theta^k - \mathbf{H}(\theta^k)^{-1} \mathbf{U}(\theta^k), k = 0, 1, 2, \dots, \quad (1.36)$$

se utilizarmos o algoritmo de Newton-Raphson ou

$$\theta^{k+1} = \theta^k + \mathbf{I}(\theta^k)^{-1} \mathbf{U}(\theta^k), k = 0, 1, 2, \dots, \quad (1.37)$$

em que $\mathbf{I}(\theta) = \mathbb{E}[-(\mathbf{H}(\theta))^2]$ é matriz de informação de Fisher, se utilizarmos o algoritmo "Scoring" de Fisher. As iterações são repetidas até que algum critério de convergência seja alcançado, por exemplo, $\|\theta^{k+1} - \theta^k\| \leq \varepsilon$ para algum $\varepsilon > 0$. Para detalhes ver Laird e Ware (1982), Jennrich e Schluter (1986), Laird et al. (1987), Lindstrom e Bates (1988),

Andreoni (1989), Jones (1993) e Wolfinger et al. (1994), Lange (1999) e Singer e Andrade (2000).

Verbeke e Molenberghs (1997) comentam que podem ocorrer problemas de convergência durante o processo de estimação dos componentes de variância. Isto acontece quando os parâmetros estão próximos da fronteira do espaço paramétrico. Para contornar esse problema sugerem uma transformação na escala das condições de avaliação, como por exemplo, dividindo as condições de avaliação por uma constante maior que 1 com o objetivo de aumentar os valores dos parâmetros de covariância estimados, tornando as equações normais mais estáveis durante o processo de maximização.

1.4 Teste de Hipótese para Componentes de Variância

Estudos com medidas repetidas completos, podem ser comparados a experimentos com blocos aleatórios e a experimentos do tipo “split-plot”. Em estudos completos ou incompletos, os testes de hipóteses sobre componentes de variância são realizados via análise de variância (ANOVA) como em Searle (1971, 411-413). A distribuição das estatísticas de teste é F no caso de dados completos e assintoticamente F no caso de dados incompletos.

Öfversten (1993) apresenta um método para obter testes exatos para componentes de variância para alguns modelos mistos não balanceados. Sua abordagem é baseada em uma transformação ortogonal. O modelo transformado define um novo vetor de componentes de erro que é usado para balancear o modelo originalmente não balanceado. Como consequência, são obtidos testes exatos análogos àqueles tradicionalmente usados em modelos de efeitos aleatórios balanceados.

Testes de hipótese sobre componentes de variância baseados na verossimilhança apresentam problemas com relação às condições de regularidade (ver Cox e Hinkley 1996, Cap. 9). O problema surge porque a primeira condição de regularidade determina que o espaço paramétrico Ω deve ter dimensão finita, ser fechado e compacto e que o verdadeiro valor do parâmetro deverá estar no interior de Ω . Se no modelo em estudo quisermos

avaliar a necessidade da inclusão do efeito aleatório $b_i \sim N(0, \sigma_a^2)$, uma hipótese de interesse é:

$$\begin{aligned} H_0 : \sigma_a^2 &= 0 \\ H_1 : \sigma_a^2 &\in (0, \infty). \end{aligned} \tag{1.38}$$

A estatística do teste da razão de verossimilhanças é dada por:

$$RV = -2\ln\left[\frac{\sup[L_{H_0}]}{\sup[L_{H_1}]}\right], \tag{1.39}$$

em que $\sup[L_{H_0}]$ é o valor máximo da verossimilhança sob a hipótese nula e $\sup[L_{H_1}]$ é o valor máximo da verossimilhança sob a hipótese alternativa. Nesses casos, a estatística da razão de verossimilhanças sob a hipótese nula não seguirá uma distribuição assintótica qui-quadrado, pois a hipótese nula coloca o valor do parâmetro na fronteira do espaço paramétrico.

O problema ocorre porque a dedução da distribuição da estatística de teste exige uma expansão em série de Taylor numa vizinhança do verdadeiro valor do parâmetro que deve estar no interior do espaço paramétrico. Como consequência, a estatística da RV terá uma distribuição degenerada na fronteira, ou seja, uma distribuição que assume o valor zero com probabilidade positiva, em contraposição à distribuição obtida quando as condições de regularidades estão satisfeitas.

Self e Liang (1987) mostraram que nesses casos a distribuição da estatística RV é uma mistura de distribuições qui-quadrado. Stram e Lee (1994) discutem o comportamento da estatística RV para componentes de variância em estudos longitudinais com efeitos aleatórios, aplicando o método descrito por Self e Liang (1987), e particularizam os resultados para várias situações de testes de hipóteses. De uma maneira geral, para testar hipótese sobre a inclusão de um efeito aleatório em um modelo com q efeitos aleatórios, a estatística do teste deverá ser comparada a uma mistura 50%:50% de distribuições qui-quadrado, χ_q^2 e χ_{q+1}^2 . O índice $q + 1$ representa o número de parâmetros adicionados ao modelo sob a hipótese alternativa, H_1 , e o nível descritivo é dado por $p = 0.5P(\chi_q^2 > -2 \ln RV) + 0.5P(\chi_{q+1}^2 > -2 \ln RV)$.

Giampaoli (1999) observou que Stram e Lee (1994) não levaram em consideração o fato de que em estudos longitudinais as variáveis não são identicamente distribuídas; este mesmo problema ocorre de uma maneira geral em estudos com medidas repetidas. Giampaoli (1999) também estudou testes de hipóteses sobre os elementos da matriz de covariância num modelo de efeitos aleatórios utilizando a estatística da razão de verossimilhanças apresentadas por Vu e Zhou (1997). Esses autores estendem os resultados obtidos por Self e Liang (1987) para o caso em que as variáveis aleatórias não são identicamente distribuídas. Giampaoli (1999) mostrou que nos modelos com até dois efeitos aleatórios, mesmo após as devidas adaptações, os resultados de Self e Liang (1987) continuam válidos.

1.5 Critérios de Informação

Como uma ferramenta auxiliar para a seleção das estruturas de covariâncias intra-unidades amostrais, consideramos os critérios de informação de Akaike (AIC_R) e de Schwarz (BIC_R) definidos por

$$AIC_R = -2l(\hat{\theta}) + 2q \quad (1.40)$$

$$BIC_R = -2l(\hat{\theta}) + q \log(n - p), \quad (1.41)$$

em que $l(\hat{\theta})$ é o máximo da função log-verossimilhança restrita e q é o número de parâmetros da matriz de covariância, p é o número de efeitos fixos e $\sum_{i=1}^N n_i = n$ é o número efetivo de observações. A estrutura que apresentar os menores valores de AIC_R e BIC_R será a mais adequada. Este critério é conveniente por permitir a comparação de quaisquer estruturas de covariâncias, exigindo apenas que o modelo para os parâmetros de localização seja o mesmo. Mais detalhes sobre o uso deste critério na seleção de estruturas de covariância e estudos com medidas repetidas são encontrados em Keselman et al. (1998).

Capítulo 2

Modelagem da Estrutura de Covariância

2.1 Introdução

Neste capítulo, o objetivo é propor um conjunto de técnicas gráficas e analíticas que auxiliem a escolha das matrizes Z_i , G e R_i do modelo (1.1). Dentro desse conjunto, existem técnicas que podem ser utilizadas em estudos longitudinais ou não, que denominaremos Técnicas Gerais e técnicas que dependem do tipo de estudo que serão denominadas de Técnicas Específicas. Para ilustrar a utilização de tais técnicas consideramos dados hipotéticos com medidas repetidas.

A escolha dessas técnicas depende de informações sobre o tipo de estudo com medidas repetidas (longitudinal ou não), sobre o planejamento (balanceado ou não) e sobre as condições de avaliação (dados completos ou não). Estudos não longitudinais, em geral, têm caráter experimental e a correspondente estrutura de covariância intra-unidades amostrais pode ser modelada com o auxílio de informações sobre o planejamento. Para estudos longitudinais, é mais razoável utilizar informações sobre o comportamento da resposta ao longo das condições de avaliação na modelagem da estrutura de covariância intra-unidades amostrais.

2.2 Técnicas Gerais

Para experimentos com medidas repetidas completos, sugerimos inicialmente considerar a análise da matriz de covariâncias e correlações amostrais para identificar: i) se os valores das variâncias ao longo das condições de avaliação são diferentes, indicando heterocedasticidade; ii) se os valores das correlações diminuem ao longo das condições de avaliação indicando uma possível correlação serial, ou ainda; iii) se os valores das correlações em cada sub-diagonal (cada diagonal acima da diagonal principal) das matrizes de correlações, apresentam um comportamento crescente, decrescente ou ambos. Para estudos com medidas repetidas com mais de um tratamento, sugerimos também a construção dessas matrizes para cada tratamento, possibilitando não só a identificação de heterocedasticidade para as observações intra-unidades amostrais, como também heterocedasticidade entre observações de tratamentos diferentes. Para ilustrar a utilização da análise da matriz de covariâncias e correlações amostrais, dos gráficos de perfis individuais das observações e das observações centralizadas e o variograma com o propósito de selecionar a estrutura de covariâncias, geramos 40 observações dados que simulam um estudo com medidas repetidas com obtidas sob cinco condições de avaliação denominadas t1, t2, t3, t4 e t5. Detalhes geração de sobre números aleatórios ver Lange (1999).

Por exemplo, consideremos um experimento com medidas repetidas e completo, em que as unidades amostrais sejam divididas em dois grupos de tratamentos com N unidades amostrais cada e que as observações de cada unidade amostral sejam obtidas sob cinco condições de avaliação. Suponhamos também que as matrizes de covariâncias e correlações amostrais para cada grupo sejam dadas por

Matriz de variâncias e correlações amostrais hipotéticas para o tratamento 1

$$V_{01} = \begin{bmatrix} 1.03 & 0.94 & 0.80 & 0.56 & 0.41 \\ & 68.10 & 0.91 & 0.72 & 0.64 \\ & & 102.98 & 0.87 & 0.74 \\ & & & 16.04 & 0.87 \\ & & & & 237.57 \end{bmatrix}. \quad (2.1)$$

Matriz de variâncias e correlações amostrais hipotéticas para o tratamento 2

$$V_{02} = \begin{bmatrix} 5.31 & 0.92 & 0.76 & 0.70 & 0.54 \\ & \mathbf{73.82} & 0.90 & 0.85 & 0.73 \\ & & \mathbf{196.27} & 0.98 & 0.88 \\ & & & \mathbf{142.15} & 0.93 \\ & & & & \mathbf{165.16} \end{bmatrix}. \quad (2.2)$$

As matrizes V_{01} e V_{02} acima apresentam: i) variâncias (em negrito), com valores distintos em cada condição de avaliação; ii) correlações positivas e decrescentes ao longo de suas colunas; iii) em cada sub-diagonal (cada diagonal acima da diagonal principal) as correlações têm valores próximos. Além disso, os valores das variâncias amostrais na matriz V_{01} são consideravelmente diferentes dos valores das variâncias na matriz V_{02} .

Essas características sugerem que o modelo para estrutura de covariâncias deverá incorporar heterocedasticidade dentro de cada tratamento, entre os tratamentos além de correlação serial. Possíveis modelos para a estrutura de covariâncias intra-unidades amostrais são, por exemplo: o modelo auto-regressivo heterogêneo (ARH1), o modelo com efeitos aleatórios mais erro aleatório com um comportamento auto-regressivo, ou ainda, o modelo com efeitos aleatórios com independência condicional. Para modelar a heterocedasticidade entre grupos é necessário considerar uma estrutura de covariâncias intra-unidades amostrais com parâmetros diferentes para cada grupo; por exemplo, podemos ajustar um modelo auto-regressivo heterogêneo sem efeito aleatório para cada grupo.

Para exemplificar outros comportamentos, consideraremos um caso com apenas um tratamento. Em estudos com medidas repetidas não longitudinais, a ordem de observação da variável resposta nas condições de avaliação é aleatória, o que sugere que as correlações das medidas realizadas na mesma unidade amostral devam ser constantes. Um exemplo desse padrão é apresentado na matriz de variâncias e correlações hipotética

abaixo.

$$V_0 = \begin{bmatrix} \mathbf{13.55} & 0.82 & 0.80 & 0.58 & 0.82 \\ & \mathbf{12.51} & 0.89 & 0.84 & 0.81 \\ & & \mathbf{12.82} & 0.84 & 0.82 \\ & & & \mathbf{14.18} & 0.65 \\ & & & & \mathbf{12.18} \end{bmatrix} \quad (2.3)$$

Podemos observar que as correlações não apresentam um padrão de decrescimento ou crescimento e os valores das variâncias (em negrito) estão próximos, indicando assim, que o modelo uniforme é adequado para a matriz de covariâncias intra-unidades amostrais. Se as variâncias fossem muito diferentes e as correlações não variassem muito, um modelo uniforme heterogêneo poderia ser adotado.

Um padrão freqüente em matrizes de covariâncias e correlações amostrais de estudos longitudinais é

$$V_0 = \begin{bmatrix} \mathbf{52.70} & \underline{0.83} & 0.68 & 0.64 & 0.59 \\ & \mathbf{85.95} & \underline{0.90} & 0.85 & 0.81 \\ & & \mathbf{97.16} & \underline{0.95} & 0.95 \\ & & & \mathbf{137.75} & \underline{0.99} \\ & & & & \mathbf{185.23} \end{bmatrix} \quad (2.4)$$

As variâncias (em negrito) na diagonal principal tendem a aumentar ao longo das condições de avaliação; as correlações são positivas e parecem decrescer monotonicamente ao longo das colunas da matriz. O padrão que a torna peculiar é o das correlações nas sub-diagonais, que tendem a crescer ao longo das condições de avaliação, como podemos observar a partir das correlações amostrais sublinhadas na primeira sub-diagonal da matriz (2.4). Um possível modelo para a estrutura de covariâncias intra-unidades amostrais é o de antedependência, como sugerem Zimmerman e Núñez (2000 e 2001).

Em estudos com medidas repetidas incompletos mas com poucos dados omissos, ainda podemos considerar a análise das matrizes de covariâncias e correlações amostrais. Neste caso, excluimos da construção dessas matrizes as unidades amostrais que não apresentam todas as observações. Quando o número dessas unidades amostrais é elevado,

sugerimos considerar ferramentas gráficas e técnicas específicas de acordo com o tipo de estudo (longitudinal ou não).

Uma importante ferramenta gráfica na análise descritiva de estudos com medidas repetidas longitudinais são os gráficos de perfis individuais e médios. O gráfico de perfis médios é um gráfico de linhas conectando as respostas médias ao longo das condições de avaliação, enquanto o gráfico de perfis individuais é um gráfico de linhas, em que, cada linha representa o comportamento da variável resposta para cada unidade amostral ao longo das condições de avaliação. O gráfico de perfis individuais fornece informação sobre o comportamento da variância ao longo das condições de avaliação, sobre possíveis efeitos aleatórios e sobre o comportamento da variável resposta para cada unidade amostral ao longo das condições de avaliação.

Os gráficos das Figuras 2.1 e 2.2 mostram perfis individuais com respostas iniciais bem diferentes e com evolução razoavelmente paralela ao longo das condições de avaliação. Um possível modelo para o problema seria $y_{ij} = \alpha + \beta t_j + a_i + \varepsilon_{ij}$, em que a_i representa o efeito aleatório da i -ésima unidade amostral, sugerindo que as curvas para as unidades amostrais são retas com mesma inclinação, porém com interceptos diferentes. O gráfico da Figura 2.3 mostra perfis individuais com pouca variabilidade para as respostas iniciais e muita variabilidade para as inclinações. Nesta situação, um modelo seria $y_{ij} = \alpha + \beta t_j + b_i t_j + \varepsilon_{ij}$, em que b_i representa o efeito aleatório do coeficiente angular associado à i -ésima unidade amostral. Os perfis são retas com mesmo coeficiente linear mas com coeficientes angulares diferentes. Além disso, as variâncias mudam ao longo das condições de avaliação. No gráfico da Figura 2.4, os perfis individuais apresentam valores iniciais variáveis e inclinações diferentes. Nesta situação, temos novamente evidência de heterocedasticidade e um possível modelo seria $y_{ij} = \alpha + \beta t_j + a_i + b_i t_j + \varepsilon_{ij}$, em que a_i e b_i representam respectivamente os efeitos aleatórios dos coeficientes linear e angular associados à i -ésima unidade amostral. No gráfico da Figura 2.5 os perfis individuais parecem apresentar valores respostas e curvaturas diferentes. Os candidatos a efeitos aleatórios são os coeficientes lineares, angulares e quadráticos. Um possível modelo seria $y_{ij} = \alpha + \beta t_j + \gamma t_j^2 + a_i + b_i t_j + c_i t_j^2 + \varepsilon_{ij}$, em que a_i , b_i e c_i representam respectivamente os

efeitos aleatórios dos coeficientes linear, angular e quadrático associados à i -ésima unidade amostral.

Figura 2.1: Gráfico de perfis sugerindo modelo com coeficiente linear aleatório

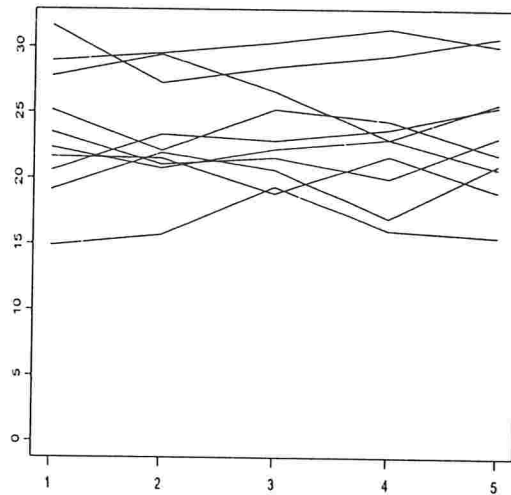


Figura 2.2: Gráfico de perfis sugerindo modelo com coeficiente linear aleatório

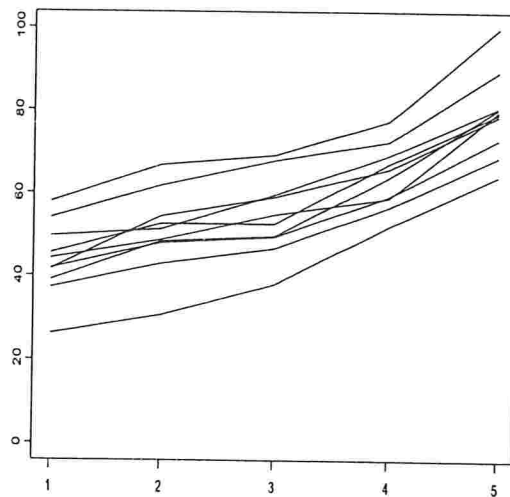


Figura 2.3: Gráfico de perfis sugerindo modelo com coeficiente angular aleatório

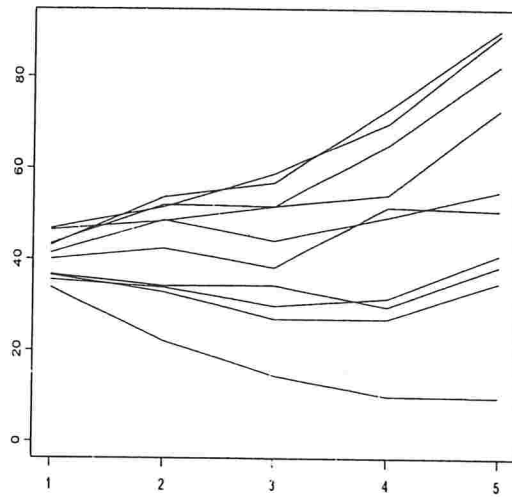


Figura 2.4: Gráfico de perfis sugerindo modelo com coeficientes linear e angular aleatórios

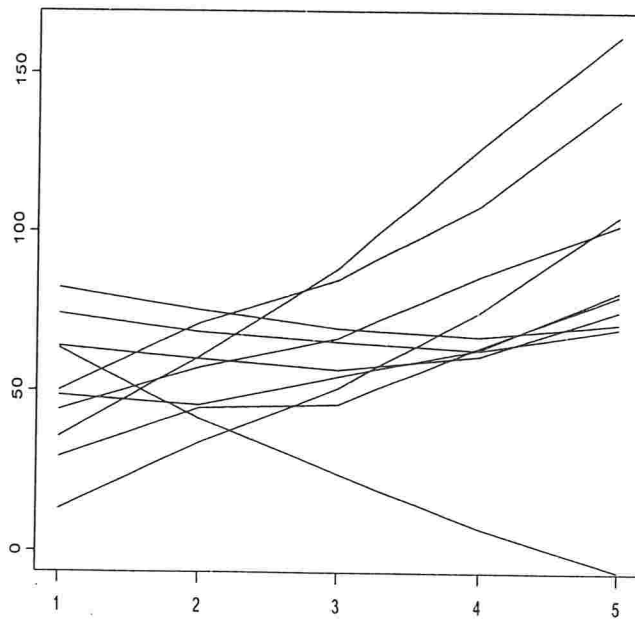
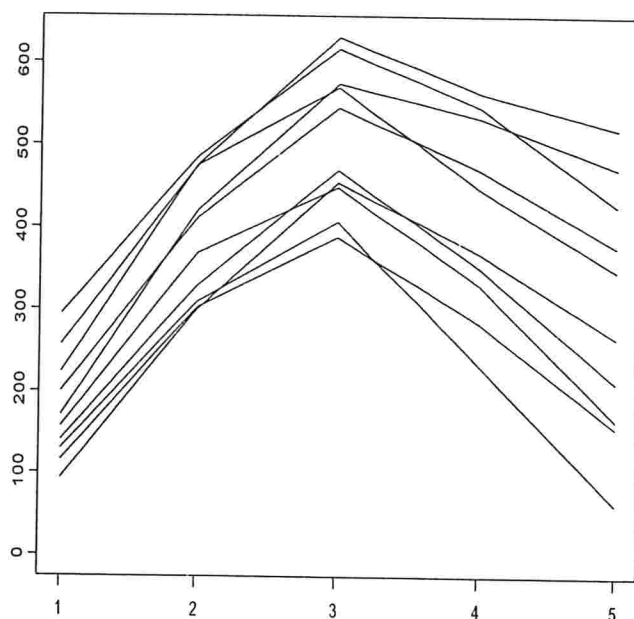


Figura 2.5: Gráfico de perfis sugerindo modelo com coeficiente linear, angular e quadrático aleatórios



Dawson et al. (1997) sugerem que o gráfico de perfis seja construído com as observações centralizadas em torno da média para diagnosticar heterocedasticidade. As observações centralizadas correspondem a $y_{ij}^* = y_{ij} - \bar{y}_j$, em que y_{ij} representa o valor da variável resposta na i -ésima unidade amostral na j -ésima condição de avaliação e \bar{y}_j representa a média das observações sob a j -ésima condição de avaliação. Os gráficos das Figuras 2.6 e 2.7 apresentam exemplos de perfis individuais de observações centralizadas que sugerem respectivamente heterocedasticidade e homocedasticidade. Dawson et al. (1997) sugerem também que o estudo do comportamento das correlações deva ser feito retirando-se não só o efeito da média, como também o da variância, padronizando as observações pelo desvio padrão amostral de cada condição de avaliação. Esse procedimento permite a análise das correlações por meio do gráfico de dispersão múltiplo, que será discutido na seção 2.3.2.

