**Identificação de um Modelo de
Efeitos Aleatórios**

Emílio Suyama

TESE DEFENDIDA
NO
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA
DA
UNIVERSIDADE DE SÃO PAULO
PARA OBTENÇÃO DO GRAU DE DOUTOR
EM
ESTATÍSTICA

Área de Concentração: **Estatística**
Orientador: **Prof. Dr. Júlio da Motta Singer**
São Paulo, Julho de 1995

Identificação de um Modelo de Efeitos Aleatórios

Este exemplar corresponde à redação
da tese devidamente corrigida
e defendida por Emílio Suyama
e aprovada pela Comissão Julgadora.

São Paulo, 11 de agosto de 1995.

Banca Examinadora:

Prof. Dr. Júlio da Motta Singer (orientador)

Prof. Dr. Dalton Francisco de Andrade

Prof. Dr. Hélio dos Santos Migon

Prof. Dr. Ruben Klein

Prof. Dr. Jorge Alberto Achcar

Resumo

No contexto de dados longitudinais gerados por modelos probabilísticos gaussianos, desenvolvemos uma técnica para identificação de um modelo para os efeitos aleatórios através da inspeção da matriz de dispersão amostral; em alguns casos particulares, essa técnica é estendida para a análise da estrutura de dispersão condicional.

Em muitos problemas envolvendo dados longitudinais nota-se que a correlação entre observações feitas em dois instantes, t_s e t_u na mesma unidade amostral é positiva e decresce à medida em que a distância entre t_s e t_u aumenta. Nesse contexto, consideramos o ajuste de retas aos dados e comparamos as estruturas de dispersão induzidas pelo modelo usual de efeitos aleatórios com independência condicional e pelo modelo de efeitos aleatórios com erros gerados por um Movimento Browniano (MB). Mostramos que dentre esses dois modelos, apenas o segundo acomoda o padrão descrito acima. A escolha entre essas estruturas de dispersão condicional pode ser feita na prática através da inspeção da matriz de dispersão amostral.

Para modelar um conjunto de dados relativos a um processo de cicatrização de úlceras corneanas induzidas em coelhos, estudamos o comportamento da matriz de dispersão marginal induzida por um modelo de efeitos aleatórios com estrutura de dispersão condicional dada por um processo de Ornstein-Ühlenbeck integrado (OU), que é identificado através da inspeção da matriz de dispersão amostral. Consideramos uma distribuição normal truncada para modelar o comportamento estritamente positivo das observações que são realizadas até completa cicatrização em cada coelho. Para contornar a intratabilidade analítica da verossimilhança associada à normal multivariada truncada, utilizamos um método de estimação em dois estágios. No primeiro, o ajuste de curvas individuais sugere que uma mistura do modelo de OU com o de MB seria mais apropriado. No segundo estágio, ajustamos um modelo de efeitos aleatórios ao par formado pelo tamanho da ferida inicial e pela velocidade média.

Abstract

In the context of longitudinal data generated by Gaussian probability models, we develop an identification technique for a random effects model by inspection of the sample dispersion matrix; in some particular cases, this technique is extended for the conditional dispersion structure analysis.

In many problems involving longitudinal data, the correlation between observations made on two different time points, t_s and t_u on the same sampling unit is positive and decreases as the distance between t_s and t_u increases. In this context, we consider the fitting of straight lines to the data and compare the dispersion structures induced by the usual conditional independence random effects model and the random effects model with the conditional errors generated by a Brownian Motion (BM). We show that between the two models, only the second one accommodates the pattern described above. The choice of one of these models for the conditional dispersion structure can be made in practice by inspection of the sample dispersion matrix.

To model a data set related to regeneration process of corneal ulcer induced in rabbits, we study the behavior of the marginal dispersion induced by a random effects model with the conditional dispersion structure given by an integrated Ornstein-Ühlenbeck process (OU), which is identified by inspection of the sample dispersion matrix. We consider a truncated normal distribution to model the strictly positive values of observations taken until complete regeneration in each rabbit. To avoid the analytical intractability associated to the truncated multivariate normal likelihood, we use a two-stage estimation method. In the first stage, the adjustment of individual curves suggest that a mixture of OU and BM models should be more appropriate. In the second stage we adjust a random effects model to the pair formed by the initial ulcer size and the mean velocity of regeneration.

Sumário

Capítulo 1: Introdução.	1
1.1 Natureza de dados longitudinais	1
1.2 Modelos para dados longitudinais	12
1.3 Inferência estatística	18
1.4 Métodos para identificação , seleção e diagnóstico de modelos .	29
1.5 Propostas deste trabalho	31
Capítulo 2: Identificação de efeitos aleatórios.	34
2.1 Introdução	34
2.2 Identificação através de regressão linear simples	36
2.3 Identificação separada de modelos para efeitos aleatórios e para efeitos fixos	40
2.4 Identificação de um modelo para efeitos aleatórios através da análise da matriz de dispersão amostral	43
2.5 Estimação e comparação entre modelos	47
2.6 Conclusão e discussão	54
Capítulo 3: Identificação da estrutura de dispersão condicional.	57
3.1 Introdução	57
3.2 Modelo de efeitos aleatórios com independência condicional . .	59
3.3 Modelo de efeitos aleatórios com Movimento Browniano	62
3.4 Identificação da dispersão condicional	69
3.5 Comparação entre os modelos	72
3.6 Conclusão e discussão	76
Capítulo 4: Uma aplicação do modelo normal truncado	77
4.1 Introdução	77
4.2 Descrição do modelo	79
4.3 Identificação do modelo	81
4.4 Análises individuais	89

4.5	Estimação dos parâmetros do modelo pa- ra efeitos aleatórios	
(4.2.3)		101
4.6	Conclusão e sugestões para análises complementares	107
Capítulo 5:	Conclusão e Sugestões	112
Apêndice A:	Processos de Ornstein-Ühlenbeck	115
Apêndice B:	Expansão linear de $\beta_{1j} (\lambda_j = \lambda)$	118
Apêndice C:	Desenvolvimento de $\text{Var}(\mathbf{y}_j)$	119
Apêndice D:	Expansão linear de $\hat{\iota}_{0j}$ e de $\hat{\beta}_{1j}$	122
Referências		123

Capítulo 1

Introdução

1.1 Natureza de dados longitudinais

Estudos longitudinais são investigações científicas em que o interesse é dirigido para a avaliação do comportamento de uma ou mais variáveis respostas ao longo do tempo ou de uma outra dimensão específica, como a distância de uma certa origem ou a dosagem de uma determinada substância (Andrade & Singer, 1986). Tais estudos têm sido considerados em diversas áreas, sob denominações diferentes: *estudo de coortes*, em pesquisas sociológicas ou biomédicas, onde também é denominado *estudo do tipo “follow-up”*, e *painel* em amostragem (Dwyer & Feinleib, 1992).

Dados (provenientes de estudos) *longitudinais* se distinguem pela sequência de duas ou mais observações em cada unidade amostral, às vezes chamada de *perfil individual de respostas* (Kenward, 1987). Diz-se que os dados longitudinais são *(ir)regulares no tempo* se o intervalo entre duas medidas consecutivas quaisquer (não) for constante ao longo do estudo, e em relação ao planejamento, *(não) balanceados com relação ao tempo* se as observações (não) forem feitas nos mesmos instantes de tempo em todas as unidades amostrais. Se (não) houver observações perdidas diz-se que a estrutura dos dados é *(completa) incompleta*. Devido à natureza de estudos longitudinais, que podem durar meses ou até anos, e requerem observações nas mesmas unidades amostrais em múltiplas ocasiões, torna-se freqüente a existência de dados incompletos, e a incorporação de dados observacionais no estudo pode torná-los não balanceados e irregulares no tempo (Grady, 1992).

Exemplos de dados longitudinais de interesse neste trabalho são:

Exemplo 1: Num estudo genético, foi observado o crescimento em peso (em gramas) de 14 camundongos nas idades de 2, 5, 8, 11, 14, 17 e 20 dias (ver Tabela 1.1). Estes dados foram analisados por Rao (1987), Izenman & Williams (1989) e Weiss & Lazaro (1992).

Tabela 1.1: Pesos de camundongos (g) - Exemplo 1.

Unidade amostral	Idade (dias)						
	2	5	8	11	14	17	20
1	19.0	38.8	62.1	82.3	107.8	113.2	119.1
2	21.8	39.3	56.8	72.9	83.9	85.2	100.4
3	21.1	39.4	54.9	70.0	78.3	87.0	92.5
4	20.9	41.9	64.5	85.0	100.1	102.6	106.9
5	19.3	36.2	52.0	53.0	64.1	64.0	75.1
6	20.1	36.1	50.2	53.0	65.7	76.2	88.8
7	20.2	37.0	49.8	65.0	79.5	85.8	91.0
8	19.0	35.0	51.0	66.6	81.9	87.9	92.9
9	21.9	39.9	57.8	69.9	70.9	82.2	95.3
10	22.5	40.0	54.5	69.0	79.6	82.5	83.6
11	22.4	38.1	57.7	75.6	86.9	92.9	99.9
12	18.7	32.9	44.1	52.5	58.9	62.1	79.6
13	27.8	47.1	60.6	77.0	88.8	100.1	110.5
14	14.1	26.0	47.2	66.2	76.0	88.5	87.8
Média	20.6	37.7	54.5	68.4	80.2	86.4	94.5

Exemplo 2: Para se estabelecer um padrão ortodôntico para o crescimento de um osso (Ramus), seu comprimento (em mm) foi medido nas idades de 8, 8.5, 9 e 9.5 anos numa coorte de 20 meninos (ver Tabela 1.2). Estes dados foram analisados por Elston & Grizzle (1962), Grizzle & Allen (1969), Lee & Geisser (1975) e Rao (1965 e 1987).

Exemplo 3: Num estudo oftalmológico, 28 coelhos foram aleatoriamente alocados em 4 grupos de tratamento com 7 animais cada e em seguida tiveram úlceras com formas circulares induzidas em cada um dos olhos. Os tamanhos das feridas foram medidos ao longo do tempo, com o objetivo de determinar o(s) melhor(es) tratamento(s) em relação ao tempo de cicatrização (ver Tabelas 1.3 a 1.10). Apesar de originalmente terem sido medidas áreas de feridas, Suyama et al. (1986) mostraram ser mais conveniente utilizar suas raízes quadradas, que são proporcionais aos respectivos diâmetros, por serem aproximadamente lineares em relação ao tempo. Por isso, as medidas apresentadas nas tabelas abaixo são das raízes quadradas das medidas originais apresentadas por Araújo (1987). Para cada animal, são apresentados os ins-

Tabela 1.2: Comprimento do osso Ramus (mm) - Exemplo 2.

Unidade amostral	Idade (anos)			
	8	8.5	9	9.5
1	47.8	48.8	49.0	49.7
2	46.4	47.3	47.7	48.4
3	46.3	46.8	47.8	48.5
4	45.1	45.3	46.1	47.2
5	47.6	48.5	48.9	49.3
6	52.5	53.2	53.3	53.7
7	51.2	53.0	54.3	54.5
8	49.8	50.0	50.3	52.7
9	48.1	50.8	52.3	54.4
10	45.0	47.0	47.3	48.3
11	51.2	51.4	51.6	51.9
12	48.5	49.2	53.0	55.5
13	52.1	52.8	53.7	55.0
14	48.2	48.9	49.3	49.8
15	49.6	50.4	51.2	51.8
16	50.7	51.7	52.7	53.3
17	47.2	47.7	48.4	49.5
18	53.3	54.6	55.1	55.3
19	46.2	47.5	48.1	48.4
20	46.3	47.6	51.3	51.8
Média	48.7	49.6	50.6	51.4

tantes de medida, os valores medidos em cada um dos olhos e os valores médios. Qualquer valor médio sublinhado corresponde a uma observação com medida apenas num dos olhos, tendo sido considerada nula a medida não realizada.

Os Exemplos 1 e 2 estão associados a conjuntos de dados univariados, completos, balanceados e regulares no tempo, ao passo que os dados do Exemplo 3 são bivariados, não balanceados, irregulares no tempo e incompletos.

Tabela 1.3: $\sqrt{Area}$ da úlcera: Tratamento 1 (continua) - Ex. 3.

Unidade amostral	variável	Observação						
		1	2	3	4	5	6	7
1	tempo (h)	0.000	11.933	19.583	29.016	37.699	41.932	47.832
	olho dir.	45.826	43.151	35.777	30.017	24.759	22.158	17.804
	olho esq.	45.453	39.724	36.633	34.132	31.032	27.803	22.803
	média	45.639	41.437	36.205	32.074	27.895	24.981	20.304
2	tempo (h)	0.000	12.233	19.483	29.066	37.583	41.900	47.700
	olho dir.	48.847	42.024	39.573	36.401	28.636	26.000	24.617
	olho esq.	48.466	44.181	39.837	31.969	27.731	23.917	20.809
	média	48.657	43.103	39.705	34.185	28.183	24.958	22.713
3	tempo (h)	0.000	12.150	19.417	28.934	37.451	41.784	47.601
	olho dir.	45.815	38.092	35.693	32.450	29.086	25.476	21.610
	olho esq.	45.111	40.902	38.145	29.326	27.203	24.454	22.294
	média	45.463	39.497	36.919	30.888	28.144	24.965	21.952
4	tempo (h)	0.000	12.100	19.383	28.900	37.417	41.734	47.584
	olho dir.	44.317	36.083	31.717	24.980	15.556	11.045	8.660
	olho esq.	48.010	41.940	33.407	25.923	23.173	18.894	14.491
	média	46.164	39.012	32.562	25.451	19.365	14.970	11.576
5	tempo (h)	0.000	12.033	19.333	28.883	37.333	41.700	47.633
	olho dir.	44.181	36.715	32.465	28.792	25.768	21.817	18.221
	olho esq.	45.749	41.073	36.837	32.818	28.337	22.694	20.273
	média	44.965	38.894	34.651	30.805	27.053	22.256	19.247
6	tempo (h)	0.000	11.950	19.250	28.783	37.250	41.600	47.533
	olho dir.	50.120	43.070	36.056	27.018	22.760	19.621	14.595
	olho esq.	47.329	42.579	37.497	30.561	22.672	19.131	15.906
	média	48.724	42.824	36.776	28.790	22.716	19.376	15.250
7	tempo (h)	0.000	11.917	19.250	28.750	37.267	41.584	47.534
	olho dir.	43.806	38.210	33.541	29.411	25.671	22.428	20.249
	olho esq.	45.738	42.662	38.328	33.362	28.213	22.672	18.412
	média	44.772	40.436	35.934	31.386	26.942	22.550	19.330
Média	tempo (h)	0.000	12.045	19.386	28.905	37.429	41.748	47.631
	olho dir.	46.130	39.621	34.975	29.867	24.605	21.221	17.965
	olho esq.	46.551	41.866	37.241	31.156	26.909	22.795	19.284
	média	46.341	40.743	36.107	30.511	25.757	22.008	18.625

Tabela 1.3: $\sqrt{Area}$ da úlcera: Tratam. 1 (continuação) - Ex. 3.

Unidade amostral	variável	Observação						
		8	9	10	11	12	13	14
1	tempo (h)	55.165	59.715	67.082	72.049	-	-	-
	olho dir.	13.856	10.440	6.557	2.646	-	-	-
	olho esq.	18.248	13.711	7.211	2.236	-	-	-
	média	16.052	12.076	6.884	2.441	-	-	-
2	tempo (h)	55.017	59.567	67.034	71.917	83.984	93.784	109.751
	olho dir.	23.345	20.125	18.028	15.100	12.329	8.775	4.690
	olho esq.	17.464	14.177	10.050	7.681	5.291	2.828	1.414
	média	20.405	17.151	14.039	11.390	8.810	5.802	3.052
3	tempo (h)	54.918	59.468	66.935	71.835	83.868	90.868	-
	olho dir.	18.303	13.856	10.863	6.403	2.000	-	-
	olho esq.	19.875	19.774	18.303	15.297	12.166	5.000	-
	média	19.089	16.815	14.583	10.850	7.083	<u>2.500</u>	-
4	tempo (h)	54.884	59.417	66.884	-	-	-	-
	olho dir.	6.928	4.000	1.732	-	-	-	-
	olho esq.	6.164	-	-	-	-	-	-
	média	6.546	<u>2.000</u>	<u>0.866</u>	-	-	-	-
5	tempo (h)	54.850	59.367	66.817	71.750	83.767	90.767	-
	olho dir.	15.492	12.000	8.602	6.325	3.464	-	-
	olho esq.	17.861	15.033	12.124	10.100	6.481	2.450	-
	média	16.676	13.517	10.363	8.212	4.972	<u>1.225</u>	-
6	tempo (h)	54.750	59.267	66.700	-	-	-	-
	olho dir.	9.219	3.162	-	-	-	-	-
	olho esq.	11.958	7.874	3.162	-	-	-	-
	média	10.589	5.518	<u>1.581</u>	-	-	-	-
7	tempo (h)	54.734	59.234	66.651	71.584	83.534	93.451	-
	olho dir.	16.031	13.454	10.583	7.874	4.583	2.646	-
	olho esq.	14.629	11.790	9.055	6.708	3.742	1.732	-
	média	15.330	12.622	9.819	7.291	4.162	2.189	-
Média	tempo (h)	54.903	59.434	66.872	-	-	-	-
	olho dir.	14.739	11.005	<u>8.052</u>	-	-	-	-
	olho esq.	15.171	<u>11.766</u>	<u>8.558</u>	-	-	-	-
	média	14.955	<u>11.385</u>	<u>8.305</u>	-	-	-	-

Tabela 1.4: $\sqrt{Area}$ da úlcera: Tratamento 2 (continua) - Ex. 3.

Unidade amostral	variável	Observação						
		1	2	3	4	5	6	7
8	tempo (h)	0.000	12.750	19.283	28.900	37.217	41.500	47.467
	olho dir.	44.170	36.000	29.631	20.494	16.371	11.402	7.874
	olho esq.	43.012	36.742	30.298	22.891	19.442	16.155	12.649
	média	43.591	36.371	29.965	21.692	17.906	13.779	10.262
9	tempo (h)	0.000	11.900	19.367	28.800	37.117	41.384	47.367
	olho dir.	47.074	39.383	30.545	22.847	16.000	11.445	6.245
	olho esq.	41.340	35.553	30.232	18.276	16.062	12.124	7.681
	média	44.207	37.468	30.389	20.561	16.031	11.785	6.963
10	tempo (h)	0.000	12.017	19.367	28.767	37.084	41.551	47.551
	olho dir.	46.141	39.674	33.045	22.159	18.439	15.653	12.490
	olho esq.	48.021	41.569	38.639	35.199	30.659	24.658	20.396
	média	47.081	40.621	35.842	28.679	24.549	20.155	16.443
11	tempo (h)	0.000	12.050	19.367	28.784	37.067	41.534	47.534
	olho dir.	45.618	40.175	34.986	26.683	22.249	16.000	11.135
	olho esq.	43.886	37.630	31.496	22.361	16.673	12.207	8.307
	média	44.752	38.902	33.241	24.522	19.461	14.103	9.721
12	tempo (h)	0.000	12.083	19.366	28.783	37.083	41.533	47.566
	olho dir.	44.497	39.535	34.235	28.107	23.388	18.248	14.697
	olho esq.	41.146	36.083	32.802	28.983	25.120	15.621	11.402
	média	42.822	37.809	33.518	28.545	24.254	16.934	13.049
13	tempo (h)	0.000	12.050	19.350	28.767	37.050	41.600	47.550
	olho dir.	46.043	43.612	35.384	32.848	26.153	19.647	16.310
	olho esq.	45.155	38.639	35.242	29.950	26.683	22.716	18.762
	média	45.599	41.126	35.313	31.399	26.418	21.181	17.536
14	tempo (h)	0.000	12.033	19.333	28.816	37.033	41.600	47.567
	olho dir.	49.639	42.732	34.627	24.104	16.492	13.229	10.000
	olho esq.	49.244	41.449	36.865	24.536	19.925	14.000	7.550
	média	49.441	42.090	35.746	24.320	18.209	13.614	8.775
Média	tempo (h)	0.000	12.126	19.348	28.802	37.093	41.529	47.515
	olho dir.	46.169	40.159	33.207	25.320	19.870	15.089	11.250
	olho esq.	44.544	38.238	33.654	26.028	22.081	16.783	12.392
	média	45.356	39.198	33.431	25.674	20.975	15.936	11.821

Tabela 1.4: $\sqrt{Area}$ da úlcera: Tratam. 2 (continuação) - Ex. 3.

Unidade amostral	variável	Observação						
		8	9	10	11	12	13	14
8	tempo (h)	54.650	-	-	-	-	-	-
	olho dir.	3.606	-	-	-	-	-	-
	olho esq.	4.123	-	-	-	-	-	-
	média	3.864	-	-	-	-	-	-
9	tempo (h)	54.550	-	-	-	-	-	-
	olho dir.	2.000	-	-	-	-	-	-
	olho esq.	2.646	-	-	-	-	-	-
	média	2.323	-	-	-	-	-	-
10	tempo (h)	54.751	58.934	66.467	71.300	83.267	-	-
	olho dir.	10.000	7.483	3.742	1.000	-	-	-
	olho esq.	15.748	11.045	8.660	6.000	3.317	-	-
	média	12.874	9.264	6.201	3.500	<u>1.658</u>	-	-
11	tempo (h)	54.734	58.917	66.450	-	-	-	-
	olho dir.	8.544	5.196	1.732	-	-	-	-
	olho esq.	4.796	2.450	-	-	-	-	-
	média	6.670	3.823	<u>0.866</u>	-	-	-	-
12	tempo (h)	55.499	58.882	66.465	-	-	-	-
	olho dir.	11.489	7.746	4.472	-	-	-	-
	olho esq.	8.307	4.243	2.000	-	-	-	-
	média	9.898	5.994	3.236	-	-	-	-
13	tempo (h)	55.467	58.934	66.434	71.184	83.134	-	-
	olho dir.	12.767	9.592	5.568	-	-	-	-
	olho esq.	14.832	10.296	7.141	4.243	2.646	-	-
	média	13.800	9.944	6.355	<u>2.121</u>	<u>1.323</u>	-	-
14	tempo (h)	55.434	-	-	-	-	-	-
	olho dir.	4.359	-	-	-	-	-	-
	olho esq.	3.162	-	-	-	-	-	-
	média	3.761	-	-	-	-	-	-
Média	tempo (h)	55.012	-	-	-	-	-	-
	olho dir.	7.538	-	-	-	-	-	-
	olho esq.	7.659	-	-	-	-	-	-
	média	7.598	-	-	-	-	-	-

Tabela 1.5: $\sqrt{Area}$ da úlcera: Tratamento 3 (continua) - Ex. 3.

Unidade amostral	variável	Observação						
		1	2	3	4	5	6	7
15	tempo (h)	0.000	11.817	18.950	28.617	36.534	41.201	47.118
	olho dir.	42.166	35.707	30.529	25.456	20.736	17.205	13.748
	olho esq.	45.122	39.141	35.440	29.275	20.905	12.845	5.745
	média	43.644	37.424	32.984	27.365	20.820	15.025	9.746
16	tempo (h)	0.000	11.750	18.900	28.550	36.450	41.100	47.050
	olho dir.	44.777	37.670	29.749	26.646	23.022	20.075	16.823
	olho esq.	45.011	37.802	30.806	26.230	22.405	18.466	13.784
	média	44.894	37.736	30.277	26.438	22.713	19.270	15.303
17	tempo (h)	0.000	11.733	19.116	28.533	34.950	41.083	47.116
	olho dir.	46.011	38.484	35.242	31.654	26.401	22.627	18.303
	olho esq.	44.452	38.197	32.435	27.386	24.413	19.925	18.083
	média	45.231	38.340	33.838	29.520	25.407	21.276	18.193
18	tempo (h)	0.000	11.733	19.133	28.516	34.949	41.099	47.082
	olho dir.	46.530	41.605	36.290	32.970	29.360	25.060	22.428
	olho esq.	47.896	42.214	37.723	33.660	30.315	27.037	20.928
	média	47.213	41.909	37.007	33.315	29.837	26.048	21.678
19	tempo (h)	0.000	11.683	19.066	28.449	34.916	41.299	47.282
	olho dir.	54.598	50.448	46.109	40.927	34.857	28.036	18.894
	olho esq.	46.658	40.423	36.180	30.050	24.839	20.833	14.629
	média	50.628	45.435	41.144	35.488	29.848	24.434	16.762
20	tempo (h)	0.000	11.683	19.233	28.466	34.916	41.283	47.283
	olho dir.	46.206	37.670	33.076	27.586	21.977	18.276	11.958
	olho esq.	45.760	39.128	30.741	25.100	20.273	15.199	8.775
	média	45.983	38.399	31.908	26.343	21.125	16.737	10.367
21	tempo (h)	0.000	11.683	19.216	28.449	34.899	41.282	47.315
	olho dir.	43.267	36.111	26.096	18.628	13.416	9.381	4.796
	olho esq.	45.989	34.914	27.477	21.494	14.967	10.863	6.557
	média	44.628	35.513	26.787	20.061	14.191	10.122	5.677
Média	tempo (h)	0.000	11.726	19.088	28.511	35.373	41.192	47.178
	olho dir.	46.222	39.671	33.870	29.124	24.253	20.094	15.279
	olho esq.	45.841	38.831	32.972	27.599	22.588	17.881	12.643
	média	46.032	39.251	33.421	28.361	23.420	18.988	13.961

Tabela 1.5: $\sqrt{Area}$ da úlcera: Tratam. 3 (continuação) - Ex. 3.

Unidade amostral	variável	Observação						
		8	9	10	11	12	13	14
15	tempo (h)	54.968	58.468	65.918	70.651	-	-	-
	olho dir.	10.100	6.856	4.243	2.236	-	-	-
	olho esq.	2.236	-	-	-	-	-	-
	média	6.168	<u>3.428</u>	<u>2.121</u>	<u>1.118</u>	-	-	-
16	tempo (h)	54.883	58.383	65.833	-	-	-	-
	olho dir.	11.619	7.746	2.828	-	-	-	-
	olho esq.	8.660	4.243	1.414	-	-	-	-
	média	10.140	5.994	2.121	-	-	-	-
17	tempo (h)	54.949	58.382	65.899	70.532	82.449	-	-
	olho dir.	14.629	11.705	7.746	5.385	2.828	-	-
	olho esq.	15.427	11.662	8.185	4.359	2.450	-	-
	média	15.028	11.683	7.966	4.872	2.639	-	-
18	tempo (h)	54.915	58.365	65.898	70.531	82.431	-	-
	olho dir.	18.166	14.177	8.124	5.099	-	-	-
	olho esq.	17.349	13.342	9.849	7.746	5.196	-	-
	média	17.758	13.760	8.986	6.422	<u>2.598</u>	-	-
19	tempo (h)	55.082	58.632	65.849	-	-	-	-
	olho dir.	12.083	5.745	-	-	-	-	-
	olho esq.	11.958	7.141	2.646	-	-	-	-
	média	12.021	6.443	<u>1.323</u>	-	-	-	-
20	tempo (h)	55.083	58.650	-	-	-	-	-
	olho dir.	8.485	5.196	-	-	-	-	-
	olho esq.	3.873	1.732	-	-	-	-	-
	média	6.179	3.464	-	-	-	-	-
21	tempo (h)	-	-	-	-	-	-	-
	olho dir.	-	-	-	-	-	-	-
	olho esq.	-	-	-	-	-	-	-
	média	-	-	-	-	-	-	-
Média	tempo (h)	-	-	-	-	-	-	-
	direito	-	-	-	-	-	-	-
	esquerdo	-	-	-	-	-	-	-
	média	-	-	-	-	-	-	-

Tabela 1.6: $\sqrt{Area}$ da úlcera: Tratamento 4 (continua) - Ex. 3.

Unidade amostral	variável	Observação						
		1	2	3	4	5	6	7
22	tempo (h)	0.000	11.783	19.183	28.533	34.866	41.249	47.299
	olho dir.	51.856	44.497	39.837	35.355	32.818	30.083	24.759
	olho esq.	47.739	41.280	38.432	35.819	31.273	25.100	20.075
	média	49.797	42.888	39.134	35.587	32.045	27.591	22.417
23	tempo (h)	0.000	11.750	19.150	28.500	34.833	41.400	47.233
	olho dir.	47.864	41.425	32.016	27.019	21.142	15.906	7.937
	olho esq.	47.948	39.281	32.031	26.000	14.765	9.592	4.472
	média	47.906	40.353	32.023	26.509	17.954	12.749	6.205
24	tempo (h)	0.000	11.700	19.033	28.400	34.733	41.250	47.117
	olho dir.	45.255	39.711	35.285	28.107	23.643	18.788	15.362
	olho esq.	46.508	41.304	34.828	29.631	22.935	19.545	17.234
	média	45.881	40.508	35.056	28.869	23.289	19.167	16.298
25	tempo (h)	0.000	11.167	18.967	28.317	34.650	41.167	47.100
	olho dir.	52.801	41.328	30.512	25.554	17.917	12.728	7.681
	olho esq.	48.156	40.804	29.648	24.495	18.412	13.784	8.944
	média	50.479	41.066	30.080	25.024	18.164	13.256	8.313
26	tempo (h)	0.000	11.267	18.900	28.250	34.733	41.100	47.000
	olho dir.	51.430	41.881	35.693	31.401	27.295	20.298	11.314
	olho esq.	46.883	39.484	34.612	28.966	22.605	17.550	11.705
	média	49.156	40.682	35.153	30.183	24.950	18.924	11.509
27	tempo (h)	0.000	11.600	18.883	28.233	34.716	41.133	48.133
	olho dir.	49.346	40.200	32.573	25.140	14.318	8.944	5.385
	olho esq.	50.030	45.354	37.802	33.616	26.096	16.124	8.426
	média	49.688	42.777	35.188	29.377	20.207	12.534	6.906
28	tempo (h)	0.000	11.583	18.883	28.300	34.717	41.150	47.033
	olho dir.	50.636	43.347	37.148	29.967	20.976	17.088	12.083
	olho esq.	47.466	36.756	29.462	25.612	20.445	13.820	9.219
	média	49.051	40.052	33.305	27.790	20.711	15.454	10.651
Média	tempo (h)	0.000	11.550	19.000	28.362	34.750	41.207	47.274
	olho dir.	49.884	41.770	34.723	28.934	22.587	17.691	12.075
	olho esq.	47.818	40.609	33.831	29.163	22.361	16.502	11.439
	média	48.851	41.189	34.277	29.049	22.474	17.096	11.757

Tabela 1.6: $\sqrt{Area}$ da úlcera: Tratam. 4 (continuação) - Ex. 3.