Uma outra ferramenta gráfica para identificar correlação serial é o variograma

Figura 2.6: Gráfico de perfis centrados sugerindo heterocedasticidade

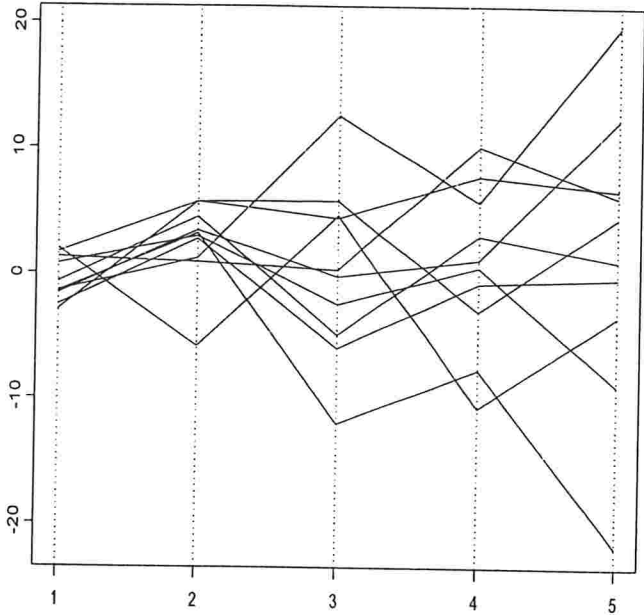
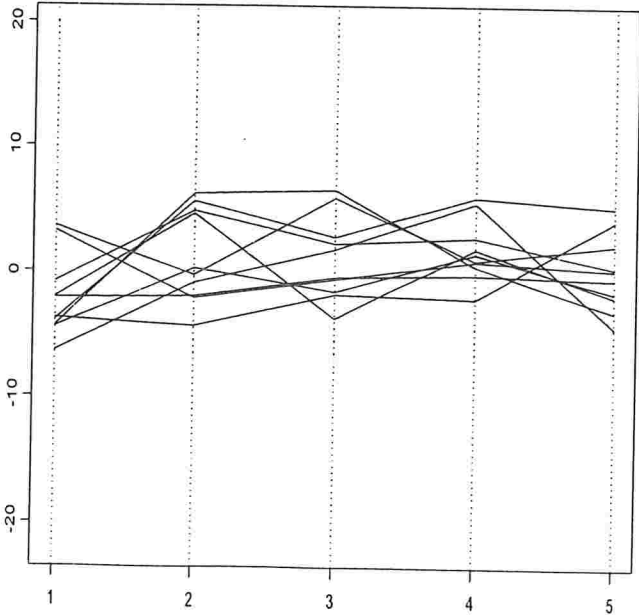


Figura 2.7: Gráfico de perfis centrados sugerindo homocedasticidade



amostral (ver Diggle (1988 e 1990) e Diggle et al. (2000)). A definição do variograma para o processo estocástico estacionário $\{Y_t, t \in \mathbb{R}\}$ é dada por

$$g(u) = \frac{1}{2} \mathbb{E}\{[Y(t) - Y(t - u)]^2\}, \quad u \geq 0.$$

Seja $\gamma(u) = \text{Cov}[Y(t), Y(t - u)]$; então o variograma pode ser escrito como

$$g(u) = \gamma(0) - \gamma(u) = \sigma^2[1 - \rho(u)].$$

Diggle (1988 e 1990) e Diggle et al. (2000) propõem a construção do variograma amostral a partir dos resíduos de mínimos quadrados ordinários, obtidos do ajuste de um modelo saturado para os parâmetros de regressão. Sua proposta tem a desvantagem de só poder ser aplicada quando os dados são homocedásticos. Neste trabalho consideraremos a construção do variograma das observações padronizadas, ou seja, o variograma construído com base em

$$\Delta_{ij} = \frac{y_{ij}^*}{s_j},$$

em que s_j representa o desvio padrão amostral associado à j -ésima condição de avaliação, o que torna o processo $\{\Delta_{ij}\}$ estacionário em relação à média e à variância.

Denotando as observações padronizadas da i -ésima unidade amostral sob a j -ésima condição de avaliação por Δ_{ij} , os pontos componentes do variograma amostral são calculados a partir de duas observações da mesma unidade amostral como

$$v_{ijk} = \frac{1}{2}(\Delta_{ij} - \Delta_{ik})^2.$$

A partir desses pontos, construímos o gráfico de dispersão de v_{ijk} em função da distância entre as condições de avaliação $u_{ijk} = |t_{ij} - t_{ik}|$. O variograma $g(u)$ pode ser estimado a partir da média dos valores v_{ijk} para um mesmo valor de u_{ijk} quando o estudo for regular no tempo, e a partir do ajuste de uma curva não paramétrica para os pares (u_{ijk}, v_{ijk}) ($j < k = 1, \dots, n_i$) em estudos irregulares. A variância amostral pode ser estimada por

$$\hat{\sigma}^2 = \frac{1}{2n} \sum_{i \neq l} \frac{1}{2} (\Delta_{ij} - \Delta_{lk})^2,$$

sendo n o número de pares de observações obtidas em unidades amostrais diferentes.

Em geral, é interessante incluir no variograma uma reta horizontal com ordenada igual ao valor da variância amostral, $\hat{\sigma}^2$. Como a correlação pode ser escrita na forma $\rho(u) = 1 - g(u)/\sigma^2$, valores de $\hat{g}(u)$ próximos de $\hat{\sigma}^2$, indicam correlação quase nula. A construção do variograma amostral não requer que o estudo seja regular, balanceado ou completo.

Os gráficos das Figuras 2.8, 2.9 e 2.10 são variogramas amostrais que sugerem situações com correlação serial, correlação constante e correlação nula, respectivamente.

Figura 2.8: Variograma amostral para dados simulados e padronizados com correlação serial

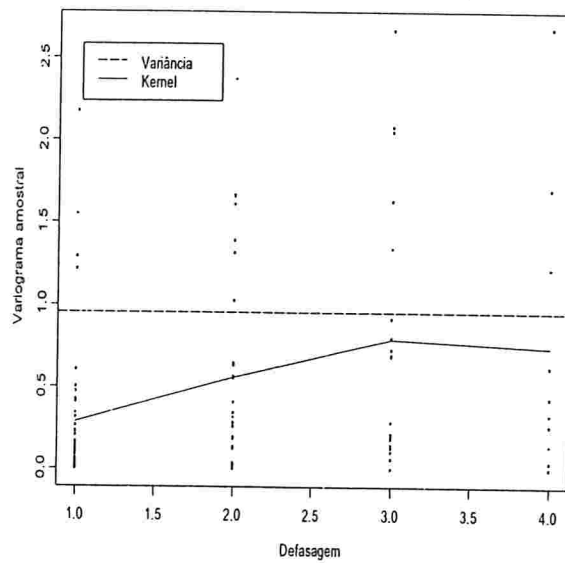


Figura 2.9: Variograma amostral para dados simulados e padronizados com correlação constante

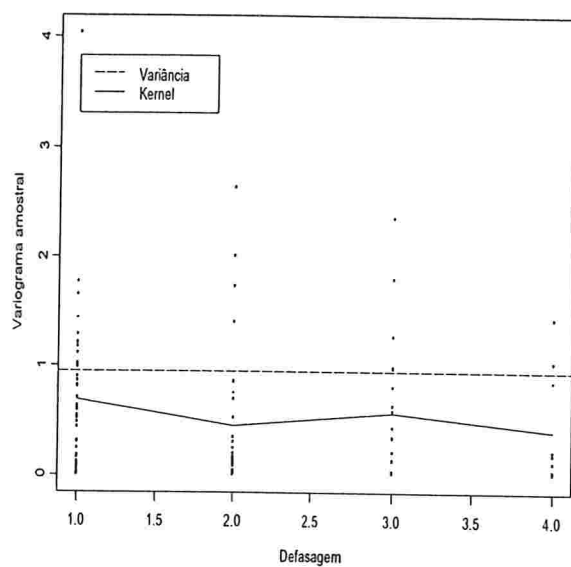
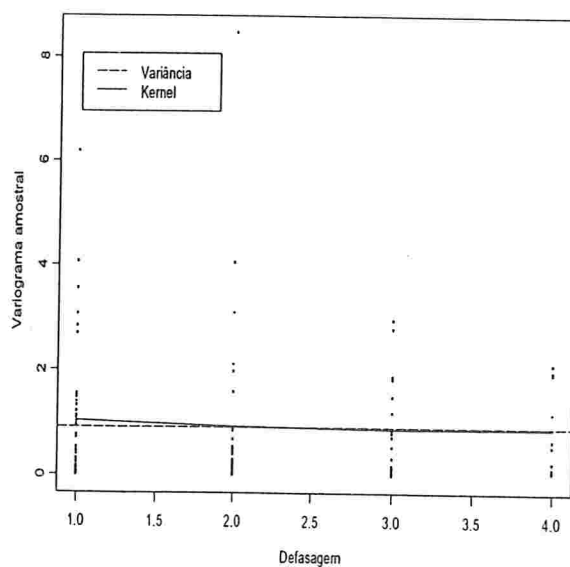


Figura 2.10: Variograma amostral para dados simulados e padronizados com correlação nula



Para situações em que o planejamento apresenta covariáveis dependentes do

tempo é importante considerar o ajuste de mínimos quadrados ordinários de um modelo saturado nos parâmetros de regressão para eliminar os efeitos das covariáveis e das médias. No caso de situações com dados completos ou com poucos dados omissos. Para propor modelos para a matriz de covariância dos dados, realizamos a análise das matrizes de covariâncias e correlações, dos gráficos de perfis individuais e a análise do variograma. A partir dos resíduos obtidos pelo ajuste do modelo acima.

No caso de estudos com medidas repetidas e covariáveis dependentes do tempo com muitos dados omissos, apenas a análise do gráficos de perfis individuais sem a influência da covariável, o gráfico de perfis individuais dos resíduos do modelo ajustado e o variograma para os resíduos padronizados do modelo ajustado podem ser considerados para diagnosticar possíveis modelos para a estrutura de covariâncias. Ajustando-se modelos saturados nos parâmetros de regressão e impondo cada estrutura de covariância proposta, pode-se escolher a mais adequada minimizando-se os critérios de informação AIC_R e BIC_R já discutidos no Capítulo 1.

2.3 Técnicas Específicas

2.3.1 Estudos Não Longitudinais

Em geral, tais estudos têm caráter experimental e constituem estudos completos ou com poucos dados omissos, facilitando a análise das matrizes de covariâncias e correlações amostrais. Além disso, por serem similares a estudos em blocos aleatórios ou do tipo split-plot, o planejamento pode fornecer informação sobre possíveis efeitos aleatórios importantes na modelagem da matriz de covariâncias. Neste contexto, possíveis modelos são o modelo uniforme e o modelo uniforme heterogêneo. No caso de estudos com dois fatores intra-unidades amostrais, uma possível estrutura é obtida a partir do produto de Kronecker entre as possíveis matrizes de covariâncias dos fatores intra-unidades amostrais como o modelo proposto por Galecki (1994). Por exemplo, consideremos um estudo com n unidades amostrais avaliadas em dois dias, e em cada dia sob cada um de três tratamentos

aleatoriamente atribuídos a cada unidade amostral. A natureza experimental desse estudo sugere uma estrutura de covariância uniforme para as medidas obtidas em dias diferentes na mesma unidade amostral e uma matriz de covariâncias não estruturada para as medidas realizadas no mesmo dia e na mesma unidade amostral sob os três tratamentos. A matriz de covariâncias intra-unidades amostrais é o produto de Kronecker entre uma matriz com estrutura uniforme e uma não estruturada. Essa estrutura pode ser implementada por intermédio do modelo com $\mathbf{G} = \mathbf{0}$ e $\mathbf{R}_i = \text{Uniforme} \otimes NE$.

Outra alternativa é considerar modelos que incorporem efeitos aleatórios sugeridos pelo planejamento. Um exemplo interessante é apresentado em Sef (1999), em que o objetivo era comparar a eficácia de 4 tratamentos quanto ao índice de placa bacteriana. Participaram do estudo 32 crianças entre 4 e 6 anos de idade, divididas em 4 grupos de 8 crianças. Todos os grupos foram submetidos aos 4 tratamentos com um intervalo de uma semana para eliminar possíveis efeitos residuais. A cada grupo foi atribuído uma ordem diferente de aplicação dos tratamentos. Esse exemplo corresponde a um estudo do tipo pré-teste/pós-teste envolvendo uma estrutura fatorial cruzada em blocos com dois fatores intra-unidades amostrais: tipo de escova (convencional ou monobloco) e dentifrício (com ou sem). Como todas as crianças foram submetidas aos 4 tratamentos, espera-se que as medidas realizadas na mesma criança sejam correlacionadas. Um possível modelo para a estrutura de covariâncias intra-unidades amostrais considera independência condicional e três efeitos aleatórios não correlacionados: criança, interação criança \times escova, interação criança \times dentifrício. Denotando por σ_c^2 , σ_e^2 e σ_d^2 as variâncias dos efeitos aleatórios correspondentes, a matriz de covariâncias intra-unidades amostrais é modelada por

$$\mathbf{V}_i(\boldsymbol{\theta}) = \begin{bmatrix} \sigma_c^2 + \sigma_e^2 + \sigma_d^2 & \sigma_c^2 + \sigma_e^2 & \sigma_c^2 + \sigma_d^2 & \sigma_c^2 \\ \sigma_c^2 + \sigma_e^2 & \sigma_c^2 + \sigma_e^2 + \sigma_d^2 & \sigma_c^2 & \sigma_c^2 + \sigma_d^2 \\ \sigma_c^2 + \sigma_d^2 & \sigma_c^2 & \sigma_c^2 + \sigma_e^2 + \sigma_d^2 & \sigma_c^2 + \sigma_e^2 \\ \sigma_c^2 & \sigma_c^2 \sigma_d^2 & \sigma_c^2 + \sigma_e^2 & \sigma_c^2 + \sigma_e^2 + \sigma_d^2 \end{bmatrix} + \sigma^2 \mathbf{I}_4. \quad (2.5)$$

2.3.2 Estudos Longitudinais Completos ou Balanceados

Para estudos longitudinais completos o gráfico de dispersão múltiplo das observações padronizadas proposto por Dawson et al (1997) expressa o comportamento da matriz de correlações amostrais e é utilizado para diagnosticar tanto a presença de correlação serial quanto o padrão de variação das correlações em cada sub-diagonal. O gráfico das linhas das matrizes de correlações e covariâncias em função da defasagem (“lag”) proposto por Grady e Helms (1995) é utilizado para diagnosticar correlação serial. Esses gráficos também podem ser obtidos para estudos com poucas unidades amostrais com observações incompletas, excluindo-se da construção tais unidades amostrais.

Para ilustrar a utilização desses gráficos, simulamos os dados de 3 estudos longitudinais em que as observações são obtidas ao longo de cinco condições de avaliação denominadas t1, t2, t3, t4 e t5. Em cada estudo as observações apresentam respectivamente correlação constante, serial e comportamento crescente nas sub-diagonais (ante-dependência). A partir dos dados hipotéticos construímos os gráficos de dispersão múltiplos das observações padronizadas apresentados nas Figuras 2.11, 2.12 e 2.13. A correlação sob as condições de avaliação parece ser constante nas Figuras 2.11. Na Figura 2.12 e 2.13, as correlações diminuem ao longo das condições de avaliação (note que a suavização por meio do lowess que significa “locally weighted regression scatter plot smoothing”, fica praticamente horizontal no gráfico da Figura 2.12, mais detalhes sobre lowess podem ser encontrados em Silveira (1992) e em Venables e Ripley (1999)), sugerindo a presença de correlação serial. Na Figura 2.13, podemos observar que as correlações decrescem ao longo das condições de avaliação e aumentam em cada sub-diagonal, sugerindo uma estrutura de ante-dependência para a matriz de covariância. Os gráficos das Figuras 2.14 e 2.15 são gráficos das linhas da matriz de correlações (Grady e Helms (1995)), sugerindo respectivamente ausência de correlação serial e a presença de correlação serial, presença esta detectada pelo decrescimento dos perfis na medida em que a defasagem (“lag”) aumenta.

Figura 2.11: Gráfico de dispersão múltiplo com sugestão de correlação constante com lowess

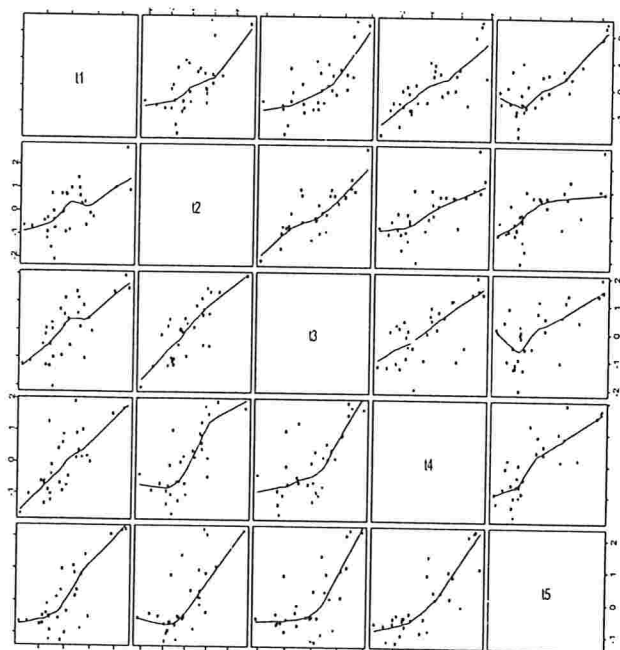


Figura 2.12: Gráfico de dispersão múltiplo com sugestão de correlação serial com lowess

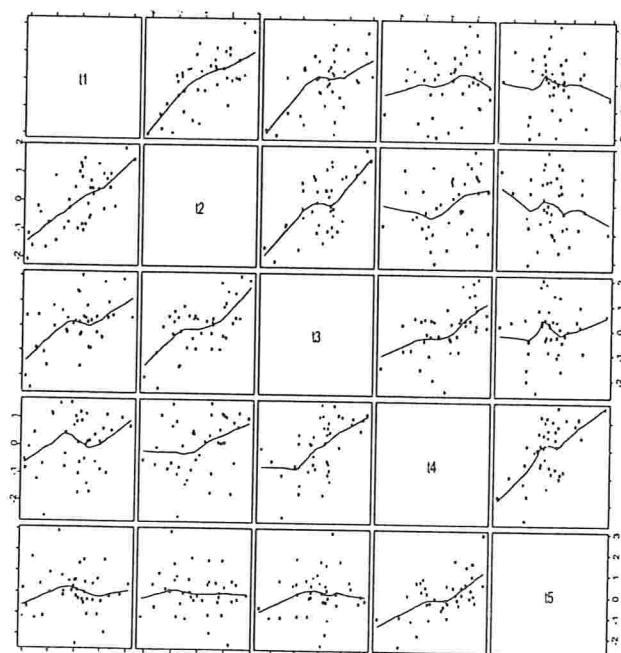


Figura 2.13: Gráfico de dispersão múltiplo com sugestão de ante-dependência com lowess

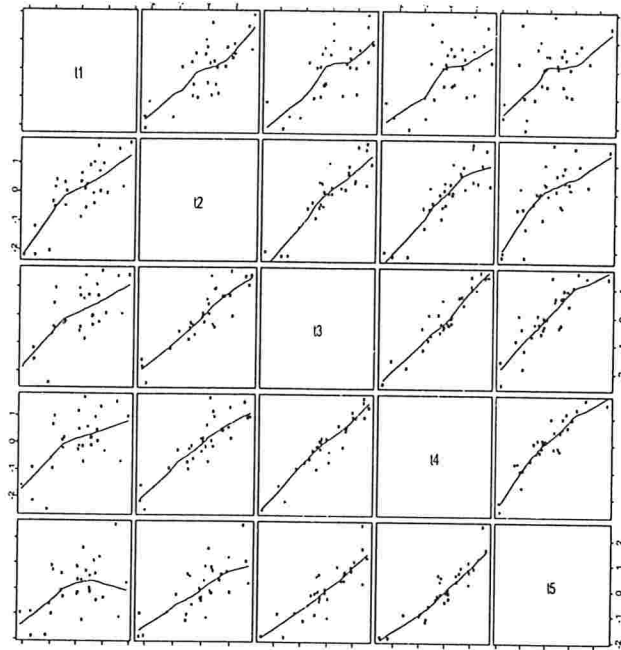


Figura 2.14: Gráfico de perfis para as linhas da matriz de correlações intra-unidades amostrais com sugestão de ausência de correlação serial

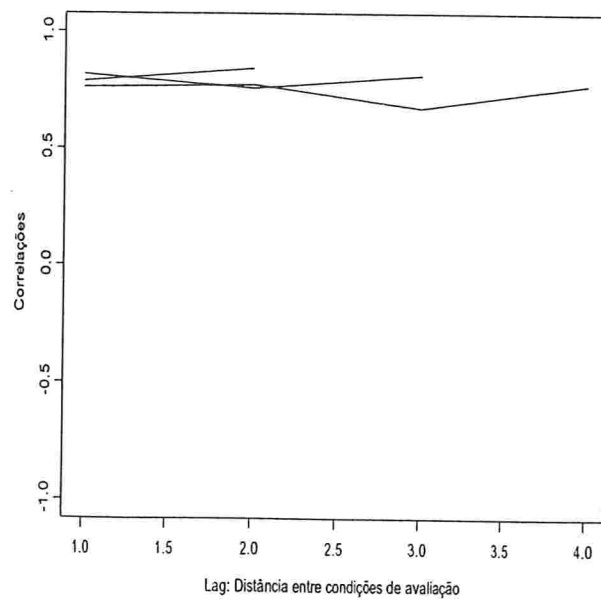
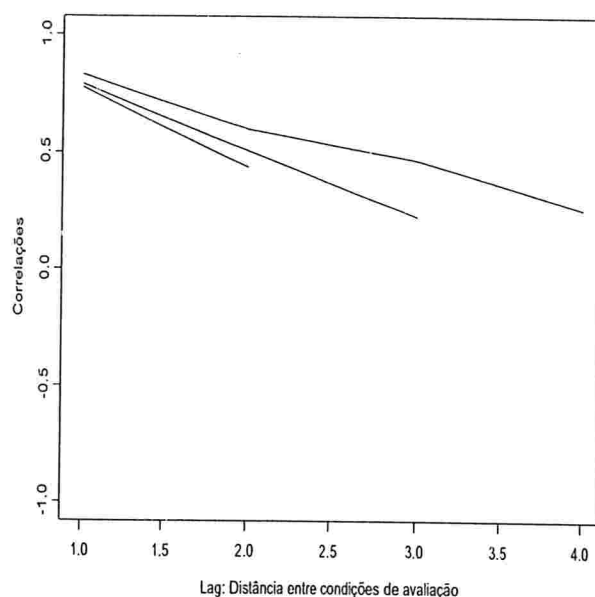


Figura 2.15: Gráfico de perfis para as linhas da matriz de correlações intra-unidades amostrais com sugestão de correlação serial



Para a identificação de possíveis efeitos aleatórios consideraremos a análise das linhas da matriz de covariância amostral proposta por Suyama e Singer (1995). Sob o modelo de independência condicional, a covariância entre observações referentes à mesma unidade amostral é modelada por meio da matriz de covariâncias associada aos efeitos aleatórios, $\mathbf{Z}_i \mathbf{G} \mathbf{Z}_i^t$, enquanto a matriz de covariâncias intra-unidades amostrais, \mathbf{R}_i , contribui apenas para a variância do modelo.