Unidade amostral	variável	Observação						
		8	9	10	11	12	13	14
22	tempo (h)	55.016	58.533	65.700	70.250	82.167	92.067	-
	olho dir.	22.023	17.889	13.000	8.307	6.083	2.646	-
	olho esq.	17.292	13.454	10.247	6.325	3.000	-	-
	média	19.657	15.671	11.623	7.316	4.541	<u>1.323</u>	-
23	tempo (h)	-	-	-	-	-	-	-
	olho dir.	-	-	-	-	-	-	-
	olho esq.	-	-	-	-	-	-	-
	média	-	-	-	-	-	-	-
24	tempo (h)	54.800	58.300	65.783	70.083	-	-	-
	olho dir.	11.747	8.485	5.099	2.000	-	-	-
	olho esq.	14.560	11.402	6.083	-	-	-	-
	média	13.154	9.943	5.591	<u>1.000</u>	-	-	-
25	tempo (h)	54.817	58.234	65.717	-	-	-	-
	olho dir.	5.568	3.742	2.236	-	-	-	-
	olho esq.	6.164	3.162	1.414	-	-	-	-
	média	5.866	3.452	1.825	-	-	-	-
26	tempo (h)	54.700	58.133	-	-	-	-	-
	olho dir.	5.831	-	-	-	-	-	-
	olho esq.	6.708	3.000	-	-	-	-	-
	média	6.270	<u>1.500</u>	-	-	-	-	-
27	tempo (h)	55.133	-	-	-	-	-	-
	olho dir.	2.000	-	-	-	-	-	-
	olho esq.	3.742	-	-	-	-	-	-
	média	2.871	-	-	-	-	-	-
28	tempo (h)	54.633	58.050	-	-	-	-	-
	olho dir.	6.481	2.450	-	-	-	-	-
	olho esq.	4.000	-	-	-	-	-	-
	média	5.240	<u>1.225</u>	-	-	-	-	-
Média	tempo (h)	-	-	-	-	-	-	-
	olho dir.	-	-	-	-	-	-	-
	olho esq.	-	-	-	-	-	-	-
	média	-	-	-	-	-	-	-

1.2 Modelos para dados longitudinais

Idealmente, os modelos estatísticos para dados longitudinais devem i) permitir a incorporação de uma estrutura de dependência entre as medidas realizadas numa mesma unidade amostral; ii) não se restringir a dados balanceados ou regulares no tempo, e iii) permitir a inclusão de covariáveis na análise. Os modelos apresentados nesta Seção obedecem, na medida do possível, à ordem cronológica em que se tornaram conhecidos.

Em geral, estudos longitudinais com dados balanceados com relação ao tempo permitem a utilização do modelo de análise de variância multivariada com fatores relacionados aos grupos de onde as unidades foram selecionadas e a distribuição de respostas multivariadas parametrizada através do seguinte modelo de médias de caselas:

$$\mathbf{Y} = \mathbf{QM} + \mathbf{E} \quad (1.2.1)$$

onde $\mathbf{Y}$ ($N \times T$) é uma matriz que contém em cada linha as T respostas de uma das $N = \sum_{i=1}^g N_i$ unidades amostrais, sendo N_i o tamanho da amostra do grupo $i, i = 1, \dots, g$, $\mathbf{Q} = \text{diag}\{\mathbf{1}_{N_1}, \dots, \mathbf{1}_{N_g}\}$, com $\mathbf{1}_{N_i}$ indicando um vetor com N_i elementos iguais a 1, é a matriz de especificação (conhecida), $\mathbf{M}$ ($g \times T$) é uma matriz de parâmetros, cujas linhas correspondem aos perfis de respostas médias de cada grupo e as linhas de $\mathbf{E}$ ($N \times T$) correspondentes aos erros aleatórios associados às unidades amostrais são independentes e idênticamente distribuídas (*iid*) com distribuição $N_T(\mathbf{0}, \mathbf{V})$, com matriz de dispersão $\mathbf{V}$ não estruturada. Quando este modelo é utilizado especificamente para dados longitudinais, a análise correspondente é conhecida como *análise (multivariada) de perfis*. Embora a parametrização de médias de caselas seja mais freqüente, outras alternativas podem ser consideradas.

Para descrever os perfis de respostas médias de *curvas de crescimento* (esta denominação é utilizada mesmo quando a dimensão ao longo da qual são feitas as observações não é o tempo) de animais tem-se recorrido a modelos polinomiais (Wishart, 1938). Para Potthoff & Roy (1964), isto equivale a estruturar a matriz $\mathbf{M}$ na forma

$$\mathbf{M} = \mathbf{BX}^t \quad (1.2.2)$$

onde $\mathbf{X}$ ($T \times r$), com $r < T$, é uma matriz de especificação comum a todos os grupos, que no caso de curvas de crescimento polinomiais, pode conter

nas colunas as diferentes potências (0 a $r - 1$) dos T tempos em que são realizadas as medidas ou outras variáveis explicativas, de tal forma que o perfil de respostas médias do grupo i seja representado pela função linear $\mathbf{X}_i \boldsymbol{\beta}_i$, onde $\boldsymbol{\beta}_i^t$ é a i -ésima linha de $\mathbf{B}$ ($g \times r$), matriz de parâmetros associados aos modelos polinomiais.

Quando os dados não são balanceados com relação ao tempo, a cada unidade amostral j , devem corresponder uma matriz de especificação e uma matriz de dispersão próprias, $\mathbf{X}_j$ e $\mathbf{V}_j$ respectivamente. Mais especificamente, sendo $\mathbf{y}_j^t$ a j -ésima linha de $\mathbf{Y}$, $\mathbf{X}_j = \mathbf{q}_j^t \otimes \mathbf{X}$ onde $\mathbf{q}_j^t$ é a j -ésima linha de $\mathbf{Q}$ e $\otimes$ é o símbolo da operação matricial (produto de Kronecker) que multiplica cada elemento da matriz à esquerda do símbolo pela matriz à direita, e $\boldsymbol{\beta} = \text{vec}(\mathbf{B})$ temos:

$$\mathbf{y}_j \sim N(\mathbf{X}_j \boldsymbol{\beta}, \mathbf{V}_j) \quad (1.2.3)$$

em que as matrizes de dispersão das observações $\mathbf{V}_j, j = 1, \dots, N$, não são estruturadas. Dificuldades na estimação sob este modelo, devido ao grande número de parâmetros desconhecidos, apontam para a necessidade de se considerar algum tipo de estrutura comum às matrizes $\mathbf{V}_j, j = 1, \dots, N$, restringindo-se o número de parâmetros necessários para modelá-las.

O *modelo (misto) de efeitos aleatórios* que será apresentado adiante foi introduzido no contexto de dados longitudinais por Rao (1965). Esta classe de modelos, conhecida também por *modelos hierárquicos* ou *modelos de dois estágios*, pode ser apresentada da seguinte forma: no primeiro estágio, define-se uma curva para cada unidade amostral; no segundo, ao se admitir que os parâmetros das curvas do primeiro estágio são aleatórios, obtém-se a distribuição marginal das observações individuais, cujo valor esperado define a curva média; sob esse enfoque, a matriz de dispersão fica estruturada com menos parâmetros que no modelo de análise de variância multivariada (1.2.1). Nas propostas originais, todos os parâmetros que definem as curvas individuais são considerados aleatórios com distribuição normal, independentes (Elston & Grizzle, 1962) ou correlacionados (Rao, 1965; Fearn, 1975). No modelo mais geral sugerido por Fearn (1977), os valores esperados de alguns parâmetros aleatórios podem se anular, permitindo, por exemplo, que a curva média seja um polinômio de grau menor que o das curvas individuais. A forma mais geral deste modelo é apresentada por Laird & Ware (1982), sendo o primeiro estágio dado por

$$y_j | \gamma_j \sim N(\mathbf{X}_j \beta + \mathbf{Z}_j \gamma_j, \sigma^2 \mathbf{R}_j), j = 1, \dots, N, \quad (1.2.4)$$

onde os vetores de parâmetros (*efeitos*) aleatórios, $\gamma_j, j = 1, \dots, N$ têm média zero e o vetor de parâmetros definidores da curva média, β , corresponde aos *efeitos fixos* e as matrizes de especificação, ($\mathbf{Z}_j$ para o efeito aleatório e $\mathbf{X}_j$ para o efeito fixo, $j = 1, \dots, N$) podem ser diferentes e variar entre unidades amostrais, estendendo o modelo para o caso de dados não balanceados em relação ao tempo e $\sigma^2 \mathbf{R}_j$, conhecida como *matriz de dispersão condicional* (associada ao *erro condicional* $\epsilon_j = y_j - \mathbf{X}_j \beta - \mathbf{Z}_j \gamma_j$), é parametrizada por σ^2 e $\mathbf{R}_j$ tem uma estrutura conhecida. Se $\mathbf{R}_j = \mathbf{I}_j, j = 1, \dots, N$, o modelo é conhecido como *modelo com independência condicional* (IC) ou *modelo com erro de medida*. Diversas estruturas de dispersão do erro condicional podem ser consideradas, com a matriz $\mathbf{R}_j$ estruturada convenientemente. O segundo estágio consiste em especificar a distribuição dos efeitos aleatórios, usualmente na forma

$$\gamma_j \sim N(\mathbf{0}, \mathbf{\Gamma}), j = 1, \dots, N, \quad (1.2.5)$$

independentemente dos erros condicionais. Conseqüentemente, a distribuição marginal de y_j é dada por

$$y_j \sim N(\mathbf{X}_j \beta, \mathbf{Z}_j \mathbf{\Gamma} \mathbf{Z}_j^t + \sigma^2 \mathbf{R}_j), j = 1, \dots, N. \quad (1.2.6)$$

O *erro marginal* é definido como $\epsilon_j = y_j - \mathbf{X}_j \beta$ e a *matriz de dispersão marginal* como $\mathbf{Z}_j \mathbf{\Gamma} \mathbf{Z}_j^t + \sigma^2 \mathbf{R}_j$.

Os modelos AR (autoregressivo) e ARMA (autoregressivo-média móvel) (Box & Jenkins, 1976) para os erros são, pela parametrização parcimoniosa e diversidade na estrutura de dispersão, apontados como uma alternativa ao modelo de efeitos aleatórios (Rao, 1967; LaVange & Helms, 1983; Ware, 1985; Schluchter, 1988, e Rochon & Helms, 1989). Através de simulação, Jones (1990) conclui que em amostras pequenas, torna-se impossível distinguir os efeitos aleatórios de erros com estrutura AR(1). Aparentemente pelo mesmo motivo, Anderson & Jones (1991) utilizam um modelo com Movimento Browniano (MB) integrado, mantendo o intercepto aleatório. Outros autores preferem utilizar esses modelos para explicar apenas a covariância condicional: Azzalini (1987) utiliza um modelo ARMA(1,1) para os erros condicionais ao analisar os dados do Exemplo 2; Chi & Reinsel (1989) utilizam um modelo AR(1) para os erros condicionais ao analisar os dados de

Zerbe (1979); ambos os trabalhos sugerem um melhor ajuste relativamente ao modelo com independência condicional. Um modelo com MB e erro de medida é utilizado por Lundbye-Christensen (1991) em curva de crescimento polinomial com o intercepto aleatório.

Para modelar os erros condicionais (vistos para cada unidade amostral como uma realização independente de um processo estocástico $\{\varepsilon_t\}$), através de um processo AR(1) com intervalos de amostragem arbitrários no tempo, Jones (1993) lança mão do processo de (velocidade) de Ornstein-Ühlenbeck (OU) (ver Durrett, 1984, para uma definição do processo OU), denominando-o *processo AR(1) contínuo* e denotando-o CAR(1). Esse processo é definido pela equação diferencial

$$d\varepsilon_t + \alpha\varepsilon_t dt = \omega dB_t \quad (1.2.7)$$

onde $\{\omega B_t\}$ é um MB com parâmetro de dispersão ω e α é uma constante. Quando $\alpha = 0$, $\{\varepsilon_t\}$ é um MB. Integrando (1.2.7), obtém-se a forma autor-regressiva

$$\varepsilon_t = \exp(-\alpha\Delta t)\varepsilon_{t-\Delta t} + u_t \quad (1.2.8)$$

onde $\{u_t\}$ é um ruído branco com variância

$$\frac{\omega^2}{2\alpha}(1 - \exp(-2\alpha\Delta t)).$$

(ver os cálculos no Apêndice A). Quando $\alpha \rightarrow \infty$ com $\omega^2/(2\alpha)$ constante, os erros condicionais se tornam independentes, podendo-se falar num *processo* $\{\varepsilon_t\}$ *de erros de medida* (PEM). Assim, tanto o MB quanto o PEM são limites do processo CAR(1) (Jones, 1993). Mandel (1957) considera estes dois processos para modelar observações repetidas univariadas, denominando o primeiro de *processo de erros intervalares* e ressaltando a importância da escolha adequada do modelo para se avaliar corretamente o erro de estimação. Box (1950) nota que a primeira diferença simplifica a estrutura de covariâncias, atribuindo este fato à presença de *erros acumulativos*. Diggle (1988) e Murray & Helms (1990) generalizam o parâmetro de autoregressão para $\exp[-\alpha(\Delta t)^c]$, o primeiro com $c = 1$ ou 2 , e os outros com $c \geq 0$, para modelar processos mais (ou menos) suaves que o CAR(1) com $c > 1$ (ou $0 \leq c < 1$, incluindo para $c = 0$, o *modelo de simetria composta*, ou seja, com padrão de covariâncias uniforme).

Outro enfoque consiste em considerar as próprias observações $\{y_{jt}, t \geq 0\}, j = 1, \dots, N$, como N realizações de um processo estocástico. Gabriel (1962) introduziu o *modelo de antedependência* para o conjunto de observações completas $\mathbf{Y} = (\mathbf{Y}_1, \dots, \mathbf{Y}_T)$. Antedependência de ordem s significa que para $t = 1, \dots, T$, $\mathbf{Y}_t | (\mathbf{Y}_{t-1}, \dots, \mathbf{Y}_{t-s})$ independe de $\mathbf{Y}_{t-k}$, para todo $k > s$; conseqüentemente, antedependência de ordem 0 implica completa independência entre as colunas de $\mathbf{Y}$ e antedependência de ordem $(T - 1)$ implica completa arbitrariedade no padrão de dependência. O modelo de antedependência é o equivalente multivariado de modelos autoregressivos, em que se torna possível relaxar a hipótese de *estacionariedade* (variâncias e correlações independentes do tempo), devido à presença de várias realizações do mesmo processo. Um exemplo de aplicação deste modelo é apresentado por Kenward (1987) no contexto de *análise de perfis*. Modelos autoregressivos (estacionários) nas observações são conhecidos por *modelos condicionais*, tendo sido utilizados com covariáveis (Rosner & Muñoz, 1988) e com efeitos aleatórios (Rosner & Muñoz, 1992).

Modelos descritos por equações diferenciais podem esclarecer o mecanismo de variação temporal das observações, sendo denominados por Izenman & Williams (1989) como *modelos mecânicos*. As soluções dessas equações diferenciais podem ser autoregressivas nas observações y_t e (ou) não lineares nos parâmetros. Vantagens destes modelos são a interpretabilidade das curvas e de seus parâmetros, em geral, em número reduzido. Ao modelar a velocidade dy_t/dt em função de y_t , ou no caso de tempos discretos, y_t em função de y_{t-s} , $s > 0$, introduz-se um mecanismo de realimentação negativa, fazendo com que o processo convirja para um estado de equilíbrio, ao passo que os modelos polinomiais são ilimitados (Dwyer & Feinleib, 1992). Um dos exemplos mais antigos, descrito por von Bertalanffy (1938), é

$$\frac{dy_t}{dt} = \alpha(\beta - y_t) \quad (1.2.9)$$

onde β é o valor para o qual y_t converge. Em forma de equação diferencial estocástica, trata-se do modelo de OU (1.2.7) em $\beta - y_t$, cuja solução é o modelo autoregressivo

$$y_t = \beta(1 - \exp(-\alpha\Delta t)) + \exp(-\alpha\Delta t)y_{t-\Delta t} + u_t \quad (1.2.10)$$

onde $\{u_t\}$ é um processo de erros independentes. O modelo *incondicional* correspondente é obtido substituindo-se Δt por t em (1.2.10), sendo o modelo

de von Bertalanffy mais conhecido nesta forma não linear. Mas neste caso, o erro acumulado é dado por $\omega \int_{u=0}^t \exp[-\alpha(t-u)]dB_u$, cuja estrutura de covariâncias é a do processo $\{y_t\}$:

$$Cov(y_s, y_t) = \frac{\omega^2}{2\alpha} [\exp(-\alpha|t-s|) - \exp(-\alpha(t+s))]. \quad (1.2.11)$$

(ver os cálculos no Apêndice A).