Admitamos que para a identificação dos efeitos aleatórios $\mathbf{Z}_i = \mathbf{Z}$, para $i = 1, \dots, N$. A metodologia para escolha dos efeitos aleatórios consiste em estudar o comportamento dos perfis das linhas da matriz de covariâncias amostral. O objetivo é encontrar a forma funcional do polinômio comum a todas as linhas da matriz de covariância amostral. Para isto podem ser utilizadas técnicas de análise de regressão. Os coeficientes significativos comuns a todas as linhas determinam a forma funcional do polinômio comum a todas as linhas da matriz de covariâncias amostral e, conseqüentemente, determinam os efeitos aleatórios que devem fazer parte do modelo.

Consideremos um exemplo em que as unidades amostrais são observadas em 3 instantes de tempo t_1 , t_2 e t_3 e em que a matriz de covariâncias amostral

$$\mathbf{V}_0 = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{bmatrix}. \quad (2.6)$$

Suponhamos também que o gráfico de perfis das linhas da matriz de covariância amostral sugira que eles possam ser representados por uma reta; nesta situação poderemos incluir como efeitos aleatórios o coeficiente linear e o coeficiente angular. Isto significa que

$$\mathbf{Z} = \begin{bmatrix} 1 & t_1 \\ 1 & t_2 \\ 1 & t_3 \end{bmatrix} \quad (2.7)$$

Apenas para simplificar os cálculos, suponhamos que a matriz de covariância dos efeitos aleatórios de intercepto e inclinação seja $\mathbf{G} = \text{diag}[\sigma_\alpha^2; \sigma_\beta^2]$, implicando

$$\mathbf{ZGZ}^t = \begin{bmatrix} \sigma_\alpha^2 + \sigma_\beta^2 t_1^2 & \sigma_\alpha^2 + \sigma_\beta^2 t_1 t_2 & \sigma_\alpha^2 + \sigma_\beta^2 t_1 t_3 \\ \sigma_\alpha^2 + \sigma_\beta^2 t_1 t_2 & \sigma_\alpha^2 + \sigma_\beta^2 t_2^2 & \sigma_\alpha^2 + \sigma_\beta^2 t_2 t_3 \\ \sigma_\alpha^2 + \sigma_\beta^2 t_1 t_3 & \sigma_\alpha^2 + \sigma_\beta^2 t_2 t_3 & \sigma_\alpha^2 + \sigma_\beta^2 t_3^2 \end{bmatrix}. \quad (2.8)$$

Podemos observar que na matriz (2.8), que os elementos da linha i correspondem aos valores de um polinômio do primeiro grau com coeficientes σ_α^2 e $\sigma_\beta^2 t_i$, calculado nos pontos t_1, t_2 e t_3 . Assim, se denotarmos por \mathbf{z}_i a i -ésima linha da matriz \mathbf{Z} , a contribuição dos efeitos aleatórios para a i -ésima coluna da matriz de dispersão é \mathbf{ZGz}_i . Desta forma, a covariância a_{ij} da matriz de covariâncias amostral (2.6) poderá ser modelada pelo elemento $\sigma_\alpha^2 + \sigma_\beta^2 t_i t_j$ da matriz (2.8), o que implica que cada linha da matriz de covariância amostral poderá ser modelada via um modelo de regressão linear em que as colunas da matriz \mathbf{Z} serão as variáveis explicativas. De um modo geral, a j -ésima coluna (ou linha) da matriz de covariâncias tem a mesma forma funcional que \mathbf{Zb} . Por exemplo, se o coeficiente $\sigma_\beta^2 t_i$ de $\sigma_\alpha^2 + \sigma_\beta^2 t_i t_j$ for significativo, considerando uma análise de regressão usual, o coeficiente angular deve ser incluído como efeito aleatório.

2.3.3 Estudos Longitudinais Incompletos ou Não Balanceados

Para identificar os efeitos aleatórios em estudos longitudinais incompletos ou não balanceados, consideramos uma análise gráfica dos perfis individuais além da técnica sugerida por Rutter e Elashoff (1994) e apresentada em Suyama e Singer (1995). Essa metodologia consiste em ajustar curvas para os perfis individuais e selecionar os efeitos fixos e aleatórios por meio de testes de hipóteses. Esta técnica não requer dados completos, mas as formas funcionais dos modelos individuais devem ser as mesmas; além disso, o número de medidas repetidas por unidade de investigação deve ser maior que o número de parâmetros individuais a serem estimados.

O modelo (1.1), com a suposição de independência condicional pode ser escrito como $y_i = \mathbf{X}_i^* \boldsymbol{\beta}_i + \mathbf{e}_i$, em que \mathbf{X}_i^* é uma matriz composta por colunas das matrizes \mathbf{X}_i e \mathbf{Z}_i construídas da seguinte maneira: colunas comuns em \mathbf{X}_i e \mathbf{Z}_i serão representadas por apenas uma coluna e as colunas diferentes formarão colunas distintas em \mathbf{X}_i^* . O vetor de parâmetros será escrito como $\boldsymbol{\beta}_i = \boldsymbol{\beta}^* + \mathbf{b}_i^*$, em que o símbolo (*) indica que alguns elementos do vetor de parâmetros $\boldsymbol{\beta}^*$ ou do vetor dos termos aleatórios \mathbf{b}_i^* são possivelmente nulos. Por exemplo, consideremos um estudo longitudinal com apenas três condições de avaliação em que o perfil individual possa ser representado por uma reta e tenhamos como efeitos aleatórios, coeficientes de polinômios do segundo grau; as matrizes \mathbf{X}_i e \mathbf{Z}_i e os vetores $\boldsymbol{\beta}$ e \mathbf{b}_i são

$$\mathbf{X}_i = \begin{bmatrix} 1 & t_1 \\ 1 & t_2 \\ 1 & t_3 \end{bmatrix}; \quad \mathbf{Z}_i = \begin{bmatrix} 1 & t_1 & t_1^2 \\ 1 & t_2 & t_2^2 \\ 1 & t_3 & t_3^2 \end{bmatrix}; \quad (2.9)$$

$$\boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}; \quad \mathbf{b}_i = \begin{bmatrix} b_{0i} \\ b_{1i} \\ b_{2i} \end{bmatrix}. \quad (2.10)$$

Com a reparametrização acima, o vetor $\mathbf{b}_i^* = \mathbf{b}_i$, a matriz \mathbf{X}_i^* , $\boldsymbol{\beta}^*$ e o vetor $\boldsymbol{\beta}_i$ são dados

por:

$$\mathbf{X}_i^* = \begin{bmatrix} 1 & t_1 & t_1^2 \\ 1 & t_2 & t_2^2 \\ 1 & t_3 & t_3^2 \end{bmatrix}, \quad \boldsymbol{\beta}^* = \begin{bmatrix} \beta_0 \\ \beta_1 \\ 0 \end{bmatrix}, \quad \boldsymbol{\beta}_i = \begin{bmatrix} \beta_0 + b_{0i} \\ \beta_1 + b_{1i} \\ b_{2i} \end{bmatrix}. \quad (2.11)$$

Expressemos o modelo em dois estgios, com o primeiro estgio especificado por $y_i | \boldsymbol{\beta}_i \sim N(\mathbf{X}_i^* \boldsymbol{\beta}_i; \sigma^2 \mathbf{I}_i)$. Nesse caso, o modelo coincide com os modelos usuais de anlise de regresso. O estimador de mxima verossimilhana de $\boldsymbol{\beta}_i$  dado por $\tilde{\boldsymbol{\beta}}_i = (\mathbf{X}_i^{*t} \mathbf{X}_i^*)^{-1} \mathbf{X}_i^{*t} \mathbf{y}_i$, com distribuio aproximada $\tilde{\boldsymbol{\beta}}_i | \boldsymbol{\beta}_i \sim N(\boldsymbol{\beta}_i; \sigma^2 (\mathbf{X}_i^{*t} \mathbf{X}_i^*)^{-1})$. Admitindo no segundo estgio que $\boldsymbol{\beta}_i \sim N(\boldsymbol{\beta}^*; \sigma^2 \mathbf{G}^*)$, em que a matriz \mathbf{G}^*  a matriz \mathbf{G} aumentada com linhas e colunas iguais a zero, correspondendo a algum elemento nulo do vetor de efeitos aleatrios \mathbf{b}_i , obtemos a distribuio aproximada marginal $\tilde{\boldsymbol{\beta}}_i \sim N(\boldsymbol{\beta}^*; \sigma^2 (\mathbf{X}_i^{*t} \mathbf{X}_i^*)^{-1} + \mathbf{G}^*)$. Assim, podemos considerar um estimador linear no viciado de $\boldsymbol{\beta}^*$ dado por

$$\bar{\boldsymbol{\beta}} = \sum_{i=1}^N \frac{\tilde{\boldsymbol{\beta}}_i}{N} = \sum_{i=1}^N \frac{(\mathbf{X}_i^{*t} \mathbf{X}_i^*)^{-1} \mathbf{X}_i^{*t} \mathbf{y}_i}{N} \quad (2.12)$$

com varincia

$$\text{Var}(\bar{\boldsymbol{\beta}}) = \sum_{i=1}^N \text{Var}\left(\frac{\tilde{\boldsymbol{\beta}}_i}{N}\right) = \sum_{i=1}^N \left[\frac{\sigma^2 (\mathbf{X}_i^{*t} \mathbf{X}_i^*)^{-1} + \mathbf{G}^*}{N^2} \right]. \quad (2.13)$$

Quando $\mathbf{X}_i^* = \mathbf{X}^*$, a varincia de $\bar{\boldsymbol{\beta}}$ ser dada simplesmente por $(\sigma^2 (\mathbf{X}^{*t} \mathbf{X}^*)^{-1} + \mathbf{G}^*)/N$.

O passo seguinte  a identificao de quais elementos do vetor $\boldsymbol{\beta}^*$ podem ser considerados nulos, ou seja, quais efeitos fixos fazem parte da forma funcional que representa o perfil mdio. Para isso, propomos testar hipteses do tipo $H_0 : \beta_j^* = 0$ para $j = 1, 2, \dots, \text{posto}(\boldsymbol{\beta}^*)$, utilizando a estatstica $t = \bar{\beta}_j / \sqrt{c_{jj}}$, com $\bar{\beta}_j$ representando o j -simo elemento do vetor $\bar{\boldsymbol{\beta}}$ e c_{jj} correspondente ao j -simo elemento da diagonal da matriz $(\sigma^2 (\sum_i \mathbf{X}_i^{*t} \mathbf{X}_i^*)^{-1})/N^2$, que ser menor do que aquele encontrado por meio da expresso (2.13). Desta forma, teremos um teste conservador no sentido de identificarmos mais candidatos a efeitos fixos do que eventualmente ter o modelo final, uma caracterstica importante em um estgio de identificao.

Para a identificao dos efeitos aleatrios, devemos observar primeiramente que a disperso esperada para o estimador do j -simo elemento do vetor $\boldsymbol{\beta}_i$, $\tilde{\beta}_{ij}$,  dada pelo

j-ésimo elemento da diagonal principal da matriz de dispersão da análise de regressão usual, isto é, $\sigma^2(\mathbf{X}_i^{*t}\mathbf{X}_i^*)^{-1}$ na ausência de efeito aleatório. Na presença de efeito aleatório, esse termo exibirá uma variabilidade maior, que passaremos a chamar de variabilidade extra. Quando $\beta_j^* = 0$, a variabilidade extra deverá ser capturada pelos parâmetros β_{ij} , $i = 1, \dots, N$, significativamente diferentes de zero. Como consequência, mais que 5% das estimativas estarão fora dos limites de dois erros-padrão em torno de zero. O erro-padrão é dado pela raiz quadrada do termo apropriado da diagonal principal da matriz de dispersão $\sigma^2(\mathbf{X}_i^{*t}\mathbf{X}_i^*)^{-1}$. Assim, se calcularmos o desvio padrão amostral das estimativas $\tilde{\beta}_{ij}$ e encontrarmos mais de 5% delas fora do intervalo de mais ou menos dois desvios padrão, concluímos que há indícios de efeito aleatório para o coeficiente β_{ij} .

Por outro lado, quando há evidência de efeitos fixos, o limite de dois erros-padrão deverá ser dado em torno da estimativa de β_j , que é $\bar{\beta}_j$, com erro padrão dado pela raiz quadrada do j-ésimo elemento da diagonal da matriz de dispersão

$$Var(\tilde{\beta}_i - \bar{\beta}) = \sigma^2[(\mathbf{X}_i^{*t}\mathbf{X}_i^*)^{-1} - \sum_{i=1}^N \frac{(\mathbf{X}_i^{*t}\mathbf{X}_i^*)^{-1}}{N^2}]. \quad (2.14)$$

Quando $\mathbf{X}_i^* = \mathbf{X}^*$, a expressão simplifica-se para

$$Var(\tilde{\beta}_i - \bar{\beta}) = \sigma^2\left[\frac{(\mathbf{X}_i^{*t}\mathbf{X}_i^*)^{-1}(N-1)}{N}\right]. \quad (2.15)$$

Portanto, para situações em que há efeitos fixos devemos calcular os desvios $\tilde{\beta}_{ij} - \bar{\beta}_j$ e repetir o mesmo procedimento anterior, ou seja, observar se mais de 5% destes desvios não pertencem ao intervalo formado por mais ou menos dois desvios padrão, sendo este desvio obtido pela expressão (2.14). Podemos ainda ser mais conservadores no sentido de identificarmos mais candidatos a efeitos aleatórios considerando, por exemplo, o intervalo compreendido por 1.64 desvios padrão, acima ou abaixo da média.

Após a identificação dos possíveis efeitos aleatórios, pode-se ajustar modelos mistos de independência condicional e modelos que possam incorporar correlação serial na estrutura de covariâncias, como o modelo de efeitos aleatórios com estrutura auto-regressiva para as medidas realizadas na mesma unidade amostral. A seleção da estrutura mais adequada dentre estas pode ser baseada nos critérios de informação AIC_R e BIC_R .

2.4 Estratégias de Análise

Em resumo, para estudos não longitudinais, consideraremos a seguinte estratégia para a seleção da estrutura de covariância intra-unidades amostrais

1. Análise das matrizes de covariâncias e correlações amostrais segundo as diretrizes apresentadas na Seção 2.2;
2. Análise do planejamento do estudo para propor um modelo para os efeitos fixos e identificação de possíveis efeitos aleatórios conforme indicado na Seção 2.3.1;
3. Ajustes do modelo saturado para os parâmetros de regressão identificados no item 2, com as matrizes de covariâncias que possam incorporar os padrões identificados nos itens 1 e 2;
4. Utilização dos critérios de informação de AIC_R e BIC_R , como ferramenta auxiliar na escolha da estrutura de covariâncias identificada em 1 e 2;

Para estudos longitudinais, uma estratégia para a identificação da estrutura de covariâncias intra-unidades amostrais é a seguinte:

1. Análise do gráfico de perfis médios para propor um modelo para os efeitos fixos segundo as diretrizes apresentadas na Seção 2.2;
2. Análise dos gráficos de perfis individuais e de perfis individuais centralizados consoante as indicações da Seção 2.2;
3. Análise das matrizes de covariâncias e correlações amostrais, quando for possível calculá-las conforme indicado na Seção 2.2;
4. Análise dos gráficos de dispersão múltiplos das observações padronizadas quando for possível construí-los conforme indicado na Seção 2.3.2;
5. Análise dos gráfico das linhas das matrizes de covariâncias e correlações amostrais versus defasagens da Seção 2.3.2;

6. Análise do variograma amostral apresentado na Seção 2.2;
7. No caso de estudos longitudinais completos, a análise do gráfico das linhas das matrizes de covariâncias para a identificação de efeitos aleatórios, apresentada na Seção 2.3.2;
8. No caso de estudos longitudinais incompletos, a análise do gráfico das linhas dos perfis individuais para a identificação de efeitos aleatórios, apresentada na Seção 2.3.3;
9. Utilização dos critérios de informação para selecionar das estruturas de covariâncias intra-unidades amostrais identificadas a que melhor representará a variabilidade dos dados.

No caso de estudos longitudinais com covariável dependente do tempo, sugerimos considerar a estratégia acima do item 2 ao 9 para os resíduos de mínimos quadrados ordinários do modelo saturado.

Capítulo 3

Exemplos de Modelagem da Estrutura de Covariância

3.1 Introdução

Neste capítulo, aplicaremos as estratégias para seleção de estruturas de covariância apresentadas no capítulo anterior aos exemplos descritos no Capítulo 1. Iniciaremos com um estudo com medidas repetidas não longitudinais, considerando o Exemplo 1.3 apenas com os dados correspondentes à profundidade de $10\mu m$ e o Exemplo 1.4, que, apesar de ter como objetivo o ajuste de uma curva de crescimento, não é um estudo longitudinal dada a aleatorização da ordem de observação das respostas sob as diferentes salinidades testes (Cúri, (1999)). Em seguida, aplicaremos a estratégia apresentada na Seção 2.4 para a identificação de estruturas de covariância em estudos longitudinais, como é o caso dos Exemplos 1.1, 1.2 e 1.3.

3.2 Modelagem da Estrutura de Covariância para Estudos Não Longitudinais Completos

Para os dados correspondentes à profundidade de $10\mu m$ no Exemplo 1.3, é razoável considerar que medidas realizadas em dentes diferentes mas observados em aparelho intrabucal do mesmo voluntário sejam correlacionadas, o que o torna o estudo equivalente a um estudo com planejamento em blocos aleatórios.

Como podemos observar na matriz de covariâncias e correlações amostrais (3.1), as variâncias amostrais correspondentes aos três tratamentos são diferentes e a correlação amostral entre os tratamentos laser e controle extrabucal é consideravelmente maior que as demais.

	controle <i>in situ</i>	controle extrabucal	laser	
controle <i>in situ</i>	326.34	10.11	67.88	(3.1)
controle extrabucal	0.05	117.51	105.66	
laser	0.18	0.48	420.83	

Como modelo inicial, consideramos

$$y_{ijk} = \mu + \beta_j + \pi_i + \varepsilon_{ijk}, \quad (3.2)$$

em que $i = 1, 2, 3, 4, 6, 7, 8$ indexa voluntário, $j = 1, 2, 3$ indexa tratamento, com níveis controle *in situ*, controle extrabucal e laser respectivamente, $k = 1, \dots, 32$ representa o índice de dente, $\sum_{j=1}^3 \beta_j = 0$ e $\pi_i \sim NI(0; \sigma_v^2)$, $\varepsilon_{ijk} \sim N(0; \sigma_j^2)$ com $cov(\varepsilon_{ijk}, \varepsilon_{ij'k}) = \sigma_{jj'}$, com π_i e ε_{ijk} independentes. Aqui, y_{ijk} denota a microdureza obtida no k -ésimo dente submetido ao i -ésimo voluntário e ao j -ésimo tratamento, μ é a média geral, μ_j é a microdureza média sob o j -ésimo tratamento, $\beta_j = \mu_j - \mu$ representa o efeito associado ao j -ésimo tratamento, π_i é o efeito aleatório para o i -ésimo voluntário e ε_{ijk} representa o erro de medida associado a cada observação y_{ijk} . Completamos a especificação do modelo

inicial com a seguinte estrutura de covariância intra-unidades amostrais

$$V = J_{12}\sigma_v^2 + I_4 \otimes \begin{bmatrix} \sigma_{11}^2 & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_{22}^2 & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & \sigma_{33}^2 \end{bmatrix}. \quad (3.3)$$

O ajuste desse modelo fornece as estimativas apresentadas na da Tabela 3.1. A coluna “Estatística de teste” contém valores da estatística de Wald para testes de covariâncias nulas.

Tabela 3.1: *Estimativas dos parâmetros do modelo (3.2) com matriz de covariâncias intra-unidades sob o modelo (3.3).*

Parâmetro	Estimativa	Erro Padrão	Estatística de teste	Nível Descritivo
σ_v^2	39.15	29.81	-	-
σ_{11}^2	270.37	79.07	-	-
σ_{12}	-33.69	33.60	1.00	0.32
σ_{22}^2	99.55	29.45	-	-
σ_{13}	16.86	61.38	0.27	0.78
σ_{23}	78.24	44.32	1.77	0.08
σ_{33}^2	394.49	105.25	-	-

Podemos observar que as covariâncias não são significativamente diferentes de zero, sugerindo que a matriz de covariâncias do modelo inicial pode ser reduzida para

$$V = J_{12}\sigma_v^2 + I_4 \otimes \begin{bmatrix} \sigma_{11}^2 & 0 & 0 \\ 0 & \sigma_{22}^2 & 0 \\ 0 & 0 & \sigma_{33}^2 \end{bmatrix} \quad (3.4)$$

As estimativas dos parâmetros da matriz de covariâncias correspondentes estão dispostos na Tabela 3.2.

Tabela 3.2: Estimativas dos parâmetros do modelo (3.2) com matriz de covariância intra-unidades amostrais sob o modelo (3.4).

Parâmetro	Estimativa	Erro Padrão
σ_v^2	42.84	30.77
σ_{11}^2	306.54	85.28
σ_{22}^2	88.80	24.76
σ_{33}^2	360.62	95.06

Sob a estrutura (3.4) realizamos testes da razão de verossimilhanças para verificar se as medidas realizadas em dentes diferentes mas submetidos ao mesmo voluntário são correlacionadas ($H_0 : \sigma_v^2 = 0$), ou seja, para avaliar a necessidade de inclusão do efeito aleatório de voluntário. A estatística de teste correspondente tem como distribuição assintótica sob H_0 uma mistura 50:50 de χ_0^2 e χ_1^2 . O valor observado da estatística foi 5.1, ($p = 0.0119$), sugerindo a manutenção do efeito aleatório de voluntário no modelo.

Os valores dos critérios de informação AIC_R e BIC_R , apresentados na Tabela 3.3 abaixo, sugerem que o modelo (3.4) é mais adequado que o modelo (3.3) para representar a estrutura de covariância dos dados.

Tabela 3.3: Estatísticas de ajuste dos modelos (3.3) e (3.4).

Estruturas	Número de parâmetros	$-2l_R$	AIC_R	BIC_R
Modelo (3.3)	7	777.6	791.6	792.2
Modelo (3.4)	4	783.1	791.1	791.4

A Tabela 3.4 apresenta as estimativas das médias dos tratamentos com seus respectivos erros padrões obtidos sob as estruturas de covariâncias dos modelos (3.3) e (3.4). As estimativas das médias são iguais nos dois modelos e seus erros padrões são similares.

Tabela 3.4: Estimativas das médias dos tratamentos com seus respectivos erros padrões.