O modelo condicional (1.2.10) tem a vantagem de apresentar uma estrutura de dispersão dos erros simples, e quando Δt é constante pode ser linearizado nos parâmetros. Quando $\{\beta_t\}$ representa a velocidade de um processo autoregressivo definido por (1.2.10), sua integração resulta num processo $\{y_t\}$ de OU integrado ou *de posição* (Taylor & Cumberland, 1991). Quando a integração é feita entre 0 e t , com as condições iniciais y_0 e $\beta_0 = 0$, obtem-se a equação

$$y_t = y_0 - \beta t + \beta[1 - \exp(-\alpha t)]/\alpha + \kappa_t \quad (1.2.12)$$

onde

$$\kappa_t = \omega \int_{\tau=0}^t \int_{u=0}^{\tau} \exp[-\alpha(\tau-u)]dB_u d\tau \quad (1.2.13)$$

tem a dispersão dada por

$$Cov(\kappa_s, \kappa_u) = \sigma^2 [\min(t_s, t_u) - (2\lambda^{t_u} + 2\lambda^{t_s} - 2 - \lambda^{|t_u-t_s|} - \lambda^{t_u+t_s})/(2 \log \lambda)]. \quad (1.2.14)$$

onde $\lambda = \exp(-\alpha)$ e $\sigma^2 = \omega^2/\alpha^2$ (ver os cálculos no Apêndice A). Quando $\lambda = 0$, $\{\kappa_t\}$ é um MB e o valor esperado de y_t se torna linear. Gamerman & Migon (1993) apresentam um modelo em que o parâmetro β é perturbado ao longo do tempo, de acordo com um MB, denominando-o *modelo hierárquico dinâmico*.

Sandland & McGilchrist (1979) propõem modelar a taxa de crescimento através de uma equação diferencial estocástica. Na prática, eles propõem identificar um modelo para as taxas de crescimento $\Delta \log y_t / \Delta t$ observadas entre os instantes sucessivos de avaliação por meio de um polinômio, uma função exponencial ou uma mistura dos dois, em t , com erros independentes ou na classe ARMA.

Entre as desvantagens do modelo mecanístico podem-se citar a inexistência de uma família paramétrica de curvas como a polinomial, o que frequentemente causa falta de ajuste, e problemas de inferência em pequenas amostras. Por outro lado, os modelos polinomiais apresentam geralmente um bom ajuste, mas às custas de muitos parâmetros, que em muitos casos carecem de interpretação clara. Izenman & Williams (1989) propõem um modelo exploratório com base em análise fatorial, cuja solução pode fornecer fatores com interpretações significativas e levar a um modelo linear parcimonioso, denominando-o *modelo linear espectral*. Sua aplicação, contudo, fica limitada a dados completos e a modelos lineares com erro de medida.

Modelos não lineares de efeitos aleatórios são apresentados por Lindstrom & Bates (1990), Palmer et al. (1991), e Vonesh & Carter (1992).

Modelos para respostas não gaussianas são considerados por diversos autores. Um modelo logístico para respostas binárias independentes é considerado por Stiratelli et al. (1984), e para respostas binárias correlacionadas, por Fitzmaurice & Laird (1993). Liang & Zeger (1986) propõem uma extensão de modelos lineares generalizados (GLM) para dados longitudinais não gaussianos. Uma extensão de GLM que inclui efeitos aleatórios é conhecida por *modelos lineares generalizados de efeitos mistos* (GLMM) (Breslow & Clayton, 1993). Uma extensão do modelo adotado por Lundbye-Christensen (1991) é apresentada por Jørgensen et al. (1991), com base em *modelos de dispersão exponencial* de Jørgensen (1987), em que o *processo de Levy* substitui o MB.

1.3 Inferência estatística

Esta Seção será dedicada aos dois modelos lineares que têm merecido mais atenção nos estudos sobre dados longitudinais: o modelo de Potthoff & Roy (1964), definido através de (1.2.1) e (1.2.2) e o modelo de efeitos aleatórios, definido através de (1.2.4) e (1.2.5). Como veremos a seguir, a inferência sobre os parâmetros do primeiro modelo é inteiramente baseada em análises multivariadas de variância e de covariância. O segundo modelo, sendo o de maior interesse neste trabalho, merecerá maior atenção.

Nas propostas originais de análise de curvas de crescimento de animais, têm sido consideradas *medidas resumo* obtidas através de uma transformação

linear dos dados originais $\mathbf{Y}$, na forma

$$\mathbf{YH}_1, \quad (1.3.1)$$

onde $\mathbf{H}_1$ é uma matriz de constantes convenientes conhecidas. Wishart (1938) propõe ajustar uma curva polinomial em função do tempo aos dados de cada animal, utilizando uma matriz de polinômios ortogonais

$$\mathbf{H} (T \times T) = [\psi_0, \psi_1, \dots, \psi_{T-1}], \quad (1.3.2)$$

onde $\psi_k, k = 0, \dots, T-1$, é um vetor contendo em cada linha o valor de um polinômio ortogonal de grau k no instante de tempo correspondente, com

$$\mathbf{H}^t \mathbf{H} = \mathbf{I}. \quad (1.3.3)$$

Os coeficientes destes polinômios ortogonais são estimados por $\mathbf{YH}$, cada linha correspondendo aos T coeficientes associados à unidade amostral correspondente. Em geral, um polinômio de grau $r-1 < T-1$, é capaz de descrever a curva de crescimento de um animal, o que induz uma partição dos coeficientes estimados dada por

$$\mathbf{YH} = [\mathbf{YH}_1 \ \mathbf{YH}_2], \quad (1.3.4)$$

onde $\mathbf{YH}_1$ contém os coeficientes de polinômios de graus menores, significativamente diferentes de zero, e $\mathbf{YH}_2$ contém os coeficientes de graus maiores, que não diferem significativamente de zero. Para finalidade de comparação entre grupos, Wishart (1938) trabalha apenas com $\mathbf{YH}_1$, recaindo num modelo de análise de variância (1.2.1) para estas medidas resumo. Potthoff & Roy (1964) consideram (1.3.1) com

$$\mathbf{H}_1 = \mathbf{G}^{-1} \mathbf{X}^t (\mathbf{X}^t \mathbf{G}^{-1} \mathbf{X})^{-1} \quad (1.3.5)$$

onde $\mathbf{G}$ é uma matriz simétrica positiva definida. O valor esperado de $\mathbf{YH}_1$ é $\mathbf{QB}$, onde $\mathbf{B}$ está definida em (1.2.2), independentemente da escolha da matriz $\mathbf{G}$. Um estimador não viciado de $\mathbf{B}$ é dado por

$$\hat{\mathbf{B}} = (\mathbf{Q}^t \mathbf{Q})^{-1} \mathbf{Q}^t \mathbf{YH}_1. \quad (1.3.6)$$

A matriz $\mathbf{G}$ pode representar um conhecimento a priori sobre $\mathbf{V}$; na ausência dessa informação, Potthoff & Roy (1964) sugerem $\mathbf{G} = \mathbf{I}$, obtendo um estimador não ponderado.

Para analisar transformações lineares dos dados originais (1.3.1), Rao (1965) sugere a utilização da análise de covariância para obtenção de estimadores mais eficientes do que (1.3.6). No caso de $\mathbf{YH}_1$ ser uma matriz de coeficientes de polinômios ortogonais, as possíveis covariáveis são os coeficientes de polinômios ortogonais de graus maiores, $\mathbf{YH}_2$, com valor esperado nulo. Deste fato, e da distribuição conjunta normal de $(\mathbf{YH}_1, \mathbf{YH}_2)$, resulta que a distribuição de $\mathbf{YH}_1$, condicionada a $\mathbf{YH}_2$, é normal, com média

$$\mathbf{E}(\mathbf{YH}_1 | \mathbf{YH}_2 = \mathbf{Z}) = \mathbf{QB} + \mathbf{ZC} \quad (1.3.7)$$

onde $\mathbf{Z}$ é o valor observado de $\mathbf{YH}_2$, dado pelas estimativas dos coeficientes de polinômios ortogonais de graus maiores, e os parâmetros $\mathbf{C}$ dependem da matriz de dispersão de $(\mathbf{YH}_1, \mathbf{YH}_2)$, desconhecida. Quando $\mathbf{H}_1$ é dada por (1.3.5), com $\mathbf{G} = \mathbf{I}$, as possíveis covariáveis são aquelas que podem gerar o *espaço de erros* (colunas de $\mathbf{Y}^t[\mathbf{I} - \mathbf{X}(\mathbf{X}^t\mathbf{X})^{-1}\mathbf{X}^t]$, por exemplo) e portanto também têm valor esperado nulo. Esta metodologia leva ao estimador ponderado de Rao (1959) dado por (1.3.5)-(1.3.6) com

$$\mathbf{G} = \mathbf{Y}^t[\mathbf{I} - \mathbf{Q}(\mathbf{Q}^t\mathbf{Q})^{-1}\mathbf{Q}^t]\mathbf{Y} \quad (1.3.8)$$

que é também obtido por Khatri (1966) como sendo o estimador de máxima verossimilhança do modelo de Pottoff & Roy (1964) dado por (1.2.1) e (1.2.2) (ver Grizzle & Allen, 1969). Como na prática, nem todas as colunas de $\mathbf{YH}_2$ estão correlacionadas com $\mathbf{YH}_1$, Rao (1965) sugere que se selecionem algumas, com base na correlação amostral, o que pode aumentar a eficiência do estimador.

No modelo (1.2.3) com as matrizes $\mathbf{V}_j, j = 1, \dots, N$ conhecidas, o estimador de mínimos quadrados generalizados de Aitken (1935) dado por

$$\hat{\beta} = \left(\sum_{j=1}^N \mathbf{X}_j^t \mathbf{V}_j^{-1} \mathbf{X}_j \right)^{-1} \sum_{j=1}^N \mathbf{X}_j^t \mathbf{V}_j^{-1} \mathbf{y}_j \quad (1.3.9)$$

corresponde ao estimador de máxima verossimilhança (MV) e ao estimador não viciado de variância mínima (Harville, 1977) com

$$\text{Var}(\hat{\beta}) = \left(\sum_{j=1}^N \mathbf{X}_j^t \mathbf{V}_j^{-1} \mathbf{X}_j \right)^{-1}. \quad (1.3.10)$$

A consistência de $\hat{\beta}$ só depende da correta especificação do valor esperado das observações .

Quando as matrizes $V_j, j = 1, \dots, N$, são desconhecidas, o procedimento usual é substituí-las por estimativas nas equações acima. O estimador em dois estágios de Aitken é obtido através da utilização de estimativas das matrizes $V_j, j = 1, \dots, N$, calculadas pelo método dos momentos a partir dos resíduos de regressão ordinária. Em muitas situações , garante-se a eficiência assintótica deste estimador (Freedman & Peters, 1984). Se os novos resíduos forem utilizados para renovar as estimativas de $V_j, j = 1, \dots, N$ e conseqüentemente de β , e se este ciclo for repetido até convergência das estimativas, elas correspondem às de máxima verossimilhança. Ware & De Gruttola (1985), ao utilizarem o modelo (1.2.3) no contexto de dados longitudinais notaram menor precisão no estimador de máxima verossimilhança de β que no de dois estágios de Aitken, provavelmente devido ao grande número de parâmetros envolvidos na matriz de dispersão não estruturada.

Em muitos casos, a estrutura das matrizes individuais de dispersão pode ser especificada em função de alguns parâmetros desconhecidos, θ , como no caso do modelo de efeitos aleatórios definido através de (1.2.6),

$$V_j = V_j(\theta) = Z_j \Gamma Z_j^t + \sigma^2 R_j, \quad (1.3.11)$$

$j = 1, \dots, N$ onde θ é um vetor contendo σ^2 , os elementos distintos de Γ e os parâmetros desconhecidos de R_j .

O logaritmo da função de verossimilhança do modelo de efeitos aleatórios é

$$l(\beta, \theta; y) = -[n \log(2\pi) + \log |V| + \epsilon^t V^{-1} \epsilon]/2 \quad (1.3.12)$$

onde $y^t = (y_1^t, \dots, y_N^t)$, $n = \sum_{j=1}^N n_j$, onde n_j representa a dimensão de y_j , $V = \sigma^2 R + ZWZ^t$, sendo $R = \text{diag}(R_1, \dots, R_N)$, $Z = \text{diag}(Z_1, \dots, Z_N)$ e $W = \text{diag}(\Gamma, \dots, \Gamma)$, e $\epsilon = y - X\beta$ com $X^t = (X_1^t, \dots, X_N^t)$.

A estimação por MV do modelo misto de análise da variância foi desenvolvida por Hartley & Rao (1967), que propuseram solucionar numericamente o problema de maximização através do método de "steepest ascent". Harville (1977) considera a estimação por máxima verossimilhança restrita (MVR) que veremos adiante e descreve os algoritmos de otimização que poderiam ser úteis para maximizar a verossimilhança. Além do método proposto anteriormente, ele cita os de Newton-Raphson (NR), de escore de Fisher e

de Marquardt (1963). Dempster, Laird & Rubin (1977) desenvolvem um método geral para obtenção de estimadores de MV a partir de observações incompletas, denominando-o *algoritmo EM*. Laird & Ware (1982) utilizam este algoritmo para estimar os parâmetros do modelo de efeitos aleatórios (1.2.6). Jennrich & Schluchter (1986) descrevem os métodos NR, escore de Fisher e uma mistura entre o último e o algoritmo EM para modelar dados incompletos, que são implementados no programa BMDP-5V (Schluchter, 1990). Laird, Lange & Stram (1987) propõem o método de aceleração de Aitken (Gerald, 1970) para melhorar a velocidade de convergência do algoritmo EM. Lindstrom & Bates (1988) comparam este algoritmo com o de NR, concluindo ser este último preferível quando implementado cuidadosamente. Descrevemos, a seguir, os algoritmos de NR e o de escore de Fisher para obtenção dos estimadores de MV e de MVR.