Médias dos tratamentos	Estruturas de covariâncias			
	Modelo (3.3)		Modelo (3.4)	
	Estimativa	Erro padrão	Estimativas	Erro padrão
μ_1	163.94	3.65	163.94	3.86
μ_2	283.45	2.83	283.45	2.85
μ_3	239.73	4.15	239.73	4.08

3.3 Estudos Não Longitudinais Incompletos

No Exemplo 1.4, como observou Cúri (1999), o gráfico de perfis médios (Figura 3.1) mostra que a amplitude média do PAC parece crescer para salinidades testes menores que a salinidade de aclimação e decrescer para salinidades testes maiores que a salinidade de aclimação em cada um dos grupos. Isto sugere o estudo da amplitude média do PAC em função das diferenças de salinidades testes e de aclimação; a representação gráfica correspondente é apresentada na Figura 3.2.

Para os efeitos fixos, o gráfico da Figura 3.2 sugere um modelo de regressão segmentada em que a mudança de coeficiente angular ocorre no ponto em que a diferença de salinidade é nula.

Figura 3.1: Perfis médios da amplitude do PAC (mV) em função das salinidades teste

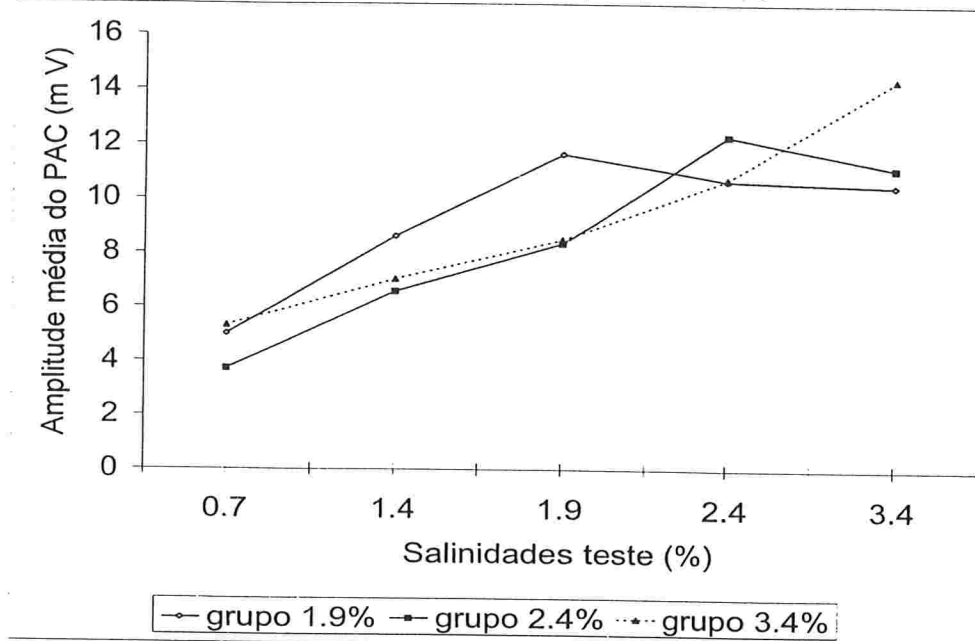
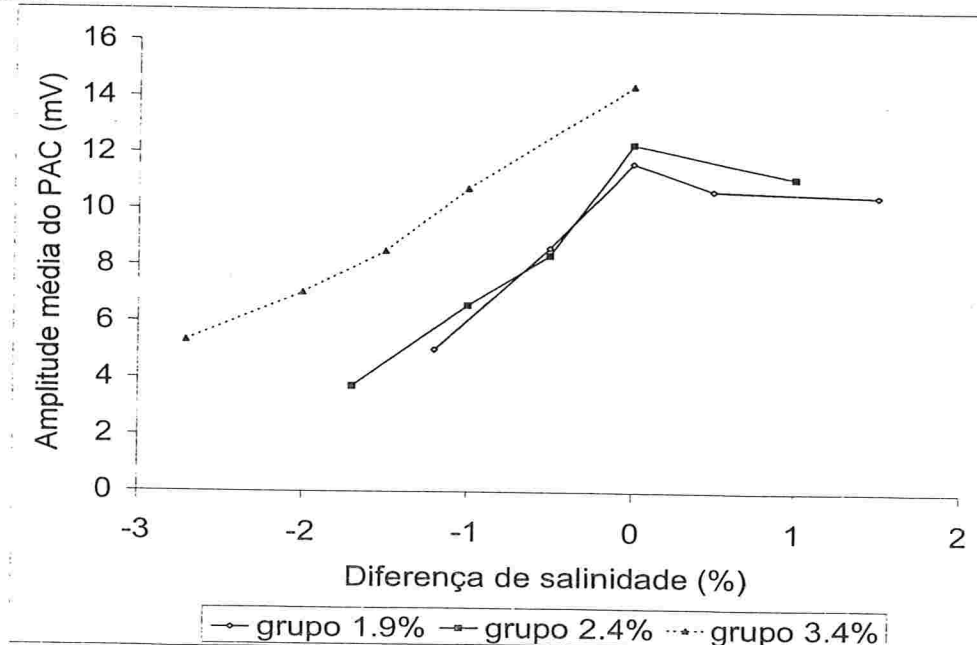


Figura 3.2: Perfis médios da amplitude do PAC (mV) em função das diferenças de salinidades teste



Para modelar a estrutura de covariâncias, começamos analisando os gráficos de perfis individuais da amplitude do PAC em função das diferenças de salinidade (Figuras

3.3, 3.4 e 3.5). Pode-se notar uma variabilidade do intercepto (para diferenças de salinidade teste e de aclimação iguais a zero) e das inclinações dos segmentos de retas, sugerindo possíveis efeitos aleatórios para os coeficientes linear e angular.

Figura 3.3: Perfis individuais da amplitude do PAC em função das diferenças de salinidades para o grupo com salinidade de aclimação 1.9%

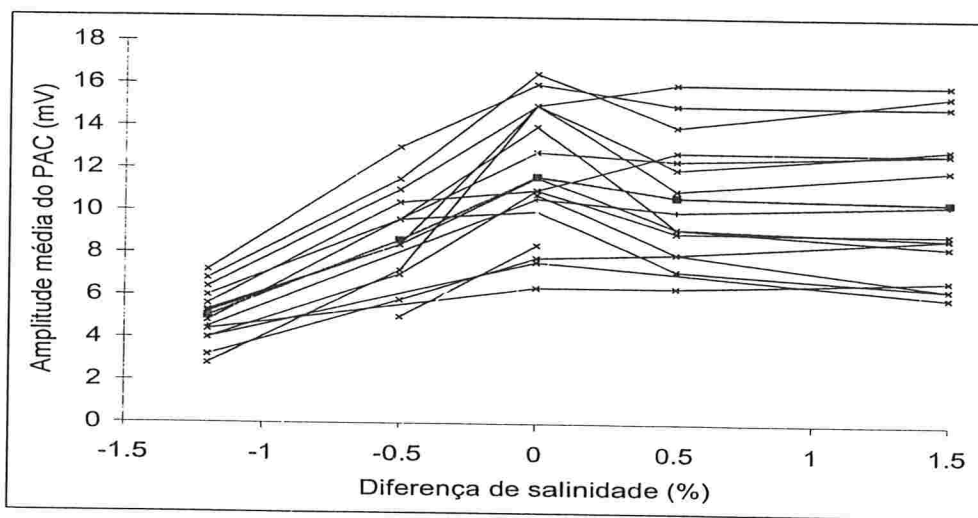


Figura 3.4: Perfis individuais da amplitude do PAC em função das diferenças de salinidades para o grupo com salinidade de aclimação 2.4%

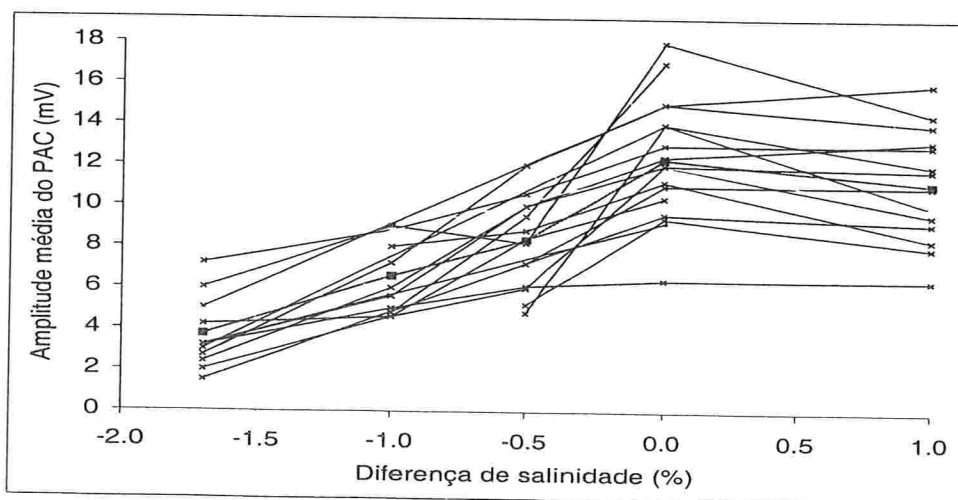
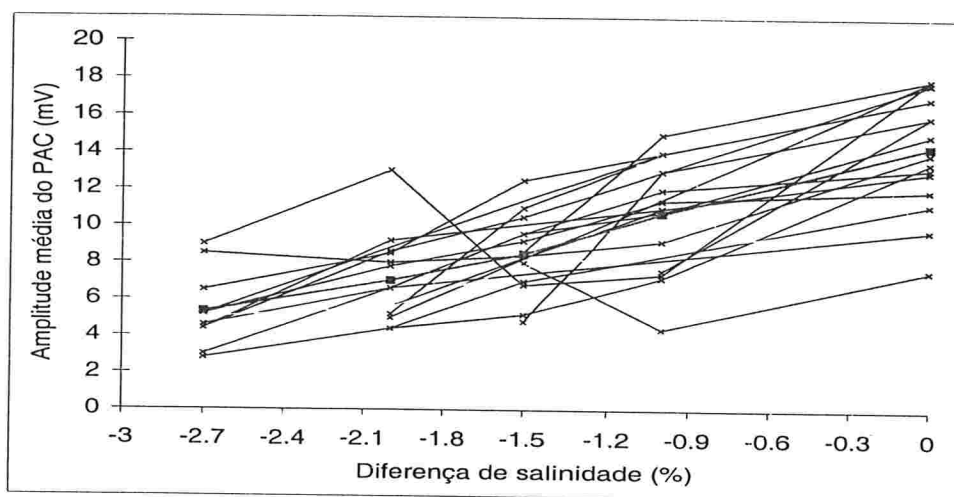


Figura 3.5: Perfis individuais da amplitude do PAC em função das diferenças de salinidades para o grupo com salinidade de aclimação 3.4%



Cúri (1999) também mostrou que as medidas controles podem ser consideradas como covariáveis e que a relação entre as medidas controles e as salinidades testes é linear. Como não se observaram as medidas controles da amplitude do PAC correspondentes às medidas sob as salinidades de aclimação, supôs-se que os valores controles são iguais aos valores obtidos na salinidade de aclimação. Isto sugere uma análise dos gráficos de perfis individuais sem a influência das salinidades controle. Para isso, consideraremos inicialmente o ajuste do modelo

$$y_{ijk} = \mu_{jk} + \beta_{jk}(x_{ijk} - \bar{x}_j) + \varepsilon_{ijk}, \quad (3.5)$$

em que $i = 1, \dots, n_j$ indexa o molusco e n_j representa o número de moluscos no grupo com salinidade de aclimação j , $j = 1.9\%, 2.4\%$ e 3.4% , $k = 0.7\%, 1.4\%, 1.9\%, 2.4\%, 3.4\%$ indexa as salinidades testes e controles e $\varepsilon_{ijk} \sim NID(0; \sigma^2)$. Além disso,

y_{ijk} denota a amplitude do PAC (mV) obtida no i -ésimo molusco submetido à j -ésima salinidade de aclimação na k -ésima salinidade teste;

x_{ijk} denota a amplitude do PAC (mV) obtida no i -ésimo molusco submetido à j -ésima salinidade de aclimação na k -ésima salinidade controle;

μ_{jk} é a média da amplitude do PAC no j -ésimo grupo na k -ésima salinidade teste ajustada pela medida controle, isto é para $x_{ijk} = \bar{x}_j$;

β_{jk} é o coeficiente de ajuste da amplitude média do PAC pela medida controle do j -ésimo grupo na k -ésima salinidade controle;

ε_{ijk} representa o erro de medida associado à y_{ijk} .

Os gráficos das Figuras 3.6, 3.7 e 3.8 apresentam os perfis individuais ajustados pelo efeito da covariável, ou seja, para a variável $y_{ijk}^* = y_{ijk} - \hat{\beta}_{jk}(x_{ijk} - \bar{x}_j)$. Podemos observar que todos os segmentos de reta têm o mesmo intercepto, o que ocorre devido à construção da covariável. Assim, para salinidades controles iguais às salinidades de aclimação, a estimativa de β_{jk} será igual a um, fazendo com que todas as unidades amostrais tenham o mesmo valor, $y_{ijj}^* = \bar{x}_j$. Além disso, como podemos observar nesses gráficos, os coeficientes angulares dos segmentos de reta apresentam uma certa variabilidade, indicando que poderiam ser incluídos como efeitos aleatórios.

Figura 3.6: Perfis individuais da amplitude do PAC em função das diferenças de salinidades ajustadas pelas medidas controles para o grupo com salinidade de aclimação 1.9%

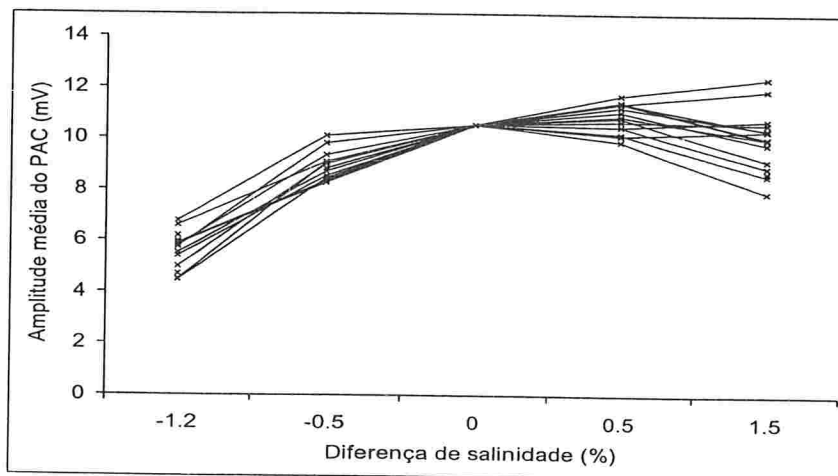


Figura 3.7: Perfis individuais da amplitude do PAC em função das diferenças de salinidades ajustadas pelas medidas controles para o grupo com salinidade de aclimação 2.4%

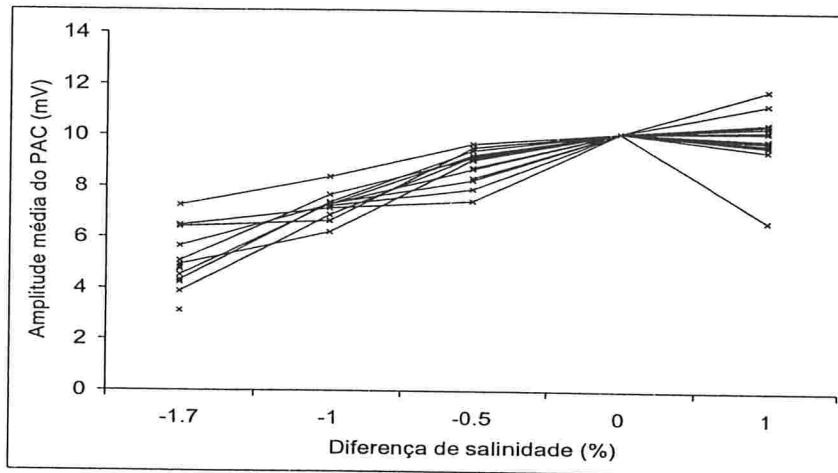
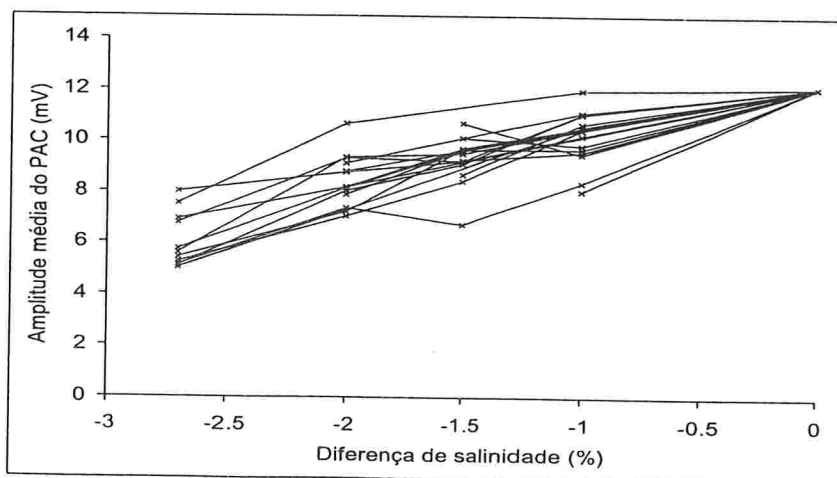


Figura 3.8: Perfis individuais da amplitude do PAC em função das diferenças de salinidades ajustadas pelas medidas controles para o grupo com salinidade de aclimação 3.4%



Para investigar o comportamento das variâncias ao longo das diferenças de salinidade, construímos os gráficos de perfis individuais dos resíduos do modelo (3.5). Esses gráficos encontram-se nas Figuras 3.9, 3.10 e 3.11 e mostram que, para cada de salinidade de aclimação, a variância parece mudar tanto para salinidades testes menores que a salinidade de aclimação quanto para salinidades testes maiores que a salinidade de aclimação.

Figura 3.9: Perfis individuais dos resíduos em função das diferenças de salinidades das medidas controles para o grupo com salinidade de aclimação 1.9%

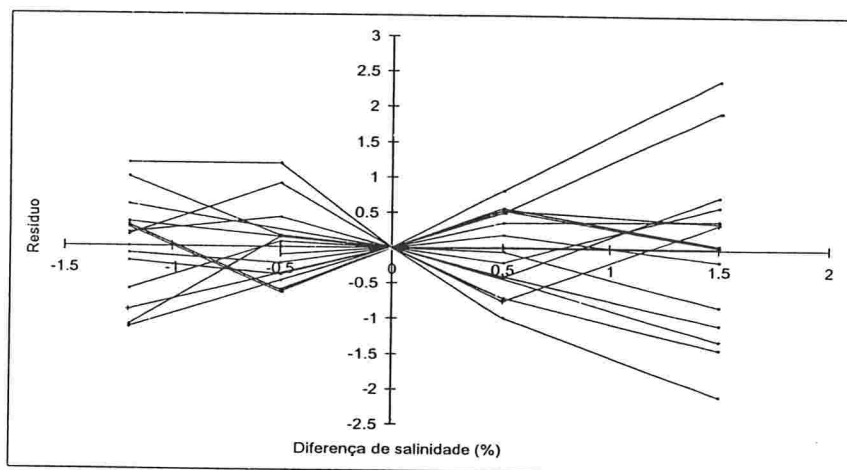


Figura 3.10: Perfis individuais dos resíduos em função das diferenças de salinidades das medidas controles para o grupo com salinidade de aclimação 2.4%

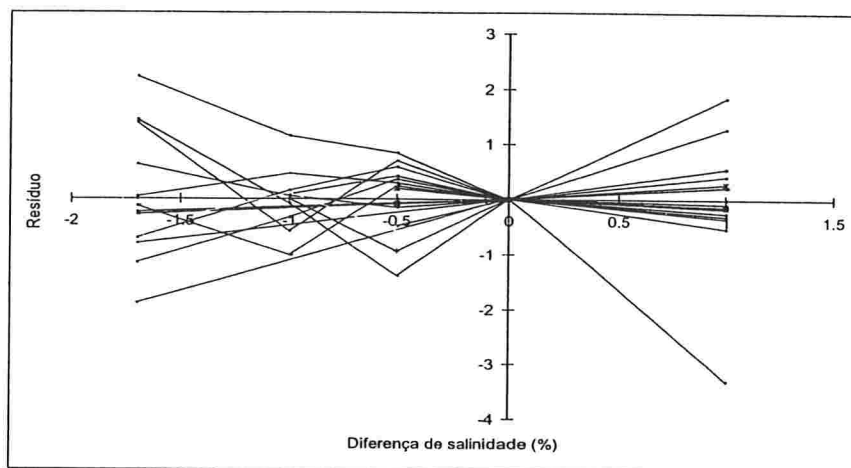
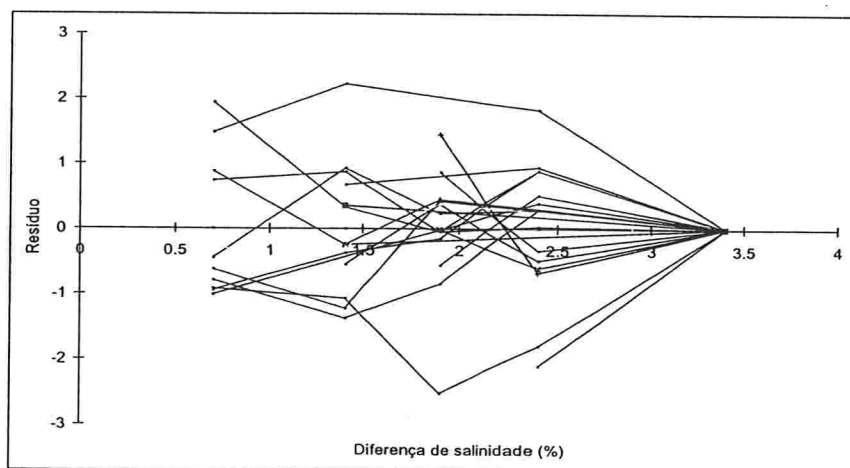


Figura 3.11: Perfis individuais dos resíduos em função das diferenças de salinidades das medidas controles para o grupo com salinidade de aclimação 3.4%



Para avaliarmos a existência de correlação serial, construímos variogramas dos resíduos padronizados do modelo (3.5) para os três grupos de salinidades de aclimação

Figuras 3.12, 3.13 e 3.14. Eles indicam que as correlações das medidas realizadas no mesmo molusco são baixas e constantes. O variograma estimado pelo ajuste de uma curva não paramétrica intercepta a linha da variância somente nas defasagens em que a variabilidade é mais alta que as demais (Diggle et al. (2002)).