A maximização do logaritmo da verossimilhança em relação aos parâmetros desconhecidos do modelo, $\vartheta = (\beta^t, \theta^t)^t$ ocorre quando se anula a derivada $\eta = \partial l / \partial \vartheta$, particionada em

$$\eta_\beta = \partial l / \partial \beta = \mathbf{X}^t \mathbf{V}^{-1} \epsilon \quad (1.3.13)$$

e $\eta_\theta = \partial l / \partial \theta$, cujos elementos são dados por

$$\partial l / \partial \theta = -[tr(\mathbf{V}^{-1} \partial \mathbf{V} / \partial \theta) - \epsilon^t \mathbf{V}^{-1} (\partial \mathbf{V} / \partial \theta) \mathbf{V}^{-1} \epsilon] / 2 \quad (1.3.14)$$

onde θ é um elemento genérico de θ . Para resolver numericamente estas equações em ϑ , parte-se de um valor inicial ϑ_0 produzindo um novo valor para ϑ utilizando a fórmula

$$\vartheta_k = \vartheta_{k-1} + \mathbf{J}_{k-1}^{-1} \eta_{k-1} \quad (1.3.15)$$

com $k = 1$ onde

$$\mathbf{J} = \begin{pmatrix} \mathbf{J}_{\beta\beta} & \mathbf{J}_{\beta\theta} \\ \mathbf{J}_{\theta\beta} & \mathbf{J}_{\theta\theta} \end{pmatrix}. \quad (1.3.16)$$

No método NR, $\mathbf{J}$ é a Hessiana

$$\begin{pmatrix} \partial^2 l / \partial \beta \partial \beta^t & \partial^2 l / \partial \beta \partial \theta^t \\ \partial^2 l / \partial \theta \partial \beta^t & \partial^2 l / \partial \theta \partial \theta^t \end{pmatrix} \quad (1.3.17)$$

resultando que

$$\mathbf{J}_{\beta\beta} = -\mathbf{X}^t \mathbf{V}^{-1} \mathbf{X}, \quad (1.3.18)$$

$$\mathbf{J}_{\beta\theta} = -\mathbf{X}^t \mathbf{V}^{-1} (\partial \mathbf{V} / \partial \theta) \mathbf{V}^{-1} \epsilon \quad (1.3.19)$$

para um elemento genérico θ de $\boldsymbol{\theta}$ e

$$\begin{aligned} \partial^2 l / \partial \theta_h \partial \theta_m &= \text{tr}[\mathbf{V}^{-1} (\partial \mathbf{V} / \partial \theta_h) \mathbf{V}^{-1} \partial \mathbf{V} / \partial \theta_m] / 2 - \\ &\quad \epsilon^t [\mathbf{V}^{-1} (\partial \mathbf{V} / \partial \theta_h) \mathbf{V}^{-1} \partial \mathbf{V} / \partial \theta_m] \epsilon \end{aligned} \quad (1.3.20)$$

para o elemento (h, m) de $\mathbf{J}_{\theta\theta}$.

No método de escore de Fisher, a Hessiana é substituída por seu valor esperado, resultando em (1.3.18) para $\mathbf{J}_{\beta\beta}$, $\mathbf{J}_{\beta\theta} = \mathbf{0}$ e o elemento (h, m) de $\mathbf{J}_{\theta\theta}$ dado por

$$\mathbf{E}(\partial^2 l / \partial \theta_h \partial \theta_m) = -\text{tr}[\mathbf{V}^{-1} (\partial \mathbf{V} / \partial \theta_h) \mathbf{V}^{-1} \partial \mathbf{V} / \partial \theta_m] / 2 \quad (1.3.21)$$

o que resulta num algoritmo mais simples que o anterior, valendo a equação (1.3.9) para β , em função dos parâmetros desconhecidos $\boldsymbol{\theta}$. Isto significa que basta inicializar com $\boldsymbol{\theta}_0$, calculando-se β_0 por (1.3.9) e $\boldsymbol{\eta}_{\boldsymbol{\theta}(0)}$ por (1.3.14) e obter um novo valor para $\boldsymbol{\theta}$ utilizando a fórmula

$$\boldsymbol{\theta}_k = \boldsymbol{\theta}_{k-1} + \mathbf{J}_{\boldsymbol{\theta}\boldsymbol{\theta}(k-1)}^{-1} \boldsymbol{\eta}_{\boldsymbol{\theta}(k-1)} \quad (1.3.22)$$

com $k = 1$.

Os passos citados são repetidos iterativamente sempre a partir do novo valor de $\boldsymbol{\theta}$ no algoritmo NR ou de $\boldsymbol{\theta}$ no algoritmo de escore de Fisher, até estabilizar os valores dos parâmetros estimados.

O estimador de máxima verossimilhança de β é o estimador de Aitken (1.3.9) com os parâmetros desconhecidos $\boldsymbol{\theta}$ das matrizes individuais de dispersão substituídos pelas estimativas de máxima verossimilhança. Sua variância assintótica também é dada por (1.3.10) (ver Laird & Ware, 1982, por exemplo). Quando $\mathbf{X}_j = \mathbf{Z}_j, j = 1, \dots, N$, e definindo-se $\beta_j = \beta + \gamma_j, j = 1, \dots, N$, esse estimador pode ser obtido pela média ponderada dos estimadores individuais de β_j ,

$$\hat{\beta}_j = (\mathbf{X}_j^t \mathbf{R}_j^{-1} \mathbf{X}_j)^{-1} \mathbf{X}_j^t \mathbf{R}_j^{-1} \mathbf{y}_j \quad (1.3.23)$$

com peso $\mathbf{X}_j^t \mathbf{V}_j^{-1} \mathbf{X}_j, j = 1, \dots, N$ (Crowder & Hand, 1990, p. 93). Hui & Berger (1983) apresentam um procedimento de estimação em dois estágios, onde primeiro se obtém $\hat{\beta}_j$ através de uma análise de regressão adequada ao modelo condicional (1.2.4), mas permitindo que a variância do erro σ_j^2 varie entre as unidades amostrais. Dados β_j e σ_j^2 , sua distribuição condicional é $N(\beta_j, \mathbf{D}_j)$ onde $\mathbf{D}_j = \sigma_j^2 (\mathbf{X}_j^t \mathbf{R}_j^{-1} \mathbf{X}_j)^{-1}, j = 1, \dots, N$ é a variância do erro de estimação de $\hat{\beta}_j$. Sob o modelo (1.2.5), obtém-se a distribuição marginal $N(\beta, \mathbf{\Gamma} + \mathbf{D}_j)$. O estimador de máxima verossimilhança de β (1.3.9) neste caso, resulta na média ponderada de $\hat{\beta}_j$ com pesos proporcionais às inversas das suas dispersões. Da análise de regressão individual que produz $\hat{\beta}_j$, obtém-se também uma estimativa s_j^2 de σ_j^2 . Quando $\sigma_j^2 = \sigma^2$, sua estimativa é dada pela média ponderada de s_j^2 com pesos proporcionais a n_j . Hui & Berger (1983) consideram apenas o caso em que $\hat{\beta}_j$ é um escalar, o que simplifica a estimação da variância $\mathbf{\Gamma}$ do efeito aleatório.

Quando $\mathbf{X}_j = \mathbf{X}$, de posto r , e $\mathbf{R}_j = \mathbf{R}, j = 1, \dots, N$, $\hat{\beta}$ pode ser expresso como sendo a média aritmética dos $\hat{\beta}_j$. No caso do modelo com independência condicional, este estimador independe de θ , e sua variância assintótica é estimada por

$$\hat{\text{Var}}(\hat{\beta}) = \mathbf{S}_\beta / N \quad (1.3.24)$$

onde $\mathbf{S}_\beta$ é a variância amostral dos $\hat{\beta}_j$ s. O estimador de MV de σ^2 é

$$\hat{\sigma}^2 = \sum_{j=1}^N \mathbf{y}_j^t (\mathbf{I} - \mathbf{X}(\mathbf{X}^t \mathbf{X})^{-1} \mathbf{X}^t) \mathbf{y}_j / [N(T - r)] \quad (1.3.25)$$

e o de $\mathbf{\Gamma}$ é

$$\hat{\mathbf{\Gamma}} = \mathbf{S}_\beta - \hat{\sigma}^2 (\mathbf{X}^t \mathbf{X})^{-1} \quad (1.3.26)$$

(Graybill, 1976, p. 465).

Os estimadores de máxima verossimilhança, em geral, subestimam θ , pois não levam em conta a estimação dos outros parâmetros. Com a finalidade de contornar esse problema, Patterson & Thompson (1971) desenvolvem o *estimador de máxima verossimilhança restrita* (MVR) que é obtido através da maximização da verossimilhança de um *contraste de erros* $\mathbf{S}\mathbf{y}$ com valor esperado nulo. Sob o ponto de vista bayesiano, Harville (1974) mostra que isto equivale a ignorar qualquer informação a priori sobre β , definindo a verossimilhança restrita por

$$L^*(\theta; y) = \int L(\beta, \theta; y) \xi(\beta) d\beta \quad (1.3.27)$$

onde L é a função de verossimilhança do modelo de efeitos aleatórios (1.2.6) e $\xi(\beta)$ é uma densidade a priori não informativa. O logaritmo desta verossimilhança é dado por

$$l^*(\theta; y) = C + l(\hat{\beta}, \theta; y) - \log |\mathbf{X}^t \mathbf{V}^{-1} \mathbf{X}|/2 \quad (1.3.28)$$

onde C é uma constante que depende apenas da escolha do contraste de erros e $\hat{\beta}$ é uma função de θ dada por (1.3.9). Definindo

$$\mathbf{P} = \mathbf{V}^{-1} - \mathbf{V}^{-1} \mathbf{X} (\mathbf{X}^t \mathbf{V}^{-1} \mathbf{X})^{-1} \mathbf{X}^t \mathbf{V}^{-1},$$

Harville (1977) mostra que

$$\partial l^* / \partial \theta = -[tr(\mathbf{P} \partial \mathbf{V} / \partial \theta) - \hat{\epsilon}^t \mathbf{V}^{-1} (\partial \mathbf{V} / \partial \theta) \mathbf{V}^{-1} \hat{\epsilon}] / 2, \quad (1.3.29)$$

onde $\hat{\epsilon} = y - \mathbf{X} \hat{\beta}$,

$$\begin{aligned} \partial^2 l^* / \partial \theta_h \partial \theta_m &= tr\{\mathbf{P}[(\partial \mathbf{V} / \partial \theta_h) \mathbf{P} \partial \mathbf{V} / \partial \theta_m - \partial^2 \mathbf{V} / \partial \theta_h \partial \theta_m]\} / 2 - \\ &\quad \hat{\epsilon}^t \mathbf{V}^{-1} [2(\partial \mathbf{V} / \partial \theta_h) \mathbf{P} \partial \mathbf{V} / \partial \theta_m - \partial^2 \mathbf{V} / \partial \theta_h \partial \theta_m] \mathbf{V}^{-1} \hat{\epsilon} / 2 \end{aligned} \quad (1.3.30)$$

e

$$\mathbf{E}(\partial^2 l^* / \partial \theta_h \partial \theta_m) = -tr[\mathbf{P}(\partial \mathbf{V} / \partial \theta_h) \mathbf{P} \partial \mathbf{V} / \partial \theta_m] / 2, \quad (1.3.31)$$

podendo-se utilizar os métodos de NR e o de escore de Fisher para a maximização da verossimilhança restrita. A partir das estimativas de θ , os efeitos fixos são estimados pelas médias das distribuições a posteriori, resultando em (1.3.9), cuja variância é dada por (1.3.10). No modelo com $\mathbf{X}_j = \mathbf{Z}_j = \mathbf{X}$, $j = 1, \dots, N$ com independência condicional, valem as fórmulas (1.3.23) a (1.3.25), com $\mathbf{S}_\beta$ sendo a variância amostral não viciada.

Quando as verossimilhanças (1.3.12) ou (1.3.28) são utilizadas para estimar os parâmetros desconhecidos da distribuição a priori, os estimadores resultantes são denominados *estimadores empíricos bayesianos* (EB).

O efeito aleatório associado ao indivíduo j pode ser estimado através do valor esperado a posteriori

$$\hat{\gamma}_j = E(\gamma_j | y_j, \hat{\beta}, \theta) = \Gamma \mathbf{Z}_j^t \mathbf{V}_j^{-1} (y_j - \mathbf{X}_j \hat{\beta}) \quad (1.3.32)$$

(Harville, 1976) em que os parâmetros θ são supostos conhecidos. Este estimador corresponde à média ponderada entre o estimador de mínimos quadrados generalizados de γ_j do modelo descrito por (1.2.4), com $\beta = \hat{\beta}$ e $\mathbf{R}_j$ suposta conhecida, com peso proporcional à inversa da sua variância, e da estimativa a priori, com valor 0 e peso proporcional à inversa de Γ (Robinson, 1991). Sua variância é dada então por (Laird & Ware, 1982)

$$\text{Var}(\hat{\gamma}_j) = \Gamma \mathbf{Z}_j^t [\mathbf{V}_j^{-1} - \mathbf{V}_j^{-1} \mathbf{X}_j \text{Var}(\hat{\beta}) \mathbf{X}_j^t \mathbf{V}_j^{-1}] \mathbf{Z}_j \Gamma. \quad (1.3.33)$$

No modelo simplificado de Hui & Berger (1983), os estimadores de β_j e de sua variância, correspondentes a (1.3.32) e (1.3.33) são, respectivamente,

$$\tilde{\beta}_j = E(\beta_j | \hat{\beta}_j, \hat{\beta}, \theta) = \frac{\Gamma}{\Gamma + D_j} \hat{\beta}_j + \frac{D_j}{\Gamma + D_j} \hat{\beta} \quad (1.3.34)$$

e

$$\text{Var}(\tilde{\beta}_j) = D_j \Gamma / (\Gamma + D_j). \quad (1.3.35)$$

Inferência sobre o efeito aleatório γ_j , $j = 1, \dots, N$ deve levar em conta também a sua variabilidade; assim, Laird & Ware (1982) propõem que o erro de estimação seja avaliado pela variância de $\hat{\gamma}_j - \gamma_j$,

$$\text{Var}(\hat{\gamma}_j - \gamma_j) = \Gamma - \text{Var}(\hat{\gamma}_j) \quad (1.3.36)$$

em vez de (1.3.33).

A função de verossimilhança (1.3.12) pode ser expressa na forma alternativa

$$L(\beta, \theta; \mathbf{y}) = \prod_{j=1}^N \int f(y_j; \beta, \theta | \gamma_j) \pi(\gamma_j; \theta) d\gamma_j \quad (1.3.37)$$

(Hartley & Rao, 1967) onde f é a densidade das observações do indivíduo j condicionadas ao efeito aleatório γ_j e π é a densidade (a priori) do efeito aleatório. As verossimilhanças (1.3.37) e (1.3.27) são suficientemente gerais para abrigar tanto modelos não lineares quanto distribuições não gaussianas,

que nem sempre podem ser desenvolvidos em forma analítica, quando então pode-se recorrer à integração numérica (Stiratelli et al., 1984; Palmer et al., 1991), à aproximação por linearização (Lindstrom & Bates, 1990; Longford, 1993, Cap. 8) ou a amostragem de Gibbs (Zeger & Karim, 1991).

Dados incompletos ocorrem freqüentemente em estudos longitudinais. Laird (1988) apresenta alguns critérios para classificar os dados incompletos e aponta as precauções necessárias na análise. Quando o mecanismo de não resposta independe da resposta, diz-se que os dados são *incompletos totalmente ao acaso* (*missing completely at random* - MCAR), não sendo necessário nenhum cuidado adicional na análise. Quando a não resposta depende apenas das respostas observadas, ela é dita *ignorável*, pois o seu mecanismo não interfere na estimação, contanto que seja baseado na verossimilhança dos dados observados. Mas deve-se esperar maior sensibilidade à falha na especificação do modelo. Quando a não resposta depende dos valores não observados, ela é dita *não ignorável*, sendo necessário especificar-se um modelo para o mecanismo de não resposta para se obterem inferências válidas com base na verossimilhança.