Figura 3.12: Variograma dos resíduos padronizados para o grupo com salinidade de aclimação 1.9%

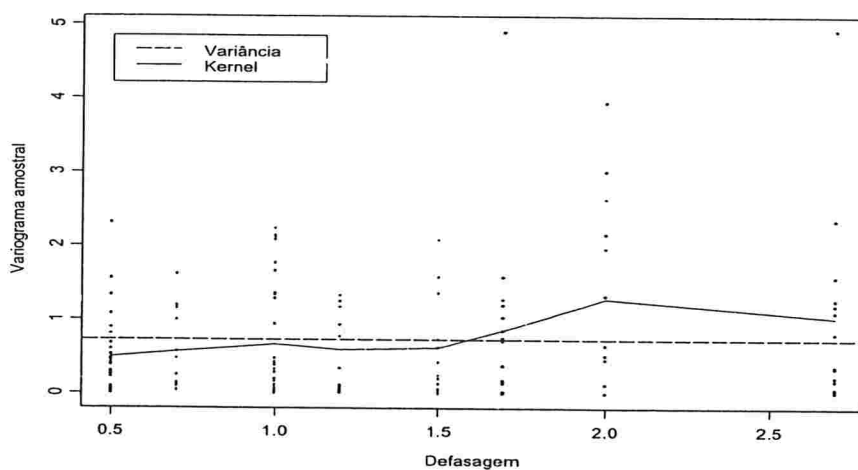


Figura 3.13: Variograma dos resíduos padronizados para o grupo com salinidade de aclimação 2.4%

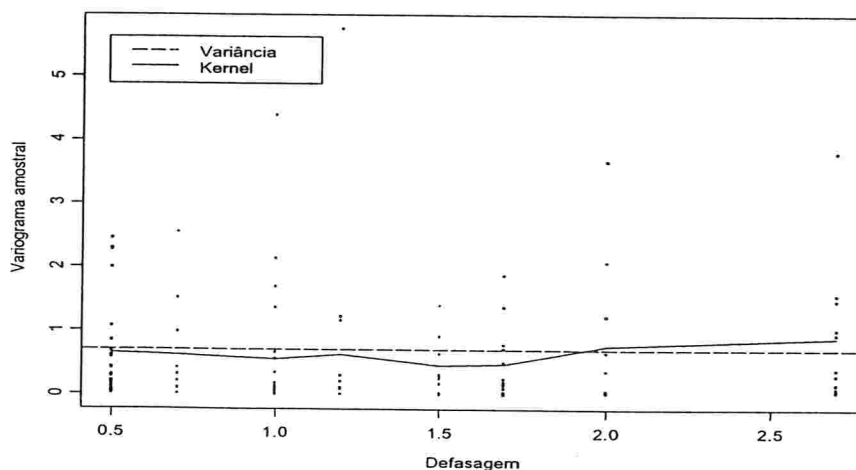
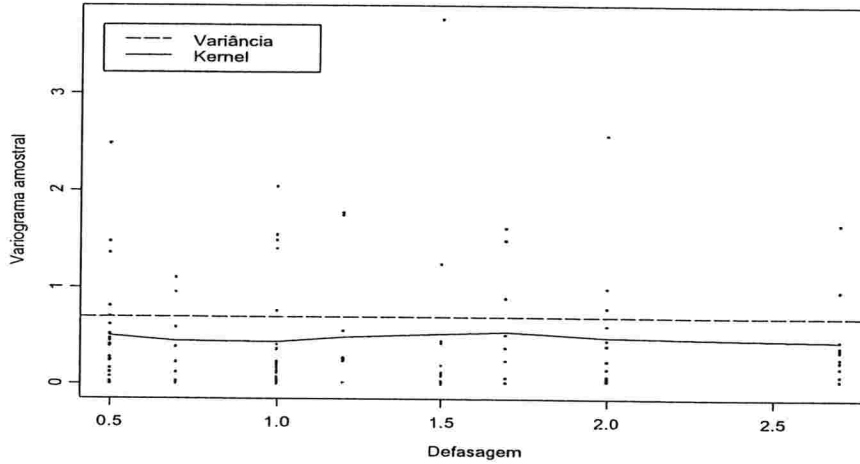


Figura 3.14: Variograma dos resíduos padronizados para o grupo com salinidade de aclimação 3.4%



Com base nos resultados acima, consideramos o modelo de efeitos aleatórios

$$y_{ijk} = \mu_{jk} + \beta_{jk}(x_{ijk} - \bar{x}_j) + \gamma_{ijs}(k - j) + \varepsilon_{ijk}, \quad (3.6)$$

em que $j = 1.9\%, 2.4\%$ e 3.4% indexa a salinidade de aclimação, $i = 1, \dots, n_j$ indexa o molusco e n_j é o número de moluscos no grupo com salinidade de aclimação j , $k = 0.7\%, 1.4\%, 1.9\%, 2.4\%, 3.4\%$ indexa as salinidades testes, s vale 1 se $k - j > 0$ e 0 em caso contrário, $\varepsilon_{ijk} \sim NI(0; \sigma^2)$, $\gamma_{ijs} \sim NI(0; \sigma_{js}^2)$ e denota o efeito aleatório do coeficiente angular para o s -ésimo segmento de reta para o i -ésimo molusco submetido à j -ésima salinidade de aclimação, com ε_{ijk} e γ_{ijs} independentes. O modelo acima fornece, para cada grupo de salinidade de aclimação, as seguintes estruturas de covariâncias intra-unidades amostrais com a seguinte especificação,

$$V_{i1.9} = \begin{bmatrix} -1.2 & 0 \\ -0.5 & 0 \\ 0 & 0 \\ 0 & 0.5 \\ 0 & 1.5 \end{bmatrix} \begin{bmatrix} \sigma_{1.9;0}^2 & 0 \\ 0 & \sigma_{1.9;1}^2 \end{bmatrix} \begin{bmatrix} -1.2 & -0.5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.5 & 1.5 \end{bmatrix} + \sigma^2 I_5$$

$$V_{i2.4} = \begin{bmatrix} -1.7 & 0 \\ -1 & 0 \\ -0.5 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \sigma_{2.4;0}^2 & 0 \\ 0 & \sigma_{2.4;1}^2 \end{bmatrix} \begin{bmatrix} -1.7 & -1 & -0.5 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} + \sigma^2 \mathbf{I}_5$$

$$V_{i3.4} = \begin{bmatrix} -2.7 \\ -2 \\ -1.5 \\ -1 \\ 0 \end{bmatrix} \begin{bmatrix} \sigma_{3.4}^2 \end{bmatrix} \begin{bmatrix} -1.7 & -2 & -1.5 & -1 & 0 \end{bmatrix} + \sigma^2 \mathbf{I}_5$$

Como estamos sob um modelo de independência condicional para verificar a suposição de homocedasticidade de ε_{ijk} , podemos considerar os resíduos condicionais do modelo (3.6) para cada grupo de salinidade de aclimação como sendo provenientes de amostras independentes, o que permite o emprego do teste de Levene para igualdade de variâncias. Neste caso, $p=0.002$, sugerindo uma diferença entre as variâncias dos três grupos de salinidade. Assim, basta considerar $\varepsilon_{ijk} \sim NI(0; \sigma_j^2)$ no modelo (3.6); o modelo resultante é rotulado por (3.7).

Para confirmar a adequabilidade da estrutura de covariâncias intra-unidades amostrais obtida sob o modelo (3.7), comparamos os correspondentes valores dos critérios de informação com os do modelo (3.6), e com outros correspondentes a modelos da forma (3.5) que consideram a modelagem da matriz de covariâncias intra-unidades amostrais exclusivamente em termos de \mathbf{R}_i . Dos modelos apresentados na Seção 1.2 incluímos o modelo uniforme (1.5), o modelo uniforme heterogêneo (1.13) e o modelo não estruturado (1.17). Os critérios AIC_R e BIC_R confirmam o ajuste da matriz de covariâncias intra-unidades amostrais do modelo (3.7) como sendo a que melhor representa a matriz de covariâncias intra-unidades amostrais dos dados (ver Tabela 3.5).

Tabela 3.5: *Estatísticas de ajuste de modelos para a estrutura de covariâncias intra-unidades amostrais para o Exemplo 1.4.*

Estruturas	Número de parâmetros	$-2l_R$	AIC_R	BIC_R
Modelo 3.6	6	509.94	521.94	534.09
Modelo 3.7	8	493.70	509.65	525.85
Uniforme	4	566.39	574.39	582.50
Uniforme heterogênea	6	551.27	563.27	575.42
Não estruturada	15	540.23	570.23	600.61

Além disso, como pode ser observado na Tabela 3.6, o modelo 3.7 apresentou os menores erros padrões para as estimativas dos efeitos fixos correspondentes às salinidades de aclimação 1.9% e 2.4% dentre as estruturas de covariâncias ajustadas. Isso não ocorre para a salinidade de aclimação 3.4% possivelmente devido ao elevado número de dados omissos ("missing data"), em torno de 26% das observações, enquanto os grupos com salinidades de aclimação 1.9% e 2.4% apresentaram respectivamente 14% e 21% de dados omissos.

Tabela 3.6: Estimativas dos parâmetros do Exemplo 1.4 com seus respectivos erros padrões sob modelos com diferentes estruturas de covariâncias intra-unidades amostrais.

Parâmetros	Estruturas de covariâncias									
	Modelo 3.6		Modelo 3.7		U		UH		NE	
	Est	EP	Est	EP	Est	EP	Est	EP	Est	EP
$\mu_{1.9,0.7}$	5.58	0.23	5.55	0.23	5.59	0.29	5.59	0.29	5.64	0.28
$\mu_{1.9,1.4}$	8.90	0.15	8.90	0.12	8.93	0.19	8.93	0.19	8.91	0.18
$\mu_{1.9,1.9}$	10.53	0.13	10.53	0.09	10.53	0.18	10.53	0.18	10.59	0.18
$\mu_{1.9,2.4}$	10.85	0.16	10.84	0.13	10.92	0.15	10.92	0.15	10.95	0.15
$\mu_{1.9,3.4}$	9.99	0.31	10.00	0.31	10.13	0.21	10.13	0.21	10.18	0.21
$\mu_{2.4,0.7}$	4.90	0.39	4.89	0.38	5.03	0.41	5.03	0.41	4.97	0.38
$\mu_{2.4,1.4}$	7.00	0.22	6.99	0.21	7.21	0.23	7.21	0.23	7.22	0.21
$\mu_{2.4,1.9}$	8.78	0.16	8.78	0.15	8.83	0.21	8.83	0.21	8.77	0.21
$\mu_{2.4,2.4}$	10.13	0.15	10.13	0.14	10.13	0.18	10.13	0.18	10.08	0.17
$\mu_{2.4,3.4}$	9.97	0.33	9.97	0.33	9.88	0.24	9.88	0.24	9.91	0.23
$\mu_{3.4,0.7}$	6.11	0.35	6.10	0.36	6.04	0.32	6.04	0.32	6.11	0.31
$\mu_{3.4,1.4}$	8.32	0.27	8.34	0.29	8.42	0.22	8.42	0.22	8.40	0.21
$\mu_{3.4,1.9}$	9.34	0.21	9.34	0.23	9.22	0.23	9.22	0.23	9.33	0.22
$\mu_{3.4,2.4}$	10.12	0.17	10.12	0.19	10.19	0.17	10.19	0.17	10.22	0.16
$\mu_{3.4,3.4}$	12.09	0.14	12.09	0.18	12.09	0.21	12.09	0.21	12.02	0.21
$\beta_{1.9,0.7}$	0.37	0.07	0.37	0.07	0.36	0.09	0.36	0.09	0.34	0.08
$\beta_{1.9,1.4}$	0.90	0.06	0.89	0.04	0.90	0.08	0.90	0.08	0.93	0.07
$\beta_{1.9,1.9}$	1.00	0.04	1.00	0.03	1.00	0.06	1.00	0.06	1.00	0.06
$\beta_{1.9,2.4}$	1.01	0.06	1.00	0.05	1.04	0.06	1.04	0.06	1.02	0.05
$\beta_{1.9,3.4}$	0.91	0.09	0.89	0.08	0.92	0.07	0.92	0.07	0.90	0.06
$\beta_{2.4,0.7}$	0.38	0.09	0.38	0.08	0.42	0.10	0.42	0.10	0.41	0.09
$\beta_{2.4,1.4}$	0.57	0.07	0.56	0.07	0.65	0.08	0.65	0.08	0.65	0.07
$\beta_{2.4,1.9}$	0.84	0.06	0.84	0.06	0.84	0.09	0.84	0.09	0.84	0.09
$\beta_{2.4,2.4}$	1.00	0.04	1.00	0.04	1.00	0.05	1.00	0.05	1.01	0.05
$\beta_{2.4,3.4}$	0.86	0.11	0.86	0.11	0.82	0.08	0.82	0.08	0.80	0.08
$\beta_{3.4,0.7}$	0.62	0.08	0.60	0.10	0.57	0.10	0.57	0.10	0.62	0.10
$\beta_{3.4,1.4}$	0.69	0.06	0.68	0.07	0.69	0.07	0.69	0.07	0.68	0.06
$\beta_{3.4,1.9}$	0.78	0.06	0.77	0.08	0.74	0.08	0.74	0.08	0.76	0.08
$\beta_{3.4,2.4}$	0.93	0.05	0.92	0.06	0.89	0.05	0.89	0.05	0.90	0.05
$\beta_{3.4,3.4}$	1.00	0.04	1.00	0.05	1.00	0.06	1.00	0.06	1.03	0.05

3.4 Estudos Longitudinais Completos

3.4.1 Exemplo 1.1 - Concentração de bilirrubina

Consideramos aqui a análise dos dados do Exemplo 1.1. O gráfico do perfil médio da Figura 3.15 sugere que um polinômio do terceiro grau poderá representar o perfil médio da concentração de bilirrubina. O gráfico de perfis individuais na Figura 3.16 e o gráfico de perfis individuais com observações centralizadas da Figura 3.17 mostram que a variância no primeiro dia parece ser menor e aumentando até o terceiro dia.

Figura 3.15: Perfil médio da concentração de bilirrubina

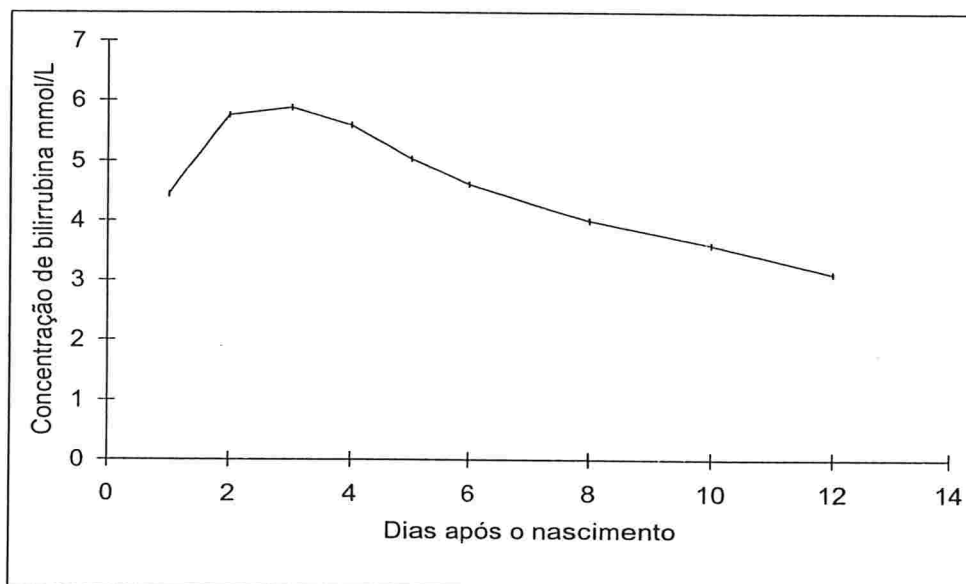


Figura 3.16: Perfis individuais da concentração de bilirrubina

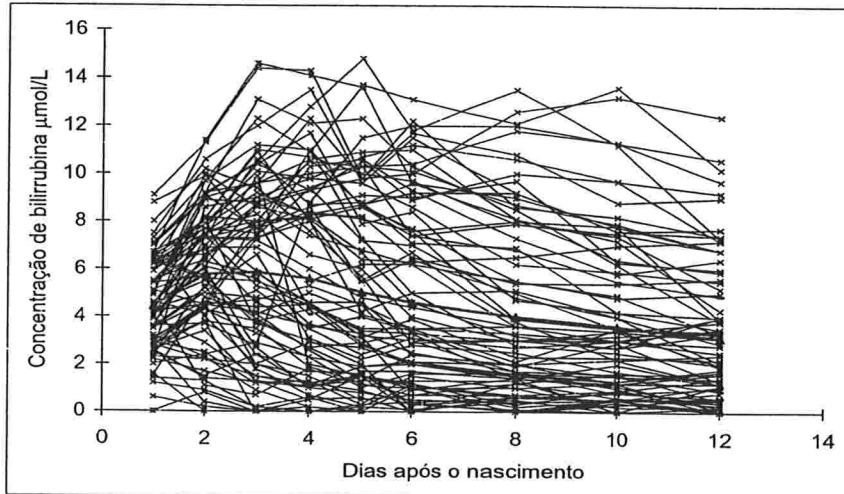
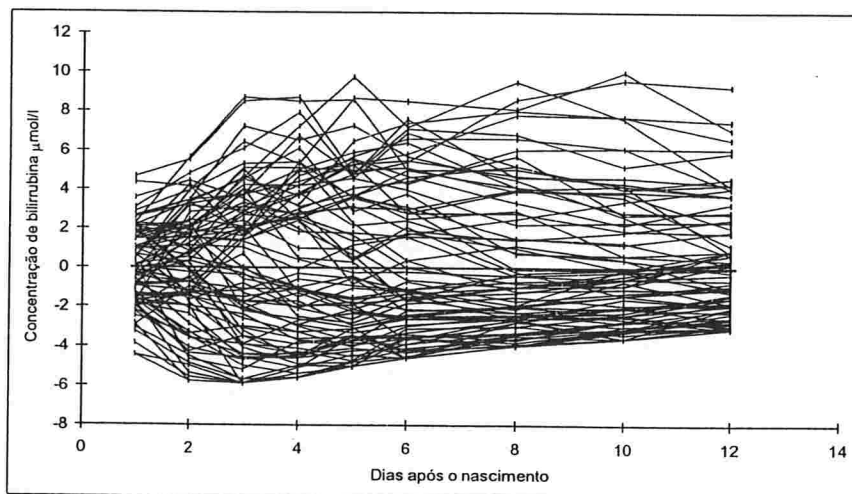


Figura 3.17: Perfis individuais da concentração de bilirrubina centralizada



O gráfico de perfis para as linhas da matriz de covariâncias amostral na Figura 3.18 sugere que elas também podem ser modeladas por polinômios cúbicos. As estimativas dos coeficientes desses polinômios, ajustados para todas as linhas da matriz de covariância amostral, estão dispostas na Tabela 3.7.

Figura 3.18: Perfis das linhas da matriz de covariância amostral

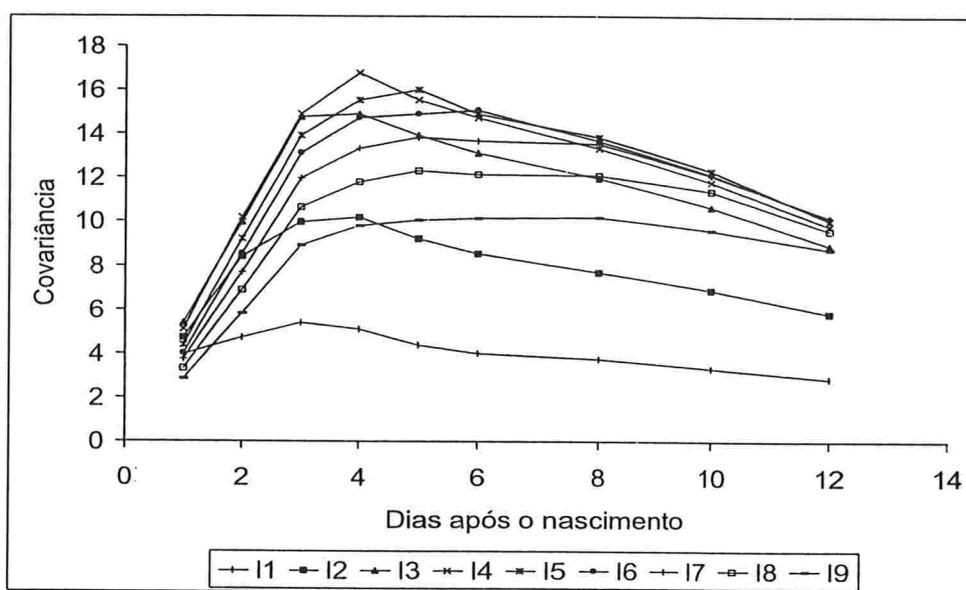


Tabela 3.7: Coeficientes dos polinômios cúbicos ajustados às linhas da matriz de covariância amostral.

Linha	Coeficientes e erros padrões (entre parênteses)				$\hat{\sigma}$	R^2
	Intercepto	Linear	Quadrático	Cúbico		
1	3.287 (0.654)	1.057 (0.431)	-0.199 (0.078)	0.009 (0.004)	0.389	86.20%
2	1.773 (1.353)	4.252 (0.891)	-0.673 (0.161)	0.029 (0.008)	0.804	88.30%
3	-0.379 (1.977)	7.310 (1.302)	-1.081 (0.235)	0.045 (0.012)	1.175	91.10%
4	-1.917 (1.748)	8.475 (1.151)	-1.214 (0.208)	0.049 (0.011)	1.039	94.90%
5	-2.571 (1.285)	8.125 (0.846)	-1.094 (0.153)	0.042 (0.008)	0.764	97.40%
6	-2.643 (0.968)	7.665 (0.637)	-1.003 (0.115)	0.038 (0.006)	0.575	98.50%
7	-1.861 (1.010)	6.466 (0.665)	-0.799 (0.120)	0.029 (0.006)	0.600	98.10%
8	-1.453 (1.007)	5.586 (0.663)	-0.681 (0.120)	0.024 (0.006)	0.598	97.60%
9	-1.124 (0.786)	4.660 (0.517)	-0.582 (0.094)	0.022 (0.005)	0.467	97.80%

Podemos observar que os valores dos coeficientes de determinação (R^2) são maiores que 85%, o que confirma o bom ajuste dos polinômios cúbicos às linhas da matriz de covariâncias amostral. Entretanto, nem todos os interceptos são significativos, o que sugere sua eliminação. Modelos sem os interceptos foram ajustados e os resultados dispostos na Tabela 3.8. Todos os coeficientes lineares, quadráticos e cúbicos são significativos e seus erros padrões se reduziram relativamente ao modelo com intercepto. Se considerarmos a

qualidade do ajuste por meio do erro padrão da regressão podemos verificar que houve aumento somente nas linhas 1 e 6 pelo fato de seus respectivos interceptos serem significativos. Assim, uma forma funcional comum adequada para todas as linhas da matriz de covariâncias amostral é um polinômio com coeficientes linear, quadrático e cúbico, sugerindo que sejam incluídos no modelo efeitos aleatórios linear, quadrático e cúbico.