Quando o modelo é corretamente especificado, a maximização da verossimilhança resulta em estimadores consistentes (Cramér, 1946). Mas a utilização da fórmula da variância assintótica, (1.3.10), com simples substituição dos parâmetros desconhecidos por suas estimativas, resulta na subestimação do seu valor pois não leva em conta o erro de estimação. Kackar & Harville (1984) e Kass & Steffey (1989) propõem fórmulas para corrigir a variância estimada. Kackar & Harville (1984) consideram a inferência sobre uma combinação linear dos efeitos fixos e aleatórios, denotada por τ . Quando τ é estimado pela substituição dos efeitos desconhecidos por seus estimadores (1.3.9) e (1.3.32) ele torna-se uma função de θ e das observações $\mathbf{y} = (\mathbf{y}_1^t, \dots, \mathbf{y}_N^t)^t$, denotada por $\tau(\mathbf{y}; \theta)$. Quando θ é conhecido, $Var[\tau(\mathbf{y}; \theta)]$ é calculada usando as expressões (1.3.10) e (1.3.33). Mas quando θ é substituído pelo estimador de MV, MVR ou, em geral, por um estimador *invariante por translação* (Kackar & Harville, 1981), obtém-se um estimador $\hat{\tau} = \tau(\mathbf{y}; \hat{\theta})$, cuja variância é dada por

$$Var(\hat{\tau}) = Var[\tau(\mathbf{y}; \theta)] + E[\hat{\tau} - \tau(\mathbf{y}; \theta)]^2 \quad (1.3.38)$$

mostrando que $Var[\tau(\mathbf{y}; \theta)]$ é apenas o limite inferior da verdadeira variância, cuja estimativa é dada pela substituição de θ por sua estimativa. Após uma

aproximação da segunda parcela pela utilização do Método Delta para expandir $\hat{\tau}$ em torno de $\tau(\mathbf{y} : \boldsymbol{\theta})$, Kackar & Harville (1984) propõem a variância aproximada

$$\widehat{Var}(\hat{\tau}) = Var[\tau(\mathbf{y}; \boldsymbol{\theta})] + tr[\mathbf{A}(\boldsymbol{\theta})\mathbf{B}(\boldsymbol{\theta})] \quad (1.3.39)$$

onde $\mathbf{A}(\boldsymbol{\theta}) = Var[\partial\tau(\mathbf{y}; \boldsymbol{\theta})/\partial\boldsymbol{\theta}]$ e $\mathbf{B}(\boldsymbol{\theta}) = E[(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})^t]$ é a matriz de dispersão (assintótica) de um estimador $\hat{\boldsymbol{\theta}}$ não viciado (consistente). Kass & Steffey (1989) desenvolvem uma fórmula semelhante, mas com

$$(\partial\tau(\mathbf{y}; \boldsymbol{\theta})/\partial\boldsymbol{\theta}) \cdot (\partial\tau(\mathbf{y}; \boldsymbol{\theta})/\partial\boldsymbol{\theta}^t)$$

no lugar de $\mathbf{A}(\boldsymbol{\theta})$. Eles mostram que os erros de aproximação dos dois termos de (1.3.39) são da ordem de $O(n_j^{-1})$ e $O(N^{-1})$ respectivamente e concluem que quando n_j é pequeno em relação a N , o primeiro termo é dominante e a prática comum de se utilizar apenas o primeiro termo como uma aproximação da variância verdadeira é apropriada; caso contrário, o segundo termo se torna relevante. Métodos de reuso amostral são propostos por alguns autores para se avaliar a variância de $\hat{\beta}$. Freedman & Peters (1984), ao aplicarem o modelo (1.2.3) num estudo econométrico, encontram uma grande discrepância entre a variância assintótica estimada usando o estimador de Aitken em dois estágios e a obtida através de *bootstrap*. Num modelo univariado de efeitos aleatórios, Lee (1991) propõe estimar esta variância através de *jackknife*, e apresenta um estimador robusto em relação à contaminação dos dados por valores aberrantes.

Amostragem de Gibbs é um método aproximado para se estimar a distribuição a posteriori dos parâmetros estimados num modelo hierárquico. Utilizando este método, Gelfand et al. (1990) encontram variâncias dos estimadores dos efeitos fixos consistentemente maiores que as estimativas obtidas através do método EB e recomendam seu uso especialmente para pequenas amostras. Este enfoque é útil para a estimação de modelos GLMM (ver pag. 18), mas seu sucesso depende da correta especificação das densidades que definem o modelo hierárquico e da capacidade de simulação das densidades condicionais resultantes.

A consistência e a eficiência assintótica dos estimadores de máxima verossimilhança dependem da correta especificação da distribuição das observações $\mathbf{y}$. Se a distribuição condicional de $\mathbf{y}|\boldsymbol{\gamma}$ estiver incorreta, ou na equação de regressão ou na distribuição dos erros, então $\hat{\boldsymbol{\gamma}}$ não é necessariamente consistente (Berk, 1966 e 1970). Quando existe falha na especificação da

distribuição do erro marginal, o estimador de máxima verossimilhança pode ainda ser consistente, com distribuição assintótica normal, mas com outra matriz de dispersão, dada por

$$\mathbf{C} = \mathbf{A}^{-1} \mathbf{D} \mathbf{A}^{-1} \quad (1.3.40)$$

onde

$$\mathbf{A} = \mathbf{E}\left[\frac{\partial^2 \log L(\boldsymbol{\beta}, \boldsymbol{\theta}; \mathbf{y})}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}^t}\right] \quad (1.3.41)$$

e

$$\mathbf{D} = \sum_{j=1}^N \frac{\partial \log f(\mathbf{y}_j; \boldsymbol{\theta}, \boldsymbol{\beta})}{\partial \boldsymbol{\beta}} \frac{\partial \log f(\mathbf{y}_j; \boldsymbol{\theta}, \boldsymbol{\beta})}{\partial \boldsymbol{\beta}^t}. \quad (1.3.42)$$

White (1982) e Royall (1986) mostram que $\mathbf{C}$ é robusta com relação à especificação do modelo para a distribuição do erro marginal. Quando o modelo estiver corretamente especificado, $-\mathbf{A} = \mathbf{D}$ e (1.3.40) resulta na variância assintótica usual. A consistência deste estimador só depende da correta especificação da média. Sob o modelo linear (1.2.3), do qual o modelo de efeitos aleatórios é um caso particular, $-\mathbf{A}^{-1}$ é dada por (1.3.10), e $\mathbf{D}$ por

$$\mathbf{D} = \sum_{j=1}^N \mathbf{X}_j^t \mathbf{V}_j^{-1} (\mathbf{y}_j - \mathbf{X}_j \boldsymbol{\beta}) (\mathbf{y}_j - \mathbf{X}_j \boldsymbol{\beta})^t \mathbf{V}_j^{-1} \mathbf{X}_j \quad (1.3.43)$$

(Rutter & Elashoff, 1992). Na prática, os parâmetros desconhecidos são substituídos pelas estimativas correspondentes.

1.4 Métodos para identificação , seleção e diagnóstico de modelos

Uma análise preliminar simples de um determinado conjunto de dados longitudinais pode sugerir um modelo para os efeitos fixos e aleatórios. Gráficos dos valores médios das observações ao longo do tempo podem sugerir a forma da curva média e o comportamento da variabilidade em função do tempo pode revelar a natureza do efeito aleatório: variância amostral constante sugere que apenas o intercepto é aleatório e variância crescente sugere a presença de outros efeitos aleatórios (Jones, 1993). Fearn (1977) ressalta a importância de se verificar a natureza das curvas individuais para se identificarem os

efeitos aleatórios. Rao & Rao (1966) representam num mesmo gráfico todas as observações disponíveis num estudo longitudinal destacando as unidades amostrais ao unir as observações individuais por meio de segmentos de reta. São os denominados *diagramas paralelos de dispersão* (*parallel plots*), que podem tanto sugerir um modelo para a curva média quanto para as individuais, como também detectar os indivíduos com observações (ou associados a curvas) aberrantes. Weiss & Lazaro (1992) sugerem a utilização deste tipo de gráfico também para os resíduos condicionais, caso em que deve refletir apenas a variabilidade dentro das unidades amostrais. Outros efeitos aleatórios não incluídos no modelo podem também ser detectados neste tipo de gráfico. Rutter & Elashoff (1992) sugerem fazer uma análise de regressão ordinária para cada indivíduo e a partir da distribuição amostral dos coeficientes estimados efetuar a seleção de efeitos fixos e aleatórios; a associação entre os parâmetros individuais com outras covariáveis pode indicar possíveis interações. Com a mesma finalidade, Waternaux et al. (1989), estudam a associação entre os efeitos aleatórios estimados numa primeira análise, com outras covariáveis. Grady (1992) propõe analisar a matriz de correlações estimadas num modelo preliminar com matriz de dispersão não estruturada, com a finalidade de identificar uma estrutura AR(1) com ou sem variância crescente. As linhas da matriz de correlações residuais são representadas em diagramas paralelos de dispersão, em função da diferença de tempo entre as medidas. Ao eliminar deste gráfico os elementos abaixo da diagonal principal, as curvas ficam com comprimentos diferentes. Uma estrutura AR(1) é identificada pelo decaimento exponencial das curvas que, se coincidirem, indicam estacionariedade; se tiverem a mesma forma, mas com as curvas ordenadas de acordo com seus comprimentos, ficando aquelas de comprimento maior abaixo daquelas de comprimento menor, indicam variância crescente, e se forem concorrentes na origem indicam variância constante e padrão de autocorrelação não estacionário. O modelo com simetria composta pode ser detectado pela constância das correlações, cujo valor é teoricamente igual à correlação intra-classe.

Testes estatísticos para seleção de modelos incluem os de *Wald* para avaliar a significância dos efeitos fixos e os de *razão de verossimilhança* para avaliar o ajuste de *modelos encaixados*, entendendo-se que um modelo está encaixado em outro quando se tratar de um caso particular deste. O primeiro teste pode estar viciado se calcado em estimativa viciada da variância. O segundo não é apropriado para efeitos fixos quando se usa MVR, pois a

verossimilhança maximizada exclui estes efeitos (Schluchter, 1992). Helms & McCarroll (1991) propõem estatísticas F para testar hipóteses envolvendo os efeitos fixos. Jones (1993) propõe a utilização do *Critério de Informação de Akaike* (AIC) para comparar modelos não encaixados. White (1982) propõe um teste para detectar falha na especificação do modelo, baseado na diferença entre os elementos das matrizes que estimam $\mathbf{A}$ e $\mathbf{D}$ (1.3.41) e (1.3.42) respectivamente.

Waternaux et al. (1989), ao utilizarem o modelo de regressão generalizada (1.2.3), propõem que o segundo passo do estimador de Aitken seja baseado na análise do modelo transformado $\mathbf{L}_j \mathbf{y}_j$, onde $\mathbf{L}_j$ é uma matriz triangular inferior tal que $\mathbf{L}_j^t \mathbf{L}_j = \hat{\mathbf{V}}_j^{-1}$, pois neste caso podem ser utilizadas técnicas de diagnóstico de regressão linear simples. Unidades amostrais aberrantes poderiam ser detectadas calculando-se as *distâncias de Mahalanobis* entre os vetores de resíduos individuais e a origem, que sob a hipótese nula teriam distribuição χ^2 com n_j graus de liberdade. Jones (1993) propõe analisar os resíduos marginais transformados $\mathbf{L}_j \hat{\boldsymbol{\epsilon}}_j$ para diagnóstico do modelo estimado.

Como os efeitos aleatórios são estimados através de transformações lineares dos resíduos marginais, Waternaux et al. (1989) chamam a seus elementos de intercepto, inclinação, etc. dos resíduos. Eles propõem utilizar a distância de Mahalanobis na forma $\hat{\boldsymbol{\gamma}}_j^t \widehat{\mathbf{Var}}(\hat{\boldsymbol{\gamma}}_j - \boldsymbol{\gamma}_j) \hat{\boldsymbol{\gamma}}_j$ para detectar unidades amostrais aberrantes. Um teste sobre a normalidade dos efeitos aleatórios, baseado na comparação entre as funções de distribuição empírica e teórica, é proposto por Lange & Ryan (1989). Para levar em conta a heterocedasticidade dos efeitos aleatórios estimados, quando $n_j, j = 1, \dots, N$ não é constante, eles propõem padronizá-los, resultando numa *função de distribuição empírica ponderada*, cujas propriedades distribucionais são apresentadas.

Medidas de influência para modelos mistos de análise da variância foram propostos por Beckman et al. (1987) e Liski (1991). Rutter & Elashoff (1992) sugerem como método de averiguação da influência de uma observação (ou de uma unidade amostral) avaliar diferenças nos parâmetros estimados e no logaritmo da verossimilhança entre as análises com e sem a sua presença.

1.5 Propostas deste trabalho

Os diagramas paralelos de dispersão nem sempre têm a clareza necessária para a identificação precisa de um modelo de efeitos aleatórios. Por outro

lado, fazer uma análise de regressão para cada indivíduo pode ter um alto custo, especialmente em amostras grandes. Além disso, Rutter & Elashoff (1992) não especificam claramente os critérios adotados para a seleção de efeitos fixos e aleatórios a partir dos parâmetros estimados das curvas individuais. Grady (1992) propõe uma análise gráfica dos elementos distintos da matriz de correlação residual de um modelo com matriz de dispersão não estruturada, mas seu trabalho está direcionado para a identificação de uma estrutura AR(1) dentro dos indivíduos. Neste trabalho, propomos identificar os efeitos aleatórios a partir dos diagramas paralelos de dispersão das linhas da matriz de dispersão amostral dos dados, que é a dispersão residual de um modelo parametrizado por médias de caselas. Para isso nos restringimos a situações com dados balanceados com relação ao tempo. Através de um exemplo no Capítulo 2, mostramos que este gráfico é mais sugestivo que o gráfico dos diagramas paralelos de dispersão dos dados originais. Como, em geral, o número de repetições por indivíduo é menor que o tamanho da amostra, o custo para a análise das linhas da matriz de dispersão amostral dos dados é menor que na análise das curvas individuais, como foi proposto por Rutter & Elashoff (1992).

O modelo clássico de efeitos aleatórios simplifica ao extremo a estrutura de dispersão intra-individual, considerando-a diagonal com variância constante. No outro extremo, modelos estocásticos com estruturas de dispersão mais complexas como ARMA (Rao, 1967; LaVange & Helms, 1983; Ware, 1985, Schluchter, 1988) ou MB integrado (Anderson & Jones, 1991) têm sido utilizados sem os efeitos aleatórios ou apenas com o intercepto aleatório. No meio termo, diversos autores consideram o modelo de efeitos aleatórios com estruturas intra-individuais variadas: ARMA(1,1) (Azzalini, 1987), AR(1) (Chi & Reinsel, 1989), AR(1) generalizado (Diggle, 1988 e Murray & Helms, 1990), AR(1) integrado (Taylor & Cumberland, 1991) e MB (Lundbye-Christensen, 1991), mas sem se preocuparem com o problema de identificação do modelo, que deve considerar concomitantemente as duas componentes da matriz de dispersão (1.3.11) (Rutter & Elashoff, 1992). Consideramos que o conhecimento mais detalhado do comportamento das matrizes de dispersão induzidas por modelos de efeitos aleatórios com diversas estruturas de dispersão intra-individual seja útil para a identificação de modelos através da matriz de dispersão amostral dos dados. No Capítulo 3, estudamos o comportamento de cada linha das matrizes de covariâncias e de correlações induzidas por modelos com intercepto e coeficiente angular

aleatórios e com erros condicionais independentes ou sob MB.

As técnicas consideradas nos Capítulos 2 e 3 são utilizadas no Capítulo 4 para identificar um modelo mais complexo (1.2.12), com erros condicionais baseados em processos OU integrados, ao analisar os dados do Exemplo 3, relativos aos tamanhos das feridas medidas em cada animal ao longo do tempo, até sua total cicatriza-ção . Devido a diferenças individuais na velocidade de cicatriza-ção , os dados se mostram incompletos. O mecanismo de não resposta, que ocorre quando o tamanho (não observado) da ferida se anula não pode ser ignorado (Laird, 1988). O modelo gaussiano não mais se aplica aos dados, pois se torna necessário considerar distribuições estritamente positivas. Nesse capítulo, mostramos uma utilização da distribuição normal truncada para modelar os dados do Exemplo 3.