Tabela 3.8: *Coefficientes dos polinômios cúbicos sem interceptos ajustados às linhas da matriz de covariância amostral.*

Linha	Coeficientes e erros padrões (entre parênteses)			$\hat{\sigma}$
	Linear	Quadrático	Cúbico	
1	3.060 (0.367)	-0.527 (0.096)	0.025 (0.006)	0.873
2	5.333 (0.357)	-0.849 (0.093)	0.037 (0.006)	0.851
3	7.079 (0.452)	-1.043 (0.118)	0.043 (0.007)	1.077
4	7.307 (0.444)	-1.023 (0.116)	0.040 (0.007)	1.056
5	6.558 (0.393)	-0.838 (0.103)	0.030 (0.006)	0.935
6	6.054 (0.348)	-0.740 (0.091)	0.025 (0.005)	0.829
7	5.332 (0.298)	-0.614 (0.078)	0.020 (0.005)	0.710
8	4.701 (0.273)	-0.536 (0.071)	0.018 (0.004)	0.650
9	3.974 (0.213)	-0.470 (0.056)	0.017 (0.003)	0.506

Analisando a matriz de covariâncias e correlações amostral (3.10), podemos observar que as variâncias crescem e depois decrescem ao longo dos dias após o nascimento, as correlações decrescem monotonicamente e as correlações em cada sub-diagonal tendem a crescer. Podemos ainda visualizar o comportamento das correlações por meio dos gráficos de dispersão múltiplos das Figuras 3.19. Este comportamento indica segundo Zimmerman et al. (1997) e Zimmerman e Núñez (2000 e 2001) que a melhor estrutura para a matriz de covariância intra-unidades amostrais é a de ante-dependência (modelo (4) da Seção 1.2).

a) Matriz de covariâncias e correlações (em negrito) amostral

$$\begin{bmatrix} 3.932 & 4.712 & 5.407 & 5.102 & 4.372 & 4.008 & 3.736 & 3.332 & 2.841 \\ \mathbf{0.819} & 8.409 & 9.979 & 10.204 & 9.216 & 8.521 & 7.700 & 6.869 & 5.812 \\ \mathbf{0.710} & \mathbf{0.896} & 14.757 & 14.890 & 13.931 & 13.120 & 11.974 & 10.665 & 8.916 \\ \mathbf{0.628} & \mathbf{0.859} & \mathbf{0.946} & 16.780 & 15.537 & 14.749 & 13.337 & 11.801 & 9.792 \\ \mathbf{0.551} & \mathbf{0.794} & \mathbf{0.906} & \mathbf{0.948} & 16.016 & 14.894 & 13.853 & 12.331 & 10.050 \\ \mathbf{0.521} & \mathbf{0.757} & \mathbf{0.880} & \mathbf{0.927} & \mathbf{0.958} & 15.077 & 13.659 & 12.139 & 10.145 \\ \mathbf{0.512} & \mathbf{0.722} & \mathbf{0.847} & \mathbf{0.885} & \mathbf{0.941} & \mathbf{0.956} & 13.543 & 12.113 & 10.199 \\ \mathbf{0.498} & \mathbf{0.702} & \mathbf{0.823} & \mathbf{0.854} & \mathbf{0.914} & \mathbf{0.927} & \mathbf{0.976} & 11.372 & 9.575 \\ \mathbf{0.485} & \mathbf{0.678} & \mathbf{0.785} & \mathbf{0.808} & \mathbf{0.849} & \mathbf{0.883} & \mathbf{0.937} & \mathbf{0.960} & 8.747 \end{bmatrix} \quad (3.7)$$

Nos gráficos de covariâncias e correlações em função da defasagem ("lag") apresentados nas Figuras 3.20 e 3.21, as correlações decrescem à medida em que as distâncias entre as observações aumentam, evidenciando uma possível correlação serial e confirmando a análise baseada na matriz de covariâncias e correlações amostral e nos gráficos de dispersão múltiplos. O variograma das observações padronizadas também apresenta indícios de correlação serial como podemos observar na Figura 3.22.

Figura 3.19: Gráficos de dispersão múltiplos com "lowess" das concentrações de bilirrubina padronizadas

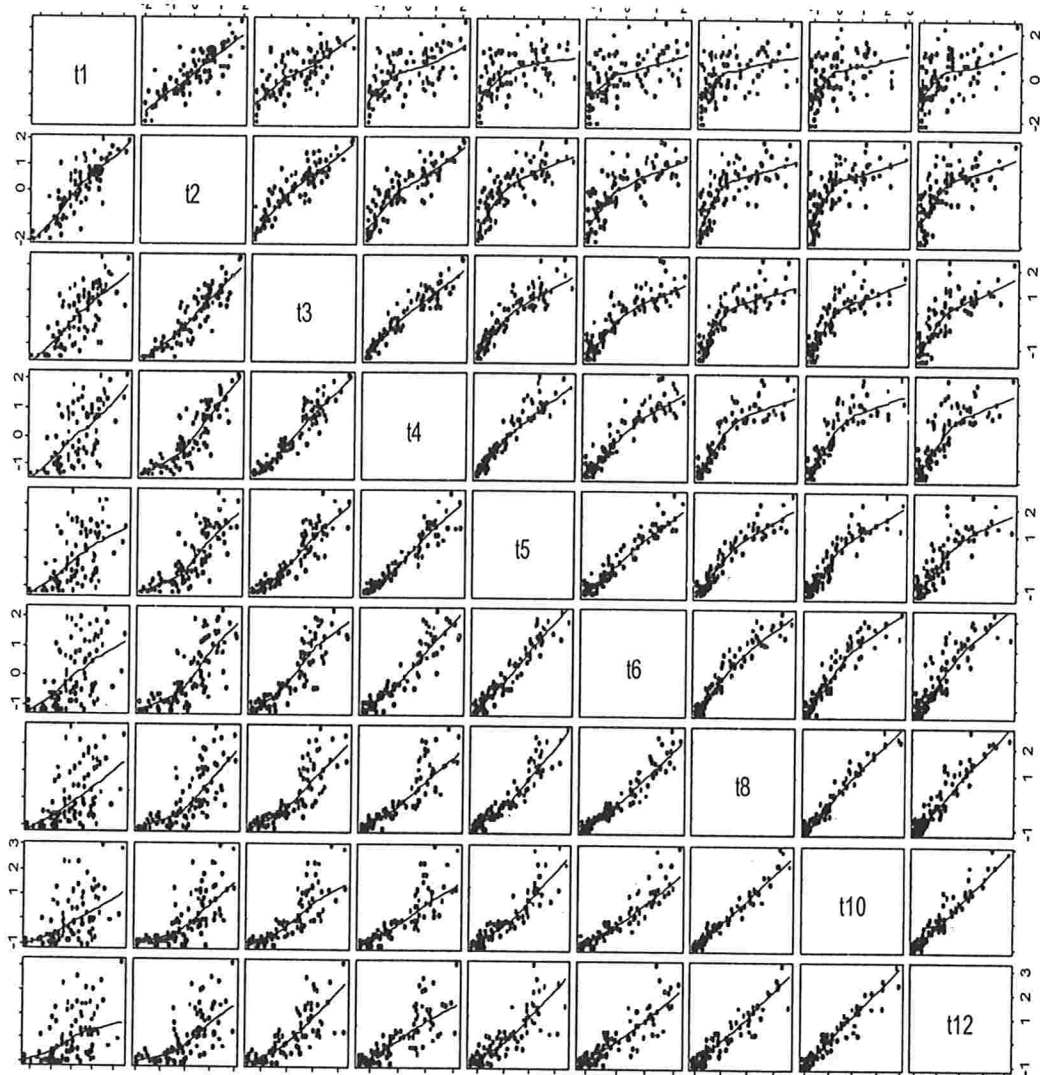


Figura 3.20: Covariâncias em função da diferença em dias

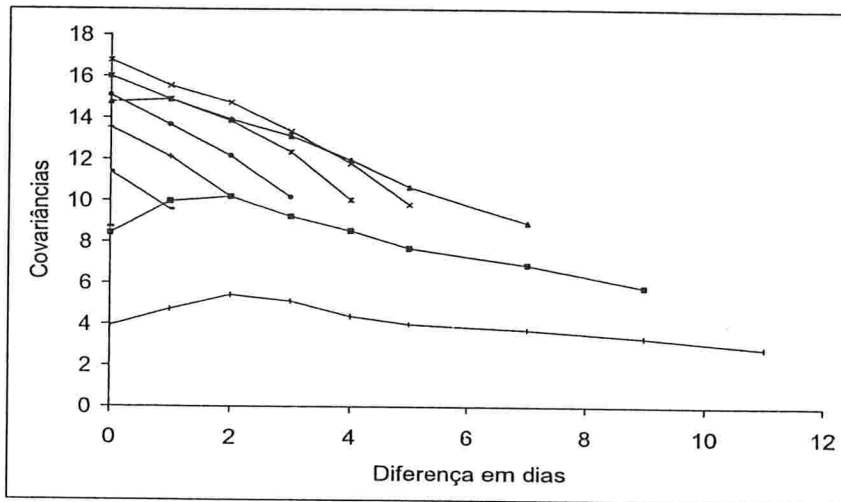


Figura 3.21: Correlações em função da diferença em dias

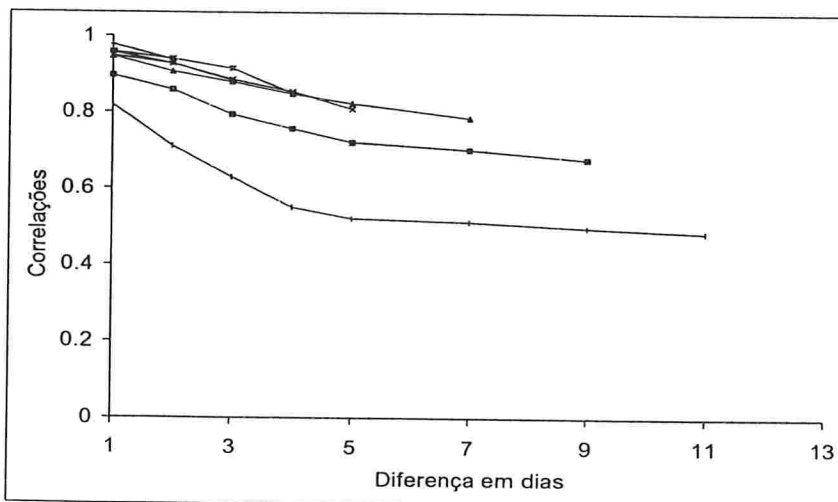
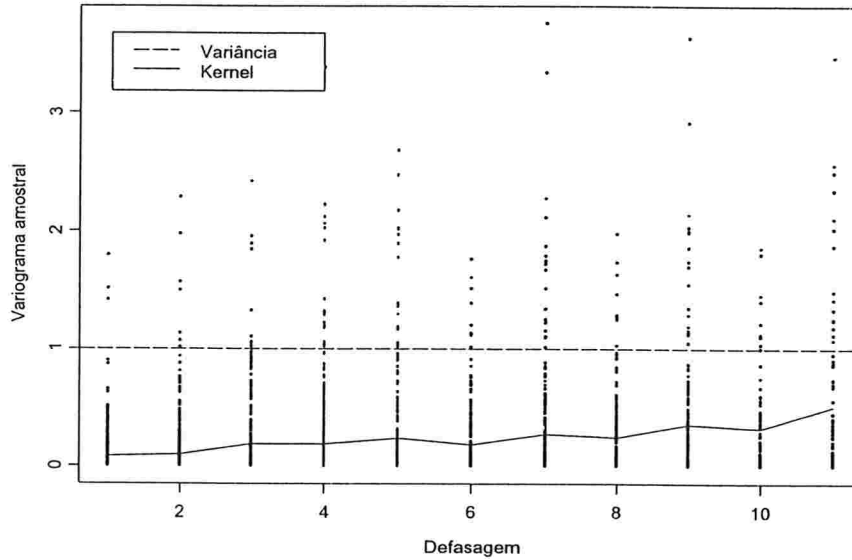


Figura 3.22: Variograma das concentrações de bilirrubina padronizadas



A princípio, consideramos um modelo com efeitos aleatórios e independência condicional a fim de observar que parcela da variabilidade amostral pode ser explicada pelos efeitos aleatórios. Propomos um modelo saturado para os efeitos fixos e um polinômio do terceiro grau sem intercepto para os efeitos aleatórios, a saber

$$y_{ij} = \mu_j + \beta_{i1}d + \beta_{i2}d^2 + \beta_{i3}d^3 + \varepsilon_{ij}, \quad (3.8)$$

em que $i = 1, \dots, 89$, $j = 1, 2, 3, 4, 5, 6, 8, 10, 12$ e d representa a variável regressora que corresponde ao número de dias após o nascimento $d = j$, $\beta_{i1} \sim N(0; \sigma_1^2)$, $\beta_{i2} \sim N(0; \sigma_2^2)$, $\beta_{i3} \sim N(0; \sigma_3^2)$ e $\varepsilon_{ij} \sim NI(0; \sigma^2)$, com os efeitos aleatórios β_{i1} , β_{i2} e β_{i3} correlacionados mas independentes de ε_{ij} . Além disso,

y_{ij} é a concentração de bilirrubina obtida na i -ésima criança no j -ésimo dia de vida;

μ_j é a concentração média de bilirrubina no j -ésimo dia;

β_{i1} é o efeito aleatório linear para a i -ésima criança;

β_{i2} é o efeito aleatório quadrático para a i -ésima criança;

β_{i3} é o efeito aleatório cúbico para a i -ésima criança;

ε_{ij} representa o erro experimental associado a cada observação y_{ij} .

A matriz de covariâncias obtida do ajuste deste modelo é apresentada a seguir.

$$\begin{bmatrix} 3.559 & 4.373 & 5.486 & 6.035 & 6.133 & 5.893 & 4.855 & 3.826 & 3.715 \\ 0.799 & 8.407 & 9.358 & 10.342 & 10.568 & 10.222 & 8.557 & 6.831 & 6.530 \\ 0.812 & 0.901 & 12.818 & 13.156 & 13.530 & 13.185 & 11.234 & 9.092 & 8.549 \\ 0.808 & 0.900 & 0.928 & 15.689 & 15.241 & 14.979 & 13.015 & 10.691 & 9.877 \\ 0.791 & 0.887 & 0.919 & 0.936 & 16.901 & 15.800 & 14.028 & 11.706 & 10.618 \\ 0.762 & 0.860 & 0.898 & 0.922 & 0.937 & 16.823 & 14.400 & 12.218 & 10.877 \\ 0.671 & 0.769 & 0.818 & 0.857 & 0.890 & 0.915 & 14.715 & 12.052 & 10.368 \\ 0.590 & 0.686 & 0.739 & 0.786 & 0.829 & 0.867 & 0.914 & 11.806 & 9.187 \\ 0.651 & 0.745 & 0.789 & 0.824 & 0.854 & 0.877 & 0.894 & 0.884 & 9.149 \end{bmatrix} \quad (3.9)$$

Podemos observar que a matriz de covariâncias e correlações estimada sob o modelo (3.8) apresenta um padrão similar ao da matriz de covariâncias amostral no que diz respeito às variâncias ao longo dos dias após o nascimento e no decréscimo das correlações ao longo das posições de cada coluna da matriz (3.9), mas não consegue modelar satisfatoriamente o crescimento das correlações ao longo das sub-diagonais. Em seguida, comparamos a estrutura de covariâncias intra-unidades amostrais correspondente ao modelo (3.8), com a estrutura covariâncias intra-unidades amostrais obtida quando o erro aleatório segue um processo de ante-dependência de ordem um (modelo (1.9) denotado por ANTE) e com outras estruturas de covariâncias intra-unidades amostrais que modelam tanto correlação serial quanto heterocedasticidade tais como: i) uma estrutura de covariâncias intra-unidades amostrais obtida quando consideramos como efeitos aleatórios os coeficientes angulares e quadráticos mas com matriz

$$\mathbf{G} = \begin{bmatrix} \sigma_{11}^2 & \sigma_{12} \\ \sigma_{12} & \sigma_{22}^2 \end{bmatrix} \quad (3.10)$$

e erro aleatório seguindo um processo auto-regressivo heterogêneo de ordem um, ou seja, com matriz \mathbf{R}_i como o modelo (1.14) (denominada EALQ+ARH1); ii) uma estrutura de covariâncias intra-unidades amostrais obtida quando o erro aleatório segue um processo auto-regressivo heterogêneo de ordem um (denominado ARH), ou seja, com \mathbf{R}_i com o modelo (1.14) e $\mathbf{G} = \mathbf{0}$; iii) uma estrutura de covariâncias intra-unidades amostrais obtida quando o erro aleatório segue um processo média móvel heterogêneo (denominado Toeplitz heterogêneo (TH)), ou seja, com \mathbf{R}_i como o modelo (1.15) e $\mathbf{G} = \mathbf{0}$ e iv) uma estrutura de covariâncias intra-unidades amostrais não estruturada (1.17) (denominada NE) para o erro aleatório.

Os resultados apresentados na Tabela 3.9 confirmam as análises anteriores, indicando que o modelo de ante-dependência de ordem um é o que melhor representa a matriz de covariâncias intra-unidades amostrais, pois também modela o comportamento das correlações ao longo das sub-diagonais.

Tabela 3.9: *Estatísticas de ajuste de modelos com diferentes estruturas de covariâncias intra-unidades amostrais.*

Estruturas	Número de parâmetros	L_c	$-2l_R$	AIC_R	BIC_R
Modelo (3.8)	7	3.2	2857.2	2871.2	2888.6
EALQ+ARH1	13	0.9	2652.9	2678.9	2711.2
ARH	10	1.5	2710.4	2730.4	2755.3
TH	17	1.1	2669.0	2703.0	2745.3
ANTE	17	0.7	2637.3	2671.3	2713.6
NE	45	0.1	2585.5	2675.5	2787.5

Na Tabela 3.10 as estimativas das médias sob cada estrutura de covariâncias intra-unidades amostrais apresentaram valores iguais, enquanto seus erros padrões sob o modelo de ante-dependência e não estruturado são iguais e próximos aos obtidos com o modelo EALQ+ARH1. Esses resultados confirmam que o modelo de ante-dependência é um forte candidato para a matriz de covariâncias intra-unidades amostrais.

Tabela 3.10: *Estimativas dos parâmetros do Exemplo 1.1 com seus respectivos erros padrões sob cada modelo para a estrutura de covariâncias intra-unidades amostrais*

Parâmetro	Estimativa	Erro Padrão					
		Inicial	EALQ+ARH1	ARH1	TH	ANTE	NE
μ_1	4.44	0.20	0.22	0.28	0.29	0.21	0.21
μ_2	5.76	0.31	0.30	0.39	0.40	0.31	0.31
μ_3	5.87	0.38	0.35	0.46	0.47	0.41	0.41
μ_4	5.59	0.42	0.38	0.44	0.45	0.44	0.44
μ_5	5.02	0.44	0.38	0.40	0.39	0.43	0.43
μ_6	4.60	0.43	0.40	0.37	0.36	0.41	0.41
μ_8	3.99	0.41	0.39	0.34	0.33	0.39	0.39
μ_{10}	3.59	0.36	0.36	0.31	0.30	0.36	0.36
μ_{12}	3.12	0.32	0.31	0.27	0.27	0.32	0.32

3.4.2 Exemplo 1.3 - Microdureza de esmalte dentário

Para esse estudo, o gráfico de perfis médios apresentado na Figura 3.23 mostra que até a profundidade de $60\mu\text{m}$ o tratamento controle *in situ* apresenta microdureza média bem inferior à dos tratamentos laser e controle extrabucal. Para profundidades maiores que $80\mu\text{m}$ não há evidências de diferença entre as microdurezas médias dos tratamentos laser e controle extrabucal; no entanto, o tratamento controle *in situ* continua associado à microdureza média inferior àquela dos outros tratamentos. Os gráficos apresentados nas Figuras 3.24, 3.25 e 3.26 sugerem um possível efeito aleatório de intercepto dado o paralelismo dos perfis individuais. O gráfico da Figura 3.27 referente às linhas da matriz de covariâncias amostral sugere que polinômios cúbicos podem se constituir em modelos adequados. Os resultados apresentados na Tabela 3.11 mostram que apenas o

intercepto é candidato a efeito aleatório, o que confirma o comportamento observado nos gráficos de perfis individuais.

Figura 3.23: Perfis médios

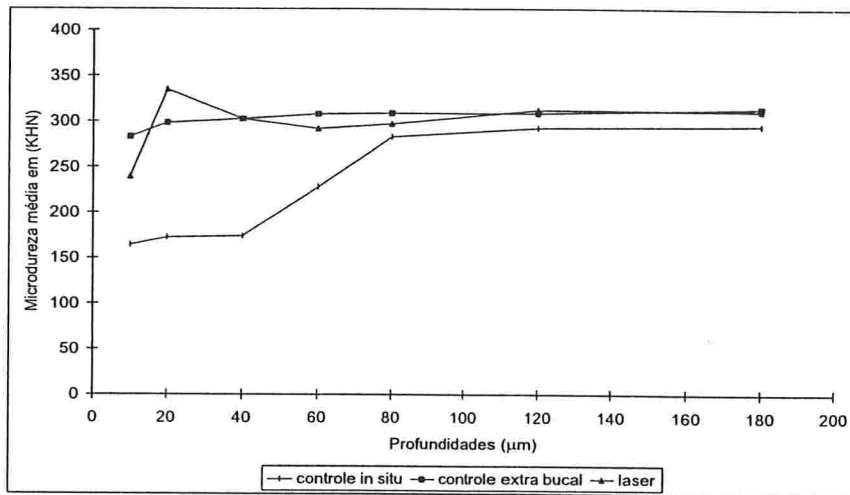


Figura 3.24: Perfis individuais para o tratamento controle *in situ*

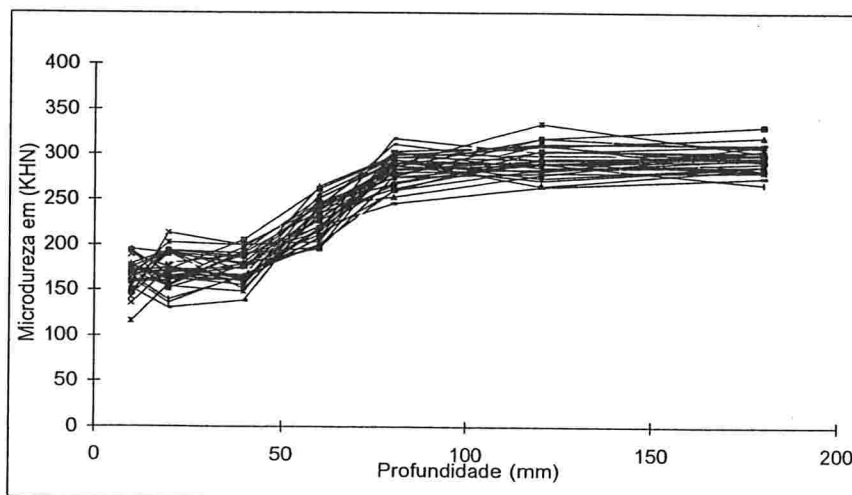


Figura 3.25: Perfis individuais para o tratamento controle extra bucal

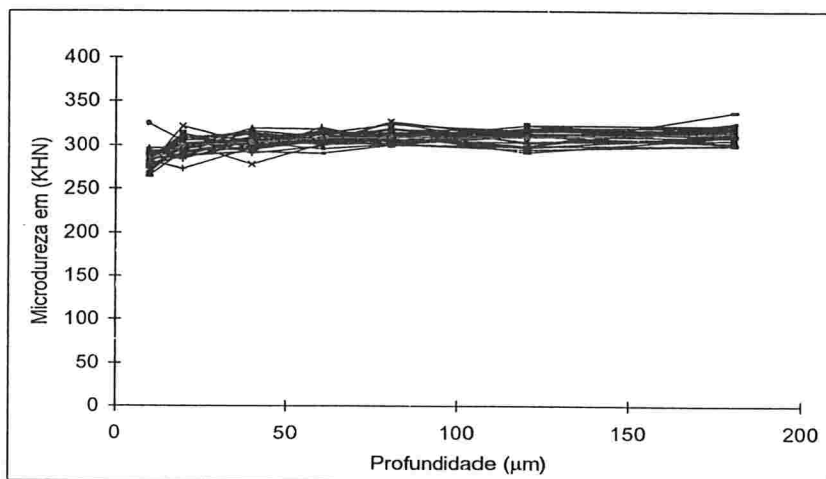


Figura 3.26: Perfis individuais para o tratamento laser

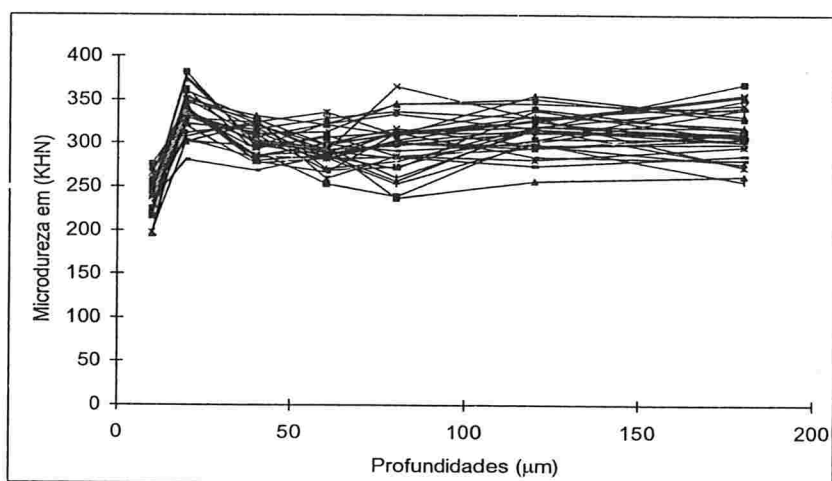


Figura 3.27: Perfis das linhas da matriz de covariâncias amostral do Exemplo 1.3

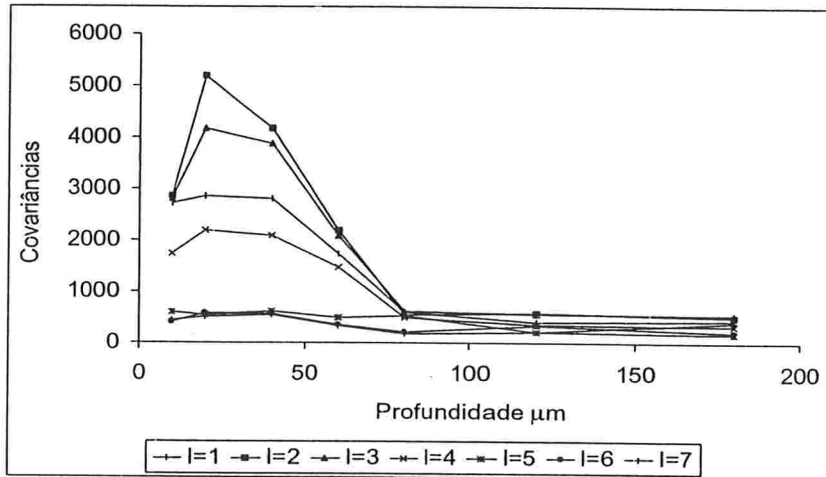


Tabela 3.11: Coeficientes dos polinômios cúbicos ajustados às linhas da matriz de covariâncias amostral (com erros padrões entre parênteses).