No Capítulo 5, apresentamos as conclusões deste trabalho, e sugestões para a continuação da pesquisa.

Capítulo 2

Identificação de efeitos aleatórios

2.1 Introdução

Na Seção 1.4, mencionamos vários métodos gráficos e analíticos correntemente utilizados para identificar modelos de efeitos aleatórios. Neste Capítulo, propomos três formas alternativas para identificação de efeitos aleatórios. Na Seção 2.2, implementamos a idéia de Rutter & Elashoff (1992) em termos de testes de hipótese formais baseados na análise de regressão ordinária. Na Seção 2.3, propomos uma modificação neste método, que recai num modelo de análise de covariância em que os testes mencionados aparecem mais naturalmente. Na Seção 2.4, desenvolvemos um método para identificar os efeitos aleatórios através da análise dos elementos da matriz de dispersão amostral dos dados. Para ilustração dos métodos propostos, utilizamos os dados do Exemplo 1 (ver os diagramas paralelos de dispersão na Figura 2.1) com a idade expressa em

$$(dias + 1)/3$$

como no trabalho de Weiss & Lazaro (1992). Na Seção 2.5, utilizando o programa BMDP-5V (Schluchter, 1990) com a opção de estimação por MV através do método de Newton-Raphson, ajustamos o modelo identificado na nossa análise e o comparamos com o que foi identificado por Weiss & Lazaro (1992). Este programa considera como parâmetros do modelo apenas os elementos do vetor θ definido em (1.3.11) no cálculo do AIC. Isto pode ser razoável quando se utiliza o método de MVR, pois nesse caso a verossimilhança maximizada exclui os efeitos fixos, mas não deve ser válido em geral. Por esta razão, o AIC que calculamos neste capítulo continua levando em consideração também o número de efeitos fixos. A matriz de dispersão assintótica robusta proposta por White (1982) e por Royall (1986) foi calculada para os efeitos fixos. A comparação desta dispersão assintótica com a dispersão assintótica usual nos dá uma idéia da falta de ajuste do modelo

(White, 1982). Para o cálculo da variância assintótica corrigida (1.3.24) do efeito fixo $\beta_h = \mathbf{e}_h^t \boldsymbol{\beta} = \tau$ (onde $\mathbf{e}_h$ é um vetor com 1 na h -ésima linha e 0 nas demais), utilizamos

$$\partial \tau(\mathbf{y}; \boldsymbol{\theta}) / \partial \theta_m = \mathbf{e}_h^t (\mathbf{X}^t \mathbf{V}^{-1} \mathbf{X})^{-1} \mathbf{X}^t \mathbf{V}^{-1} (\partial \mathbf{V} / \partial \theta_m) \mathbf{V}^{-1} (\bar{\mathbf{y}} - \mathbf{X} \hat{\boldsymbol{\beta}}), \quad (2.1.1)$$

onde $\mathbf{X}$ e $\mathbf{V}$ são respectivamente a matriz de especificação dos efeitos fixos e a matriz de dispersão marginal, comuns às unidades amostrais, θ_m é o m -ésimo elemento de $\boldsymbol{\theta}$ e $\bar{\mathbf{y}}$ é o vetor das médias das observações individuais.

Além da função de verossimilhança e do AIC, medidas usuais para comparação de modelos, propomos também comparar a dispersão obtida através dos resíduos marginais ($\mathbf{G} = \sum_{j=1}^N \hat{\epsilon}_j \hat{\epsilon}_j^t / N$), com a dispersão amostral ($\hat{\mathbf{V}}_0 = \sum_{j=1}^N (\mathbf{y}_j - \bar{\mathbf{y}})(\mathbf{y}_j - \bar{\mathbf{y}})^t / N$) para diagnosticar o modelo para os efeitos fixos, e com a dispersão induzida pelo modelo ($\hat{\mathbf{V}} = \mathbf{V}(\hat{\boldsymbol{\theta}})$, dada por (1.3.11)) para diagnosticar o modelo para a dispersão marginal. Esta comparação deve ser feita sequencialmente, a segunda apenas quando o modelo para os efeitos fixos for aceitável, para não haver prejuízo no diagnóstico sobre a dispersão marginal. As variâncias generalizadas podem ser úteis para esta finalidade. Para comparar $\mathbf{G}$ e $\hat{\mathbf{V}}_0$ propomos

$$L_f(\hat{\mathbf{V}}_0, \mathbf{G}) = \log |\mathbf{G}^{-1} \hat{\mathbf{V}}_0|, \quad (2.1.2)$$

e para comparar $\mathbf{G}$ e $\hat{\mathbf{V}}$ propomos

$$L_v(\hat{\mathbf{V}}, \mathbf{G}) = |\log |\mathbf{G}^{-1} \hat{\mathbf{V}}||, \quad (2.1.3)$$

que se anulam quando $\hat{\mathbf{V}} = \mathbf{G}$, que é o caso do modelo parametrizado por médias de caselas e com matriz de covariâncias não estruturada (Modelo 0). A estatística L_f é equivalente à Λ de Wilks, obtida no teste da razão de verossimilhança para testar a estrutura da média num modelo normal multivariado com estrutura de dispersão livre. Como a estimativa de $\boldsymbol{\beta}$ dada por (1.3.9) é não viciada para quaisquer matrizes $\mathbf{V}_j, j = 1, \dots, N$ utilizadas, o valor crítico que se usa para $\log(\Lambda)$ deve servir para testar se o modelo especificado para os efeitos fixos é equivalente ao Modelo 0.

Figura 2.1: Diagramas paralelos de dispersão - Exemplo 1.

2.2 Identificação através de regressão linear simples

Esta metodologia consiste em identificar a forma funcional do valor esperado de (1.2.4) por meio de análise de regressão separada para cada unidade amostral. Sob a hipótese de independência condicional, (1.2.4) se reduz então a um modelo de regressão ordinária com a variância do erro comum a todas as unidades amostrais. Reparametrizações podem ser necessárias em (1.2.4) quando as matrizes $\mathbf{X}_j$ e $\mathbf{Z}_j$ tiverem algumas colunas em comum. Para que a forma funcional da resposta seja igual para todas as unidades amostrais, é necessário que os modelos individuais tenham as mesmas variáveis explicativas. Esta metodologia não requer dados balanceados, mas é necessário que o número de observações repetidas em cada unidade amostral seja maior que o número de parâmetros individuais estimados. Após a estimação, propõem-se testes para verificar se alguns parâmetros têm os mesmos valores para todas as unidades amostrais, e caso contrário, se eles têm distribuição normal. Propomos testar seqüencialmente:

a) se o parâmetro correspondente a uma determinada variável explicativa tem média (entre as unidades amostrais) diferente de zero, para verificar se ele pode ser considerado como um efeito fixo;

b) se existe diferença significativa entre algum parâmetro individual e a média, quando esta for significativamente diferente de zero; ou caso contrário, se algum parâmetro individual difere significativamente de zero, para verificar a necessidade de incluir parâmetros individuais no modelo;

c) a normalidade dos parâmetros individuais, caso algum teste do item b seja significativo, para verificar a possibilidade de tratá-los como efeitos aleatórios com distribuição normal e/ou se existe algum "outlier".

No ajuste de um modelo aos dados do Exemplo 1, Weiss & Lazaro (1992) consideram como efeitos fixos os parâmetros de um polinômio de segundo grau e como efeito aleatório apenas o coeficiente linear, mas Izenman & Williams (1989) sugerem um modelo polinomial de terceiro grau para as curvas individuais, sem no entanto mostrar o resultado da análise. Assim, tomamos como ponto de partida o ajuste de polinômios individuais de grau 3, em função da idade. Os coeficientes estimados e seus desvios em relação às respectivas médias, quando estas são significativamente diferentes de zero, aparecem na Tabela 2.1 (os coeficientes significativos estão em negrito). O erro padrão de um coeficiente médio é igual ao desvio padrão do coeficiente individual dividido pela raiz quadrada de N .

A análise dos coeficientes médios indica que tanto o intercepto quanto o termo cúbico não são significativos com $p > 0.10$, sugerindo que os efeitos fixos devem ser modelados apenas pelos coeficientes linear e quadrático. Não há nenhum intercepto individual significativamente diferente de zero com $p > 0.05$. Mas existem desvios dos coeficientes individuais em relação ao coeficiente médio no caso dos termos linear e quadrático, e alguns coeficientes individuais, no caso do termo cúbico, significativamente diferentes de zero, sugerindo que todos os coeficientes do polinômio cúbico, exceto o intercepto devem ser considerados como efeitos aleatórios. Nas Figuras 2.2 a 2.4 apresentamos gráficos de probabilidade normal, respectivamente para os coeficientes lineares, quadráticos e cúbicos; também indicamos os coeficientes de correlação entre cada coeficiente estimado e seu escore normal padrão. Podemos então concluir que o modelo normal é aceitável, já que o valor crítico da correlação para a aceitação da normalidade é de 0.9347 para $N = 10$ e de 0.9506 para $N = 15$ (Ryan et al., 1981), valores superados por todas as correlações calculadas.

Tabela 2.1: Estimativas individuais de polinômios cúbicos.

Animal	Coeficientes e desvios em relação à média					
	intercepto	linear	desvio	quadrático	desvio	cúbico
1	5.586	9.027	-12.425	4.900	5.945	-0.5556
2	-5.186	28.827	7.375	-3.089	-2.044	0.1556
3	-0.571	22.970	1.518	-1.462	-0.417	0.0111
4	-2.257	21.643	0.191	1.114	2.159	-0.2861
5	-10.143	34.914	13.462	-6.351	-5.306	0.4417
6	-2.657	27.146	5.694	-4.548	-3.503	0.3639
7	7.229	12.115	-9.337	1.473	2.518	-0.2139
8	5.657	11.582	-9.870	2.043	3.088	-0.2750
9	-12.243	39.444	17.992	-6.949	-5.904	0.5000
10	5.143	17.152	-4.300	0.407	1.452	-0.1806
11	3.286	17.734	-3.718	0.682	1.727	-0.1806
12	-7.557	31.006	9.554	-5.961	-4.916	0.4694
13	7.771	21.524	0.072	-1.270	-0.225	0.0417
14	4.171	5.240	-16.212	4.380	5.425	-0.4889
D.p.	(7.265)	(7.330)	(7.063)	(2.061)	(1.986)	(0.1703)
Média	-0.126	21.452		-1.045		-0.0141
D.p.	(1.942)	(1.959)		(0.551)		(0.0455)

Figura 2.2: Gráfico em papel de probabilidade normal dos coeficientes lineares individuais. Correlação = 0.994.

Figura 2.3: Gráfico em papel de probabilidade normal dos coeficientes quadráticos individuais. Correlação = 0.981.

Figura 2.4: Gráfico em papel de probabilidade normal dos coeficientes cúbicos individuais. Correlação = 0.974.

2.3 Identificação separada de modelos para efeitos aleatórios e para efeitos fixos

Quando os dados são balanceados, pode-se identificar um modelo para os efeitos aleatórios utilizando um modelo de análise de covariância em que a parametrização é feita através de médias de caselas e as covariáveis, comuns às unidades amostrais, são as potências dos tempos em que as observações foram feitas. Na prática, as médias das caselas, que estimam os efeitos fixos, são subtraídas das observações originais, e os desvios são explicados pelas covariáveis através de uma análise de regressão nas covariáveis. Os testes sobre os coeficientes deste modelo de regressão equivalem aos testes sobre os desvios entre os coeficientes individuais estimados e as respectivas médias na metodologia anterior. O resultado desta análise com os dados do Exemplo 1 e a idade elevada às potências de 0 a 3 como covariáveis é apresentado na Tabela 2.2, obtendo-se estimativas dos parâmetros quase idênticas às obtidas para os desvios dos coeficientes em relação às respectivas médias na Tabela 2.1 (os valores significativos encontram-se em negrito). Por esta razão, evitamos repetir os gráficos equivalentes às Figuras 2.2 a 2.4. Os testes são menos

conservadores pois os desvios em relação às médias das caselas são menos variáveis.

Tabela 2.2: Estimativas individuais de polinômios cúbicos ajustados aos desvios dos dados do Exemplo 1 em relação às médias das caselas.

Animal	Coeficientes			
	intercepto	linear	quadrático	cúbico
1	5.712	-12.425	5.945	-0.5415
2	-5.059	7.375	-2.044	0.1696
3	-0.445	1.518	-0.417	0.0252
4	-2.131	0.192	2.159	-0.2720
5	-10.016	13.463	-5.306	0.4558
6	-2.531	5.694	-3.503	0.3780
7	7.355	-9.337	2.518	-0.1998
8	5.784	-9.870	3.088	-0.2609
9	-12.116	17.992	-5.904	0.5141
10	5.269	-4.300	1.452	-0.1665
11	3.412	-3.718	1.727	-0.1665
12	-7.431	9.554	-4.916	0.4835
13	7.898	0.072	-0.225	0.0558
14	4.298	-16.212	5.425	-0.4748
D.p.	(6.397)	(6.454)	(1.815)	(0.1499)

Figura 2.5: Diagrama de dispersão dos pesos médios.

Tabela 2.3: Estimativas dos parâmetros dos polinômios ajustados às médias dos instantes de observação (Exemplo 1). Entre parênteses estão os desvios padrões.

Modelo	Coeficientes				$\hat{\sigma}$
	intercepto	linear	quadrático	cúbico	
cúbico completo	-0.127 (3.444)	21.452 (3.475)	-1.045 (0.977)	-0.014 (0.080)	1.186
cúbico sem intercepto	-	21.330 (0.895)	-1.013 (0.376)	-0.017 (0.038)	1.028
quadrático completo	-0.634 (1.609)	22.029 (0.922)	-1.214 (0.113)	-	1.032
quadrático sem intercepto	-	21.694 (0.322)	-1.177 (0.056)	-	0.941

Um modelo para o efeito fixo pode ser identificado através do ajuste de um polinômio aos valores médios das caselas, cujo diagrama de dispersão é apresentado na Figura 2.5. Um polinômio cúbico com todos os termos foi ajustado; as estimativas correspondentes estão na primeira linha da Tabela 2.3. Apenas o termo linear é significativo, podendo-se eliminar alguns coeficientes. Quando o intercepto é eliminado, o termo quadrático também se mostra significativo, como pode ser notado na segunda linha da Tabela 2.3. Modelos quadráticos com e sem o intercepto são apresentados nas outras linhas da Tabela 2.3. Devido à escala utilizada para a idade, o intercepto não pode ser interpretado como peso médio ao nascer, não havendo necessidade de sua inclusão no modelo, a menos que seja significativo. O modelo com o menor erro quadrático médio ($\hat{\sigma}^2$) é o modelo quadrático sem o intercepto, como havia sido identificado na seção 2.2. Já se podia esperar por esta conclusão, devido à coincidência das estimativas nas Tabelas 2.1 e 2.2, mostrando que o modelo para efeitos fixos identificado na Seção 2.2 é basicamente coincidente com o modelo de médias das caselas.