Linha	Intercepto	Linear	Quadrático	Cúbico	$\hat{\sigma}$	$R^2(\%)$
1	3004.20 (631.60)	-3.43 (33.51)	-0.40 (0.44)	0.0019 (0.0016)	463.5	92%
2	3889.00 (1799.00)	18.88 (95.43)	-0.90 (1.26)	0.0039 (0.0045)	1320	75.7%
3	3399.00 (1323.00)	20.36 (70.18)	-0.83 (0.93)	0.0035 (0.0033)	970.7	81.1%
4	1882.20 (542.50)	14.41 (28.78)	-0.48 (0.38)	0.0020 (0.0014)	398.1	88.4%
5	505.61 (94.410)	6.70 (5.01)	-0.13 (0.07)	0.0004 (0.0002)	69.28	85.7%
6	526.30 (188.50)	-1.54 (10.00)	-0.02 (0.13)	0.0001 (0.0005)	138.3	56.8%
7	460.60 (136.20)	2.93 (7.23)	-0.10 (0.10)	0.0005 (0.0003)	99.96	75.2%

Tendo em conta o planejamento do estudo, vamos considerar uma estrutura de covariâncias que contemple os seguintes aspectos: i) para medidas realizadas em um mesmo dente sob diferentes tratamentos, não há motivo para supor, por exemplo, que a covariância entre as observações sob o tratamento laser e controle *in situ* seja igual à covariância entre as observações sob os tratamentos laser e controle extrabucal; ii) para as medidas realizadas sob um mesmo tratamento ao longo das profundidades espera-se, por seu caráter longitudinal, um padrão de correlações decrescentes com o espaçamento temporal. Como resultado, consideramos a estrutura de covariâncias (1.20) para a matriz $\mathbf{R}(\boldsymbol{\theta})$. Medidas realizadas em um mesmo voluntário podem ser correlacionadas, o que pode ser modelado por um efeito aleatório de voluntário.

À luz dessas observações, consideramos inicialmente o seguinte modelo

$$y_{ivjp} = \mu + \alpha_v + \beta_j + \gamma_p + \beta\gamma_{jp} + \varepsilon_{ivjp} \quad (3.11)$$

em que $i = 1, \dots, 32$, $v = 1, 2, 3, 4, 5, 6, 7, 8$, $j = 1, 2, 3$ representa o índice para o fator tratamento, com níveis laser, controle *in situ* e controle extrabucal respectivamente, $p = 10, 20, 40, 60, 80, 120, 180$ são as profundidades, $\sum_{v=1}^8 \alpha_v = 0$, $\sum_{j=1}^3 \beta_j = 0$, $\sum_p \gamma_p = 0$, $\sum_{j=1}^3 \sum_p \beta\gamma_{jp} = 0$ e $\varepsilon_{iv} \sim NI(0; \mathbf{R}(\boldsymbol{\theta})_i)$ com $\mathbf{R}(\boldsymbol{\theta})_i$ tendo estrutura igual à apresentada em (1.20) (Não estruturada Kronecker auto-regressiva de ordem 1 - $NE \otimes AR(1)$). Além disso,

y_{ivjp} é o valor da microdureza em KHN obtida no i -ésimo dente submetido ao v -ésimo voluntário e ao j -ésimo tratamento na p -ésima profundidade;

μ é a microdureza média geral;

α_v é o efeito fixo do v -ésimo voluntário;

β_j é o efeito fixo do j -ésimo tratamento;

γ_p é o efeito fixo da p -ésima profundidade;

$\beta\gamma_{jp}$ é o efeito fixo de interação entre o j -ésimo tratamento e a p -ésima profundidade;

\mathbf{y}_{iv} é o vetor dos valores da microdureza em KHN obtida no i -ésimo dente submetido ao v -ésimo voluntário;

$\boldsymbol{\varepsilon}_{iv}$ é o vetor de erros experimentais associados ao vetor de observações \mathbf{y}_{iv} .

Comparamos a estrutura de covariância intra-unidades amostrais inicial com a que considera uma combinação entre um modelo não estruturado (1.17) para as medidas realizadas no mesmo dente mas em diferentes tratamentos e uma estrutura de covariâncias uniforme (1.5) para as medidas realizadas no mesmo tratamento mas em diferentes profundidades. A estrutura resultante é apresentada em (1.21) ($\text{NE} \otimes \text{Uniforme}$).

Analisando as Tabelas 3.12 e 3.13, verificamos que a estrutura de covariâncias que melhor representa a variabilidade dos dados é a $\text{NE} \otimes \text{AR}(1)$, por apresentar os menores valores nos critérios de informação e de, maneira geral, fornecer erros padrões menores para as estimativas das médias dos tratamentos em cada profundidade.

Tabela 3.12: *Estatísticas de ajuste de modelos para a estrutura de covariâncias intra-unidades amostrais*

Estruturas	Número de parâmetros	$-2\ln R$	AIC_R	BIC_R
$\text{NE} \otimes \text{AR}(1)$	7	5399.7	5413.7	5424.0
$\text{NE} \otimes \text{Uniforme}$	7	5402.1	5416.1	5426.4

Tabela 3.13: *Estimativas das microdurezas médias para os 3 tratamentos ao longo das profundidades do Exemplo 1.3 com seus respectivos erros padrões sob cada modelo para a estrutura de covariâncias intra-unidades amostrais*

Médias	Estimativa	Erro Padrão	
		NE \otimes AR(1)	NE \otimes Uniforme
$\mu_{1,10}$	239.73	4.07	4.11
$\mu_{1,20}$	334.57	4.07	4.11
$\mu_{1,40}$	302.44	4.07	4.11
$\mu_{1,60}$	292.05	4.07	4.11
$\mu_{1,80}$	296.94	4.07	4.11
$\mu_{1,120}$	312.83	4.07	4.11
$\mu_{1,180}$	311.28	4.07	4.11
$\mu_{2,10}$	163.94	3.12	3.14
$\mu_{2,20}$	171.99	3.12	3.14
$\mu_{2,40}$	174.19	3.12	3.14
$\mu_{2,60}$	227.87	3.12	3.14
$\mu_{2,80}$	282.80	3.12	3.14
$\mu_{2,120}$	293.27	3.12	3.14
$\mu_{2,180}$	295.60	3.12	3.14
$\mu_{3,10}$	283.45	1.55	1.53
$\mu_{3,20}$	298.25	1.55	1.53
$\mu_{3,40}$	302.64	1.55	1.53
$\mu_{3,60}$	307.78	1.55	1.53
$\mu_{3,80}$	309.38	1.55	1.53
$\mu_{3,120}$	309.31	1.55	1.53
$\mu_{3,180}$	314.28	1.55	1.53

3.5 Estudos Longitudinais Incompletos

Nas Figuras 3.28 e 3.29 apresentamos os gráficos de perfis médio e individuais para os dados do Exemplo 1.2. O gráfico da Figura 3.28 sugere que uma reta pode ser utilizada para modelar o perfil médio. Os perfis individuais da Figura 3.29 evoluem ao longo das semanas de forma razoavelmente paralela, sugerindo o coeficiente linear como um possível efeito aleatório. O variograma amostral na Figura 3.30 apresenta evidência de correlação serial nos dados.

Figura 3.28: Perfil médio do diâmetro diastólico da aorta por unidade de peso (AO) para os dados do Exemplo 1.2

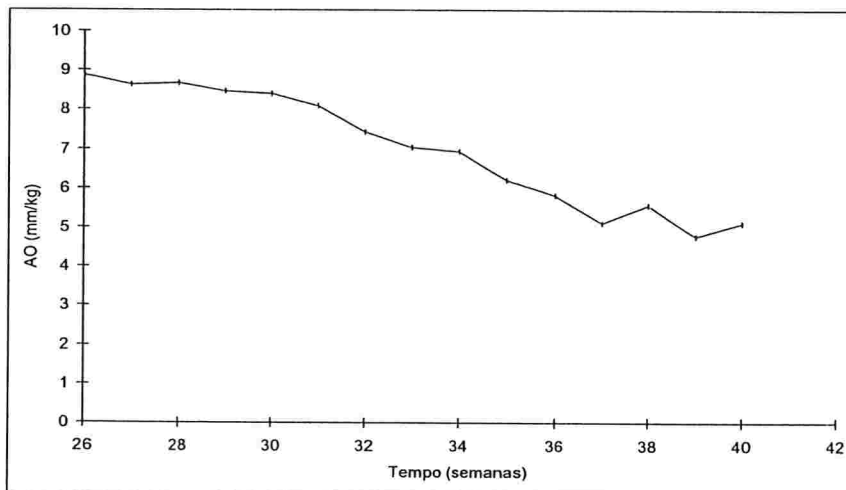


Figura 3.29: Perfis individuais do diâmetro diastólico da aorta por unidade de peso (AO) para os dados do Exemplo 1.2

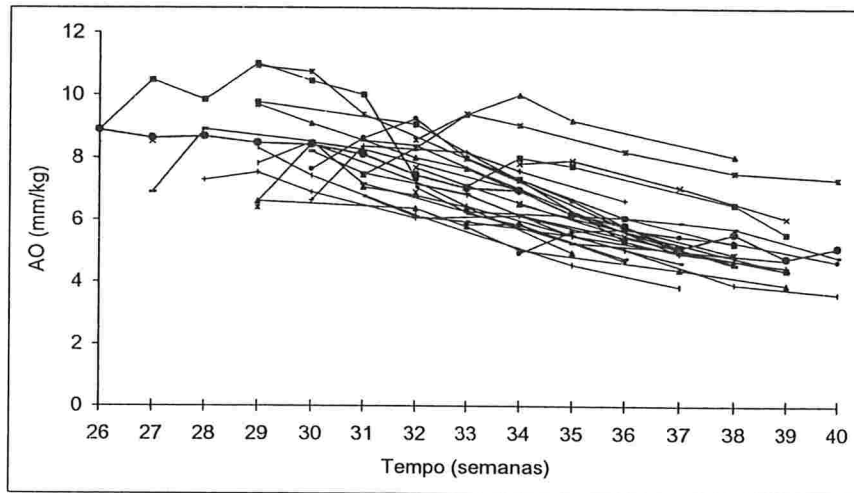
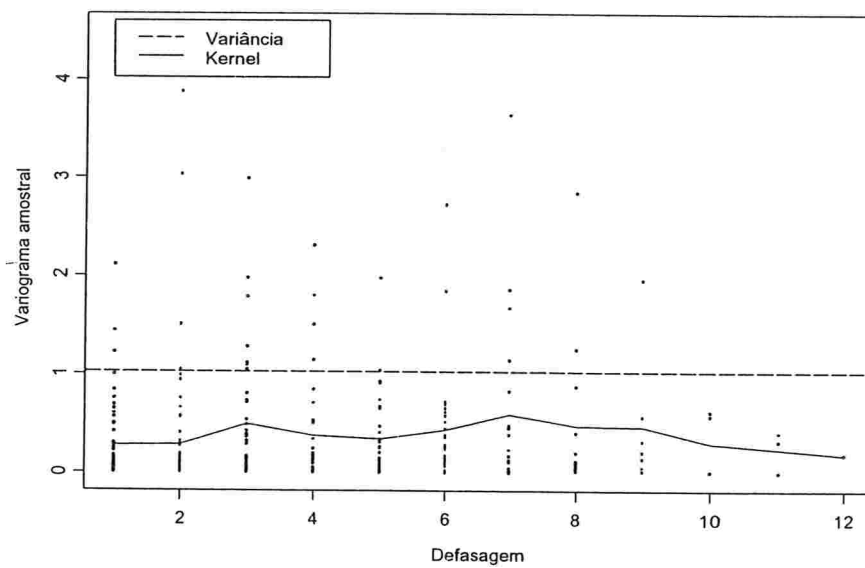


Figura 3.30: Variograma do diâmetro diastólico da aorta por unidade de peso (AO) padronizado para os dados do Exemplo 1.2



Consideramos o ajuste do modelo linear para cada perfil individual com o obje-

tivo de diagnosticar possíveis efeitos aleatórios para os coeficientes lineares e angulares. A Tabela 3.14 apresenta as estimativas dos parâmetros de cada reta ajustada. Como as médias dos coeficientes angulares e a dos coeficientes lineares são significativas ($p < 0.001$), concluímos que uma reta poderia representar o perfil médio. Para a identificação dos efeitos aleatórios, foram obtidos os desvios de cada estimativa em relação a sua respectiva média. A coluna dos desvios em torno da média das estimativas individuais dos coeficientes lineares apresentam três valores (em negrito) com desvios maiores (em valores absolutos) do que 1.96, o que representa mais de 5% das estimativas fora do intervalo, sugerindo a presença de coeficiente linear aleatório. A coluna dos coeficientes angulares apresenta apenas uma observação fora desse intervalo o que representa quase 5% das observações fora do intervalo de mais ou menos 1.96 desvio padrão dos desvios em torno da média das estimativas individuais dos coeficientes angulares, sugerindo um possível coeficiente angular aleatório. Verificamos também que mais de 10% das estimativas individuais dos coeficientes lineares e angulares possuem desvios maiores (em valor absoluto) do que 1.64 desvio padrão.

Tabela 3.14: *Estimativas dos parâmetros do modelo linear ajustado para o perfil de cada recém nascido pré-termo (RNPT) e o desvio em torno da média (valores maiores que 1.96 desvio padrão em negrito)*

RNPT	Coefficiente Angular	Desvio Coeficiente Angular	Coefficiente Linear	Desvio Coeficiente Linear
1	-0.4206	-0.0382	20.5610	0.9113
2	-0.3094	0.0730	18.0510	-1.5987
3	0.0581	0.4405	6.7639	-12.8858
4	-0.5142	-0.1318	24.0880	4.4383
5	-0.4241	-0.0417	20.9170	1.2673
6	-0.4706	-0.0882	23.0540	3.4043
7	-0.0737	0.3087	9.8982	-9.7515
8	-0.2260	0.1564	14.1150	-5.5347
9	-0.6330	-0.2506	28.3810	8.7313
10	-0.5551	-0.1727	24.1560	4.5063
11	-0.6318	-0.2494	28.5600	8.9103
12	-0.2935	0.0889	15.3310	-4.3187
13	-0.3740	0.0084	20.7560	1.1063
14	-0.4152	-0.0328	19.7280	0.0783
15	-0.3588	0.0236	17.8240	-1.8257
16	-0.4850	-0.1026	22.2500	2.6003
17	-0.3327	0.0497	17.3640	-2.2857
18	-0.3913	-0.0089	19.9200	0.2703
19	-0.5136	-0.1312	24.7030	5.0533
20	-0.5383	-0.1559	23.9240	4.2743
21	-0.2385	0.1439	16.8100	-2.8397
22	-0.0835	0.2989	8.3440	-11.3057
23	-0.2714	0.1110	15.0590	-4.5907
24	-0.3242	0.0582	16.6830	-2.9667
25	-0.4377	-0.0553	21.2020	1.5523
26	-0.3916	-0.0092	19.2130	-0.4367
27	-0.1763	0.2061	14.8290	-4.8207
28	-0.5556	-0.1732	25.8070	6.1573
29	-0.7088	-0.3264	31.5490	11.8993
Média	-0.3824	0.0000	19.6497	0.0000
Desvio padrão	0.1767	0.1736	5.8376	5.7361
Erro padrão	0.0328		1.0840	

Os gráficos envelope das Figuras 3.31 e 3.32 para os coeficientes ajustados angular e linear sugerem que as estimativas não violam a suposição de normalidade. Essa suposição é utilizada quando testamos a ausência de efeito aleatório verificando, por exemplo, se 5% dos desvios são maiores (em valores absolutos) que 1.96 desvio padrão.

Figura 3.31: Gráfico normal de probabilidades com envelope para as estimativas dos coeficientes lineares

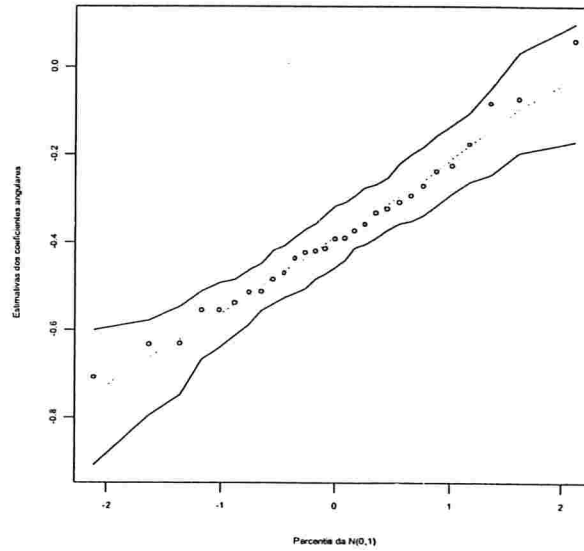
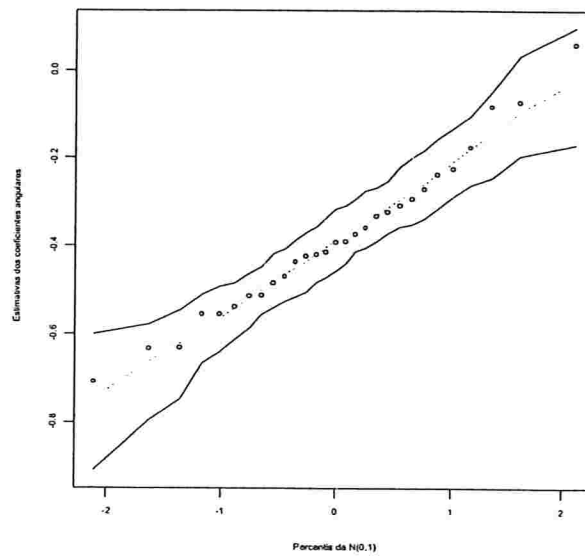


Figura 3.32: Gráfico normal de probabilidades com Envelope para as estimativas dos coeficientes angulares



Essas análises iniciais sugerem o seguinte modelo:

$$y_{is} = \mu_s + \alpha_i + \beta_i s + \varepsilon_{is} \quad (3.12)$$

$i = 1, \dots, 32$, $s = 26, \dots, 40$, $\alpha_i \sim N(0; \sigma_1^2)$, $\beta_i \sim N(0; \sigma_2^2)$, $\varepsilon_{is} \sim NI(0; \sigma^2)$ com os efeitos aleatórios α_i e β_s correlacionados ($cov(\alpha_i, \beta_s) = \sigma_{12}$) mas independentes de ε_{ij} , em que.

y_{is} é o diâmetro diastólico por unidade de peso (AO) em mm/kg obtida no i -ésimo recém nascido pré-termo (RNPT) na s -ésima semana de vida;

μ_s é o diâmetro diastólico médio por unidade de peso (AO) em mm/kg na s -ésima semana;

α_i é o efeito aleatório linear para o i -ésimo RNPT;

β_i é o efeito aleatório angular para o i -ésimo RNPT;

ε_{is} representa o erro experimental associado a cada observação y_{is} .

Após o ajuste, realizamos o teste da razão de verossimilhanças para verificar se a inclusão dos efeitos aleatórios para os coeficientes lineares e angulares é razoável. Para o coeficiente linear, o valor da estatística da razão de verossimilhanças foi 77.2, que quando comparado a uma mistura 50% : 50% de distribuições qui-quadrado χ_0^2 e χ_1^2 fornece nível descritivo ($p < 0.001$), indicando que devemos incluir efeito aleatório de coeficiente linear. Para o coeficiente angular na presença do efeito aleatório para o coeficiente linear, o valor da estatística razão de verossimilhanças foi 20.4, que comparado a uma mistura 50% : 50% de distribuições qui-quadrado χ_1^2 e χ_2^2 resulta no nível descritivo ($p < 0.001$), indicando também que devemos incluir no modelo o efeito aleatório de coeficiente angular. Esses resultados confirmam a análise dos perfis individuais realizada anteriormente.

Em seguida, comparamos a estrutura de covariâncias intra-unidades amostrais obtida sob modelo (3.15), com uma estrutura de covariâncias intra-unidades que modela tanto heterocedasticidade quanto correlação serial, obtida quando consideramos os efeitos aleatórios do modelo (3.15) mais erros aleatórios seguindo um processo auto-regressivo de ordem um, que denominamos (EALA+AR1), além de estruturas que modelam apenas a

correlação serial como as estruturas de covariâncias intra-unidades amostrais obtidas quando o erro aleatório segue um processo auto-regressivo de ordem um (AR1 - modelo (1.6)), markoviano de ordem um (MARKOV - modelo (1.12)) ou auto-regressivo de médias móveis de ordem um (ARMA11 - modelo (1.7)). A Tabela 3.16 abaixo apresenta os valores dos critérios de informação para os ajustes desses modelos. Podemos observar que as estruturas de covariâncias AR1 e MARKOV apresentaram os mesmos valores no número de parâmetros e nos critérios de informação e os menores valores de AIC_R e BIC_R dentre as estruturas ajustadas, indicando que uma dessas estruturas poderá ser escolhida como a mais adequada para modelar a variabilidade dos dados. Realizamos o teste da razão de verossimilhanças usual, para verificarmos se o modelo ARMA11 pode ser reduzido para o modelo AR1, uma vez que são estruturas encaixadas. O valor observado da estatística foi 0.92 ($p=0.3375$), indicando que podemos ficar apenas com a estrutura auto-regressiva de ordem 1. Isso permite restringir as estruturas de covariância possíveis apenas para a estrutura de efeito aleatório, efeito aleatório mais o erro aleatório seguindo um processo auto-regressivo de ordem um, a Markov e a auto-regressiva de ordem 1.

Concluimos por meio das Tabelas 3.16 e 3.17 que as estruturas de covariâncias intra-unidades que melhor modelam a matriz de covariâncias intra-unidades amostrais dos dados são as obtidas quando consideramos o erro aleatório seguindo um processo auto-regressivo de ordem um ou um processo markoviano, por apresentarem os menores valores dos critérios de informação e produzirem, em geral, erros padrões das médias menores em relação aos das outras estruturas.

Tabela 3.16: Estatísticas de ajuste dos modelos ajustados para o exemplo 1.2

Estruturas	Número de parâmetros	$-2l_R$	AIC_R	BIC_R
EALA	4	366.76	374.76	380.62
EALA+AR1	5	359.28	369.28	376.61
AR1	2	362.18	366.18	369.11
MARKOV	2	362.18	366.18	369.11
ARMA11	3	361.26	367.26	371.65

Tabela 3.17: Estimativas e erros padrões (EP) dos diâmetros diastólicos médios por unidade de peso (em mm/kg) ao longo das semanas com seus respectivos erros padrões sob cada modelo para a estrutura de covariâncias intra-unidades amostrais.

Parâmetros	Estruturas de covariâncias							
	EALA		EALA+AR1		AR1		MARKOV	
	Estimativa	EP	Estimativa	EP	Estimativa	EP	Estimativa	EP
μ_{26}	5.97	0.82	5.68	0.86	6.46	0.77	6.46	0.77
μ_{27}	7.50	0.56	7.29	0.59	7.63	0.53	7.63	0.53
μ_{28}	8.32	0.47	7.85	0.47	7.98	0.44	7.98	0.44
μ_{29}	8.24	0.31	8.13	0.31	8.22	0.30	8.22	0.30
μ_{30}	8.14	0.28	8.07	0.27	8.11	0.27	8.11	0.27
μ_{31}	7.77	0.26	7.81	0.25	7.81	0.26	7.81	0.26
μ_{32}	7.38	0.22	7.36	0.22	7.35	0.24	7.35	0.24
μ_{33}	7.03	0.21	6.99	0.21	6.98	0.23	6.98	0.23
μ_{34}	6.75	0.22	6.65	0.22	6.64	0.24	6.64	0.24
μ_{35}	6.17	0.24	6.12	0.23	6.12	0.25	6.12	0.25
μ_{36}	5.66	0.26	5.69	0.25	5.70	0.26	5.70	0.26
μ_{37}	5.32	0.28	5.25	0.27	5.28	0.27	5.28	0.27
μ_{38}	4.98	0.31	5.05	0.30	5.12	0.28	5.12	0.28
μ_{39}	4.59	0.39	4.48	0.39	4.61	0.35	4.61	0.35
μ_{40}	4.20	0.48	4.45	0.52	4.74	0.46	4.74	0.46

Capítulo 4

Conclusões e Sugestões

Em geral, para estudos não longitudinais, por possuírem caráter experimental, a matriz de covariâncias pode ser modelada com o auxílio de informações sobre o planejamento e obtidas da análise da matriz de covariâncias amostral.

Os gráficos de perfis médios e individuais das observações são ferramentas descritivas importantes por fornecerem respectivamente informações sobre o comportamento das respostas médias e individuais ao longo das condições de avaliação. Os gráficos de perfis individuais das observações e das observações centralizadas também podem indicar heterocedasticidade e possíveis efeitos aleatórios, além de poderem ser utilizados em estudos incompletos e não balanceados.

Em estudos longitudinais, a análise descritiva das matrizes de covariâncias e correlações amostrais poderá fornecer informações sobre heterocedasticidade intra e entre unidades amostrais, correlação serial e também sobre o comportamento das correlações ao longo das sub-diagonais. O gráfico de dispersão múltiplo das observações padronizadas sugerido por Dawson et al. (1997) é uma ferramenta importante por expressar graficamente o comportamento da matriz de correlações amostrais.

O gráfico de perfis das linhas da matriz de correlações amostrais em função da defasagem como sugerem Gray e Helms (1995) é útil no diagnóstico de correlação serial. Entretanto, o gráfico de perfis das linhas da matriz de correlações amostrais e o de dispersão múltipla apresentam o inconveniente de só poderem ser aplicados em estudos

longitudinais completos. Assim também consideramos a análise do variograma amostral que pode ser aplicado tanto em estudos longitudinais completos quanto incompletos como proposto por Diggle (1988 e 1990) e Diggle et al. (2000). Nesse trabalho propomos a construção do variograma amostral a partir das observações padronizadas visando o diagnóstico de correlação serial.

As técnicas apresentadas para a identificação de possíveis efeitos aleatórios em estudos longitudinais apresentadas no Capítulo 2 são importantes principalmente quando o estudo é incompleto, permitindo, em conjunto com as análises acima citadas, indicar possíveis estruturas de covariâncias intra-unidades amostrais. Para selecionarmos aquelas que melhor representam a variabilidade dos dados, utilizamos os critérios de informação.

Por fim, notamos que a modelagem da estrutura de covariância intra-unidades amostrais em geral fornece estimativas mais precisas para os efeitos fixos, possibilitando assim inferências mais confiáveis. Além disso, as características induzidas pelo planejamento do experimento podem ser modeladas de forma mais adequada, sem admitirmos as usuais suposições de independência ou homocedasticidade muitas vezes distantes da realidade.

Tanto na identificação de padrões nas correlações das medidas realizadas na mesma unidade amostral, quanto em novos modelos para a estrutura de covariâncias intra-unidades amostrais que incorporem os padrões diagnosticados, existem muitos aspectos que precisam ser estudados, tais como: i) técnicas gráficas para diagnosticar, por exemplo, o erro aleatório seguindo um processo auto-regressivo de ordem maior que um e ii) novos modelos de estruturas de covariâncias para estudos com medidas repetidas com dois ou mais fatores intra-unidades amostrais que possam incorporar heterocedasticidade em ambos os fatores.

Bibliografia

- [1] Arnold, S. F. (1980). *The Theory of Linear Models and Multivariate Analysis*. New York: John Wiley & Sons.
- [2] Andrade, D. F. (1984). *ML Estimation and LR Tests for the Multivariate Normal Distribution with Patterned Mean and Covariance Matrix. Complete and Incomplete-Data Cases*. Tese de Doutorado. North Carolina: Chapel Hill. Department of Biostatistics.
- [3] Bartlett, M. S. (1937). Some examples of statistical methods of research in agriculture and applied biology. *Journal of the Royal Statistical Society, Supplement*, **4**, 137-170.
- [4] Callanan, T. P. and Harville, D. A. (1991). Some new algorithms for computing restricted maximum likelihood estimates of variance components. *Journal of Statistical Computation and Simulation*, **38**, 239-259.
- [5] Chi, E. M. and Reinsel, G. C. (1989). Models for longitudinal data with random effects and AR(1) errors. *Journal of the American Statistical Association*, **84**, 452-459.
- [6] Chiu, T. Y. M., Leonard, T. and Tsui, K. (1996). The matrix-logarithmic covariance model, *Journal of the American Statistical Association*, **91**, 198-210.
- [7] Cook, N. R. and Ware, J. H. (1983). Design and analysis methods for longitudinal research. *Annual Review of Public Health*, **4**, 1-24.

- [8] Cox, D. R. and Hinkley, D. V. (1974) *Theoretical Statistics*. London: Chapman & Hall/CRC.
- [9] Christopher, H. M., Jay D. P. and Larry, J. B. (1997). Linear transformations of linear mixed model. *American Statistical Association*, **51**, 338-343.
- [10] Crowder, M. J. and Hand, D. J. (1990). *Analysis of Repeated Measures*. London: Chapman & Hall.
- [11] Crowder, M. J. and Hand, D. J. (1996). *Practical Longitudinal Data Analysis*. London: Chapman & Hall.
- [12] Cúri, M. (1999). *Medidas Repetidas com Covariáveis Dependentes do Tempo: Um Exemplo com Dados Incompletos*. Dissertação de Mestrado. São Paulo: IME-USP.
- [13] Dawson, K. S., Gennings, C. and Carter, W. H. (1997). Two graphical techniques useful in detecting correlation structure in repeated measures data. *Statistical Computing and Graphics*, **51**, 275-283.
- [14] Dempster, A. P., Laird, N. M. and Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, B*, **39**, 1-38.
- [15] Diggle, P. J. (1988). An approach to the analysis of repeated measurements. *Biometrics*, **44**, 959-971.
- [16] Diggle, P. J. (1989). Testing for dropouts in repeated measurement data. *Biometrics*, **45**, 1255-1258.
- [17] Diggle, P. J. (1990). *Time Series A Biostatistical Introduction*. Oxford: Clarendon Press.
- [18] Diggle, P. J. and Kenward, M. G. (1994). Informative dropouts in longitudinal data analysis (with discussion). *Applied Statistics*, **43**, 49-93.

- [19] Diggle, P. J.; Liang, K. and Zeger, S. L. (2002). *Analysis of Longitudinal Data*. Oxford: Clarendon Press.
- [20] Davis, C. S. (2002). *Statistical Methods for the Analysis of Repeated Measurements*. New York: Springer.
- [21] Draque, C. M., Almeida, M. F. B., Sañudo, A., Peres, C. A., Santos, A. M. N. e Kopelman, B.I. (2001). Curva-padrão de bilirrubina total sérica até 12 dias de vida em recém-nascidos a termo saudáveis em aleitamento materno. *Trabalho apresentado no XVII Congresso Brasileiro de Perinatologia*. Florianópolis.
- [22] Gabriel, K. R.G. (1962). Ante-dependence analysis of an ordered set of variables. *Annals of Mathematical Statistics*, **33**, 201-212.
- [23] Galecki, T. G. (1994). General class of covariance structures for tow or more repeated factors in longitudinal data analysis. *Communications in Statistics -Theory and Methods*, **23**, 3105-3119.
- [24] Giampaoli V. (1999). *Inferência estatística para modelos lineares com restrições nos parâmetros em condições regulares e não regulares*. Dissertação de Mestrado. São Paulo: IME-USP.
- [25] Graybill, F. A. (1976). *Theory and Application of the Linear Model*. North Scituate: Duxbury Classic Series.
- [26] 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.
- [27] Gregoire, T. G., Brillinger, D. R. Diggle, P. J. Russek-Cohen, E., Warren, W. G. and Wolfinger, R. D. (editores). *Modelling longitudinal and spatially correlated data - methods, applications, and future directions*. New York: Springer.
- [28] Goodnight, J. H. and Hemmerle, W. J. (1979). A simplified algorithm for the W-transformation in variance component estimation. *Thechnometrics*, **21**, 265-268.

- [29] Hartley, H. O. and Hocking, R. R. (1967). Maximum likelihood estimation for the mixed analysis of variance model. *Biometrika*, **54**, 93-108.
- [30] Harville, D. A. (1974). Bayesian inference for variance components using only error contrasts. *Biometrika*, **61**, 383-385.
- [31] Harville, D. A. (1976). Extension of the Gauss-Markov theorem to include the estimation of random effects. *Annals of Statistics*, **4**, 384-395.
- [32] Harville, D. A. (1977). Maximum likelihood approaches to variance component estimation and to related problems. *Journal of the American Statistical Association*, **72**, 320-340.
- [33] Harville, D. A. (1997). *Matrix Algebra from a Statistician's Perspective*. New York: Springer-Verlag.
- [34] Henderson, C. R. (1975). Best linear unbiased estimation and prediction under a selection model. *Biometrics*, **31**, 423-447.
- [35] Henderson, C. R. Jr. (1982). Analysis of covariance in the mixed model: Higher-level, nonhomogeneous, and random regression. *Biometrics*, **38**, 623-640.
- [36] Henderson, C. R. (1984). *Applications of Linear Models in Animal Breeding*. Guelph: University of Guelph.
- [37] Hamilton, J. D. (1994). *Time Series Analysis*. New Jersey: Princenton University Press.
- [38] Jennrich, R. I. and Schluchter, M. D. (1986). Unbalanced repeated-measures models with structured covariance matrices. *Biometrics*, **42**, 805-820.
- [39] Jones, R. H. (1990). Serial correlation or random subject Effects? *Communications in Statistics Simulation*, **19**, 1105-1123.

- [40] Jones, R. H and Boadi-Boateng, F. (1991). Unequally spaced longitudinal data with AR(1), *Biometrics*, **47**, 161-175.
- [41] Jones, R. H. (1993). *Longitudinal Data with Serial Correlation: A State-Space Approach*. London: Chapman & Hall.
- [42] Kackar, R. N. and Harville, D. A. (1984). Fixed and random effects in mixed linear models. *Journal of the American Statistical Association*, **79**, 854-862.
- [43] Keselman, H. J., James A., Rhonda K. K. and Russell D. W. (1998). A comparison of two approaches for selecting covariance structures in the analysis of repeated measurements. *Communications in Statistics Simulation*, **27**, 591-604.
- [44] Khattree, R. and Naik, D. N. (1999). *Applied Multivariate Statistics with SAS® software*. Cary: SAS Institute.
- [45] Kenward, M. G. (1987). A Method for Comparing Profiles of Repeated Measurements. *Applied Statistics*, **36**, 296-308.
- [46] Korytnicki, L. D. (2002). *Esmalte irradiado ou não com laser de Nd:YAG: análise microbiológica do biofilme dental, mensuração da dureza knoop em seção transversal e observação ao microscópio de luz polarizada: estudo 'in situ'*. Tese de Doutorado. São Paulo: Faculdade de Odontologia-USP.
- [47] Laird, N. M. (1982). Computation of variance components using EM algorithm. *Journal of Statistical Computation and Simulation*, **14**, 295-303.
- [48] Laird, N. M., Lange, N. and Stram, D. (1987). Maximum likelihood computations with repeated measures: Application of the EM algorithm. *Journal of the American Statistical Association*, **82**, 97-105.
- [49] Laird, N. M. and Ware, J. H. (1982). Random-effects models for longitudinal data. *Biometrics*, **38**, 963-974.
- [50] Lange, K. (1999). *Numerical Analysis for Statisticians*. New York: Springer-Verlag.

- [51] Laster, L. I. and Pickands III, J. (1993). Repeated measurement designs with random selections. *Biometrics*, **49**, 75-83.
- [52] Lindley, D. V. and Smith, A. F. M. (1972). Bayes estimates for the linear model. *Journal of the Royal Statistical Society, B*, **34**, 1-41.
- [53] Lindstrom, M. J. and Bates, D. M. (1988). Newton-Raphson and EM algorithms for linear mixed-effects models for repeated-measures data. *Journal of the American Statistical Association*, **83**, 1014-1022.
- [54] Littell, R. C., Milliken, G. A., Stroup, W. W. and Wolfinger, R. D. (1996). *SAS System for Mixed Models*. Cary: SAS Institute.
- [55] Littell, R. C., Pendergast, J. and Natarajan, R. (2000). Tutorial in Biostatistics - Modelling covariance structure in the analysis of repeated measures data. *Statistics in Medicine*, **19**, 1793-1819.
- [56] Little, R. J. A. and Rubin, B. D. (2002). *Statistical Analysis with Missing Data*. Second Edition. New Jersey: John Wiley & Sons.
- [57] Little, R. J. A. (1988). A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association*, **83**, 1198-1202.
- [58] Longford, N. T. (1993). *Random Coefficient Models*. Oxford: Clarendon Press.
- [59] McLean, R. A., Sanders, W. L. and Stroup, W. W. (1991). A unified approach to mixed linear models. *The American Statistician*, **45**, 54-65.
- [60] Morrison, D. F. (1990). *Multivariate Statistical Methods*. New York: McGraw-Hill.
- [61] Nelder, J. A. and Mead, R. (1965) A simplex method for function minimization. *The Computer Journal*, **7**, 941-946.
- [62] Noble, B. e Daniel, J. W. (1977). *Applied Linear Algebra*. Englewood Cliffs: Prentice Hall.

- [63] Öfversten, J. (1993). Exact tests for variance components in unbalanced mixed models. *Biometrics*, **49**, 45-57.
- [64] Patterson, H. D. and Thompson, R. (1971). Recovery of interblock information when block sizes are unequal. *Biometrika*, **58**, 545-554.
- [65] Robinson, D. L. (1987). Estimation and use of variance components. *The Statistician*, **36**, 3-14.
- [66] Robinson, G. K. (1991). That BLUP is a good thing: the estimation of random effects. *Statistical Science*, **6**, 15-51.
- [67] Rochon, J. and Helms, R. W. (1989). Maximum likelihood estimation for incomplete repeated-measurements under an ARMA covariance structure. *Biometrics*, **45**, 207-218.
- [68] Rubin, D. B. (1976). Inference and missing data. *Biometrika*, **63**, 581-592.
- [69] Rutter, C. M. and Elashoff, R. M. (1994). Analysis of longitudinal data: random coefficient regression modelling. *Statistics in Medicine*, **13**, 1211-1231.
- [70] Schwarz, C. J. (1993). The mixed-model ANOVA: the truth, the computer packages, the books - Part I: Balanced data. *The American Statistician*, **47**, 48-64.
- [71] Searle, S. R. (1971). *Linear Models*. New York: John Wiley & Sons.
- [72] Searle, S. R., Casella, G. and McCulloch, C. E. (1992). *Variance Components*. New York: John Wiley & Sons.
- [73] Sef, H. (1999). *Modelos multiplicativos de regressão para dados pré-teste/pós-teste em blocos*. Dissertação de Mestrado. São Paulo: IME-USP.
- [74] Self, S.G. and Liang, K. (1987). Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions. *American Statistical Association*, **82**, 605-610.

- [75] Singer, J. M. and Andrade, D. F. (1986). *Análise de Dados Longitudinais*. Campinas: VII Simpósio Nacional de Probabilidade e Estatística.
- [76] Singer, J. M. and Andrade, D. F. (2000). Analysis of Longitudinal Data. *Handbook of Statistics*, 18, 115-160.
- [77] Silveira, G. B. (1992). *Estimação de Densidades e de Funções de Regressão*. Rio de Janeiro: 10 Simpósio Nacional de Probabilidade e Estatística.
- [78] Singer, J. D. (1998). Using SAS PROC MIXED to fit multilevel models, hierarchical models, and individual growth models. *Journal of Educational and Behavioral Statistics*, 24, 323-355.
- [79] Smith, A. F. (1973). A general bayesian linear model. *Journal of the royal statistical society, B*, 35, 67-75.
- [80] Solange, A. (1989). *Modelos de Efeitos Aleatórios para Análise de dados Longitudinais Não Balanceados em Relação ao Tempo*. Dissertação de Mestrado. São Paulo: IME-USP.
- [81] Suyama, E. (1995). *Identificação de um Modelo de Efeitos Aleatórios*. Tese de Doutorado. São Paulo: IME-USP.
- [82] Stram, D. O. and Lee, J. W. (1994). Variance components testing in the longitudinal mixed effects model. *Biometrics*, 50, 1171-1177.
- [83] 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-USP.
- [84] Timothy, G. G., Brillinger, D. R., Diggle, P. J., Russek-Cohen, E., Warren, W. G. and Wolfinger, R. D. (1997). *Modelling Longitudinal and Spatially Correlated Data - Methods , Applications, and Future Directions*. New York: Springer-Verlag.

- [85] Verbeke, G. and Molenberghs (1997). *Linear Mixed Models in Practice A SAS-Oriented Approach*. New York: Springer-Verlag.
- [86] Verbeke, G., Spiessens, B. and Lesaffre, E. (2001). Conditional linear mixed models. *The American Statistician*, **55**,25-34.
- [87] Vanleeuwen, Dawn M. (1998). A note on the covariance structure in linear model. *The American Statistician*, **51**, 140-144.
- [88] Vonesh, E. F. and Chinchilli, V. M. (1997). *Linear and Non-Linear Models for the Analysis of Repeated Measurements*. New York: Marcel Dekker.
- [89] Venables, W. N. and Ripley, B. D. (1999). *Modern Applied Statistics with S-Plus*. New York: Springer.
- [90] Vu, H. T. V. and Zhou, S. (1997). Generalization of likelihood ratio tests under nonstandard conditions. *The Annals of Statistics*, **25**, 897-916.
- [91] Ware, J. H. (1985). Linear models for the analysis of longitudinal studies. *American Statistical Association*, **39**, 95-101.
- [92] Ware, J. H., Dockery, D. W., Louis, T. A., Xu, X., Ferris Jr., B. G. and Speizer, F. E. (1990). Longitudinal and cross-sectional estimates of pulmonary function decline in never-smoking adults. *American Journal of Epidemiology*, **132**, 685-700.
- [93] Winer, B. J. (1971). *Statistical Principles in Experimental Designs*. New York: McGraw-Hill.
- [94] Wolfinger, R. D. (1993). Covariance structure selection in general mixed model. *Communications in Statistics Simulation*, **22**, 1079-1106.
- [95] Wolfinger, R., Tobias, R. and Sall, J. (1994). Computing gaussian likelihoods and their derivatives for general linear mixed models. *SIAM Journal of Scientific Computation*, **15**, 1294-1310.

- [96] Wolfinger, R. D. (1996). Heterogeneous variance-covariance structures for repeated measures. *Journal of Agricultural, Biological, and Environmental Statistics*, **1**, 205-230.
- [97] Zimmerman, D. L., Nuñez-Antón, V. and El-Barmi, H. (1997). Computational aspects of likelihood-based estimation of first-order antedependence order. *Journal of Statistical Computation and Simulation*, **60**, 67-84.
- [98] Zimmerman, D. L. and Nuñez-Antón (2000). Modeling nonstationary longitudinal Data. *Biometrics*, **56**, 699-705.
- [99] Zimmerman, D. L. and Nuñez-Antón (2001). Parametric modelling of growth curve data: An Overview. *Test*, **10**, 111-999.