2.4 Identificação de um modelo para efeitos aleatórios através da análise da matriz de dispersão amostral

A matriz de dispersão do modelo de efeitos aleatórios é dada por (1.3.11). A hipótese de independência condicional transfere para os efeitos aleatórios a capacidade de modelar as covariâncias entre as observações, sendo a dispersão intra-individual responsável apenas por uma parcela da variância das observações. Assim, com a finalidade de identificação dos efeitos aleatórios, propomos modelar a dispersão amostral observada, quando $\mathbf{Z}_j = \mathbf{Z}$, $j = 1, \dots, N$, pela dispersão devida apenas aos efeitos aleatórios, $\mathbf{Z}\mathbf{\Gamma}\mathbf{Z}^t$, cujo elemento (s, u) é dado pela forma quadrática $\mathbf{z}_s^t \mathbf{\Gamma} \mathbf{z}_u$ (onde $\mathbf{z}_s^t$ e $\mathbf{z}_u^t$ são respectivamente as linhas s e u da matriz $\mathbf{Z}$). A linha s de $\mathbf{Z}\mathbf{\Gamma}\mathbf{Z}^t$ é obtida fixando-se $\mathbf{z}_s$ e fazendo variar $\mathbf{z}_u$. Num modelo polinomial, $\mathbf{z}_u$ é um vetor composto pelas potências de t , que é o tempo correspondente à u -ésima observação, o que faz com que os elementos da linha s de $\mathbf{Z}\mathbf{\Gamma}\mathbf{Z}^t$ sejam gerados por um polinômio em t . Ao variar s , podem ser observados os elementos de uma família de polinômios em t , indexados por t_s . Assim, o problema de identificação dos

efeitos aleatórios coincide com o problema de identificação das possíveis colunas de $\mathbf{Z}$ que estejam correlacionadas simultaneamente com todas as linhas da matriz de dispersão observada. Vamos abordar este problema através de uma análise de regressão em cada linha da matriz de dispersão observada.

Tabela 2.4: Matriz de dispersão amostral (Exemplo 1).

Idade	1	2	3	4	5	6	7
1	8.231						
2	12.173	21.105					
3	9.132	19.950	31.577				
4	10.241	24.312	47.964	97.625			
5	8.541	26.638	57.989	118.140	166.066		
6	8.777	24.868	56.320	119.721	162.825	174.897	
7	12.826	29.238	52.870	99.659	133.177	141.837	135.206

Figura 2.6: Diagramas paralelos de dispersão para as linhas da matriz de dispersão amostral do Exemplo 1.

A matriz de dispersão amostral dos dados do Exemplo 1 se encontra na Tabela 2.4; os diagramas paralelos de dispersão das suas linhas estão

dispostos na Figura 2.6, podendo-se notar uma forma sigmoidal das curvas correspondentes às últimas linhas ($s = 3$ a 7). Isto sugere que elas podem ser modeladas por polinômios cúbicos. As estimativas dos coeficientes destes polinômios, ajustados a todas as linhas da matriz de dispersão amostral se encontram na Tabela 2.5.

Tabela 2.5: Coeficientes dos polinômios cúbicos ajustados às linhas da matriz de dispersão amostral do Exemplo 1. Entre parênteses estão os desvios padrões.

Linha	Coeficientes				$\hat{\sigma}$	R^2
	intecepto	linear	quadrático	cúbico		
1	2.27 (4.04)	9.61 (4.08)	-2.936 (1.147)	0.2568 (0.0947)	1.39	75.3%
2	2.56 (7.14)	13.70 (7.20)	-2.758 (2.026)	0.1978 (0.1674)	2.46	91.6%
3	4.34 (7.38)	0.94 (7.45)	4.907 (2.094)	-0.5697 (0.1730)	2.54	99.3%
4	28.04 (20.26)	-38.26 (20.44)	22.992 (5.750)	-2.2785 (0.4749)	6.98	99.0%
5	49.27 (24.14)	-72.98 (24.35)	37.352 (6.849)	-3.5786 (0.5657)	8.31	99.3%
6	56.09 (15.90)	-81.14 (16.05)	39.410 (4.513)	-3.6936 (0.3728)	5.48	99.7%
7	32.19 (15.89)	-38.41 (16.03)	22.540 (4.509)	-2.1097 (0.3724)	5.47	99.6%

Os valores do coeficiente de determinação (R^2) mostram que os polinômios cúbicos se ajustam bem às linhas da matriz de dispersão amostral dos dados do Exemplo 1. Nem todos os parâmetros, entretanto, são significativos, indicando a possibilidade de eliminação de alguns. Esta simplificação deve ser

feita simultaneamente em todas as linhas, para que os modelos associados tenham as mesmas variáveis explicativas.

Tabela 2.6: Coeficientes dos polinômios cúbicos sem os interceptos ajustados às linhas da matriz de dispersão amostral do Exemplo 1. Entre parênteses estão os desvios padrões.

Linha	Coeficientes			$\hat{\sigma}$
	linear	quadrático	cúbico	
1	11.798 (1.104)	-3.513 (0.463)	0.3016 (0.0463)	1.27
2	16.163 (1.895)	-3.408 (0.796)	0.2484 (0.0794)	2.17
3	5.123 (2.026)	3.803 (0.850)	-0.4839 (0.0849)	2.32
4	-11.250 (6.741)	15.863 (2.830)	-1.7240 (0.2826)	7.74
5	-25.514 (9.696)	24.825 (4.070)	-2.6043 (0.4065)	11.13
6	-27.115 (9.376)	25.151 (3.936)	-2.5846 (0.3931)	10.76
7	-7.404 (6.355)	14.356 (2.668)	-1.4732 (0.2664)	7.29

Os coeficientes de graus maiores, em geral, são significativos, sendo os interceptos, em geral, não significativos, exceto na linha 6. Estes termos foram retirados do modelo e os resultados da análise apresentados na Tabela 2.6. Podemos notar que quase todos os coeficientes, com exceção de um coeficiente cúbico (linha 6), tiveram seus desvios padrões estimados diminuídos. O ajuste, medido através dos desvios padrões estimados dos erros, piorou

significativamente apenas para a penúltima linha, resultado já esperado por ser a única linha para a qual o intercepto (e todos os demais coeficientes) se mostrara(m) significativamente diferente(s) de zero. O modelo identificado para os efeitos aleatórios é portanto o mesmo das Seções anteriores.

2.5 Estimação e comparação entre modelos

Tabela 2.7: Modelos propostos para os dados do Exemplo 1.

Modelo	Efeitos fixos	Efeitos aleatórios
0	médias das caselas	estrutura de covariâncias livre
1	quadrático	linear sem o intercepto
2	quadrático	cúbico sem o intercepto
3	quadrático sem o intercepto	cúbico sem o intercepto

Nesta Seção, vamos ajustar o modelo identificado nas Seções anteriores (Modelo 3 da Tabela 2.7) e compará-lo com o modelo identificado por Weiss & Lazaro (1992) (Modelo 1 da Tabela 2.7). Para facilitar a comparação, propomos também estimar um modelo híbrido (Modelo 2 da Tabela 2.7), em relação ao qual ambos os modelos citados ficam encaixados. Um modelo de referência para todos os outros é aquele que tem a matriz de dispersão amostral como estimativa da dispersão, ou seja o modelo de médias de caselas com estrutura de covariâncias livre (Modelo 0), em relação ao qual todos os modelos citados ficam encaixados. Como os parâmetros deste modelo já foram utilizados neste Capítulo, vamos apenas compará-lo a outros modelos na Tabela 2.18. As estimativas dos parâmetros dos Modelos 1 a 3 se encontram nas Tabelas 2.8 (efeitos fixos) e 2.9 (vetor θ).

Na Tabela 2.8, os desvios padrões assintóticos robustos (White, 1982; Royall, 1986) dos efeitos fixos, dados pelas raízes quadradas dos elementos da diagonal da matriz de dispersão assintótica C (1.3.40), foram calculados e apresentados juntamente com os desvios padrões assintóticos usuais, dados pelas raízes quadradas dos elementos da diagonal de $\text{Var}(\hat{\beta})$ (1.3.10). A discrepância entre os dois tipos de desvios padrões assintóticos é uma indicação de falta de ajuste do modelo utilizado. Nos Modelos 1 e 2, notam-se diferenças, mas em sentidos opostos, entre os desvios padrões assintóticos usuais e robustos, no primeiro, devido a uma possível falha na especificação dos efeitos aleatórios, e no segundo, devido a um excesso presumível na

parametrização dos efeitos fixos. Esta discrepância deixa de existir no Modelo 3, que, em relação aos Modelos 1 e 2, apresenta sensível melhora na precisão das estimativas dos efeitos fixos.

Tabela 2.8: Parâmetros fixos estimados no ajuste de modelos quadráticos no efeito fixo, com intercepto (modelos 1 e 2) ou sem (modelo 3). Entre parênteses estão os desvios padrões assintóticos (d.p.), tradicionais e robustos.

Parâmetros	Modelo 1	Modelo 2	Modelo 3
intercepto	-0.63367	-0.29439	-
(d.p.)	(1.64563)	(1.65121)	-
(d.p. robusto)	(2.16557)	(1.06462)	-
linear	22.02925	21.80078	21.61852
(d.p.)	(1.07688)	(1.22733)	(0.67918)
(d.p. robusto)	(1.59174)	(0.91492)	(0.67842)
quadrático	-1.21412	-1.19474	-1.17774
(d.p.)	(0.11522)	(0.12268)	(0.07719)
(d.p. robusto)	(0.14546)	(0.10129)	(0.07497)

Do Modelo 1 para os Modelos 2 e 3, pode-se notar uma sensível melhora no ajuste das observações individuais, medido pela variância estimada do erro condicional, $\hat{\sigma}^2$ (Tabela 2.9). O Modelo 3 pode ainda ser melhorado eliminando-se uma possível redundância representada pelo parâmetro σ_{32} , cujo valor estimado aponta para a parametrização

$$\sigma_{32} = (\sigma_2^2 \sigma_3^2)^{1/2}.$$

Os valores estimados para cada lado da expressão acima diferem apenas na terceira casa decimal. Isto indica que, mesmo sob a nova parametrização, os valores estimados não se alterariam sensivelmente, razão pela qual evitamos o ajuste do novo modelo. A estimativa da matriz de dispersão assintótica do estimador de θ do novo modelo pode ser obtida a partir da estimativa da matriz de dispersão assintótica do Modelo 3 (Tabela 2.10). Como as estimativas das covariâncias assintóticas do estimador de σ_{32} com os demais parâmetros são nulas, basta eliminar a linha e a coluna correspondentes a σ_{32} para se obter a nova matriz de dispersão assintótica.

Tabela 2.9: Estimativa dos elementos do vetor θ . Entre parênteses estão os desvios padrões assintóticos.

Parâmetros	Modelo 1	Modelo 2	Modelo 3
σ_1^2	3.78351 (1.47228)	21.1494 (6.0064)	21.1912 (6.0119)
σ_{21}	-	-9.7149 (1.7154)	-9.7387 (1.7172)
σ_2^2	-	5.5449 (0.1949)	5.5585 (0.1952)
σ_{31}	-	0.9291 (0.1449)	0.9315 (0.1451)
σ_{32}	-	-0.5513 (4.1815)	-0.5527 (4.1725)
σ_3^2	-	0.0554 (0.0019)	0.0556 (0.0020)
σ^2	15.6114 (2.4089)	5.9247 (1.1175)	5.9246 (1.1175)

Tabela 2.10: Dispersão assintótica da estimativa do vetor θ do Modelo 3.

par.	σ_1^2	σ_{21}	σ_2^2	σ_{31}	σ_{32}	σ_3^2	σ^2
σ_1^2	36.1426						
σ_{21}	-9.8556	2.9489					
σ_2^2	0.8557	-0.2695	0.0381				
σ_{31}	0.7944	-0.2440	0.0199	0.0210			
σ_{32}	0.0000	0.0000	0.0000	0.0000	17.4100		
σ_3^2	-0.0081	0.0027	-0.0004	-0.0002	0.0000	0.0000	
σ^2	-0.4227	0.0915	-0.0024	-0.0071	0.0000	-0.0000	1.2489

Com esta providência, foram calculadas as correções de Kackar & Harville (1984) e de Kass & Steffey (1989) para os desvios padrões assintóticos dos efeitos fixos (Tabela 2.11). A correção de Kackar & Harville (1984) difere da correção de Kass & Steffey (1989) pois a primeira envolve a estimativa da matriz de dispersão do erro marginal médio $\bar{y} - \mathbf{X}\beta$, induzida pelo modelo, e a segunda utiliza sua matriz de dispersão residual $(\bar{y} - \mathbf{X}\hat{\beta})(\bar{y} - \mathbf{X}\hat{\beta})^t$. A discrepância entre as duas correções pode indicar uma falta de ajuste do Modelo 3.

Tabela 2.11: Correção de Kackar & Harville (1984) (KH) e de Kass & Steffey (1989) (KS) para as variâncias assintóticas dos efeitos fixos do Modelo 3.

Parâmetro	Correção para a var. assintótica	
	KH	KS
linear	$3.287 \cdot 10^{-2}$	$0.200 \cdot 10^{-2}$
quadrático	$4.257 \cdot 10^{-4}$	$0.258 \cdot 10^{-4}$

As matrizes de dispersão marginal induzidas pelos Modelos 1 a 3 estão respectivamente nas Tabelas 2.12 a 2.14. Enquanto as duas últimas são quase iguais, a primeira difere das duas, elemento por elemento, devido à diferença no modelo de efeitos aleatórios. As respectivas matrizes de dispersão residual estão nas Tabelas 2.15 a 2.17. Todas as três matrizes são semelhantes à matriz de dispersão amostral (Tabela 2.4), indicando que as esperanças marginais estão igualmente bem estimadas. A diferença entre a dispersão marginal induzida e a dispersão residual devido a um modelo é uma indicação da falta de ajuste do modelo para a dispersão marginal, quando o modelo para a esperança marginal está correto. Estas comparações são mais facilmente feitas através das variâncias generalizadas apresentadas na Tabela 2.18, junto com as medidas de ajuste dos modelos. O Modelo 1 é pior que os outros modelos propostos em relação a todas as medidas de ajuste. L_f e L_v acompanham respectivamente a variação da magnitude de AIC e de $\log(L)$ mas enquanto L_f não varia muito entre os Modelos polinomiais, L_v muda de patamar quando passa do Modelo 1 para os Modelos 2 e 3, e destes para o Modelo 0. O valor crítico para L_f com $\alpha = 0.10$ é de 12.427 (11.352) para o modelo de efeitos fixos com 2 (3) parâmetros - Modelo(s) 3 (1 e 2) - sugerindo que os modelos de efeitos fixos propostos são aceitáveis.

Tabela 2.12: Estimativa da dispersão marginal induzida pelo Modelo 1.

idade	1	2	3	4	5	6	7
1	19.395						
2	7.567	30.745					
3	11.350	22.701	49.663				
4	15.134	30.268	45.402	76.147			
5	18.917	37.835	56.753	75.670	110.199		
6	22.701	45.402	68.103	90.804	113.505	151.818	
7	26.485	52.969	79.454	105.942	132.423	158.907	201.003

Tabela 2.13: Estimativa da dispersão marginal induzida pelo Modelo 2.

idade	1	2	3	4	5	6	7
1	14.000						
2	9.308	21.798					
3	6.296	20.274	43.794				
4	1.640	23.091	55.016	94.001			
5	-2.061	24.903	67.648	112.934	153.441		
6	-2.208	26.287	71.701	120.251	158.153	177.548	
7	3.799	27.824	63.108	100.689	131.601	146.880	143.485

Tabela 2.14: Estimativa da dispersão marginal induzida pelo Modelo 3.

idade	1	2	3	4	5	6	7
1	14.010						
2	9.311	21.799					
3	6.284	20.227	43.809				
4	1.611	23.081	55.052	94.090			
5	-2.103	24.888	67.700	113.062	153.625		
6	-2.251	26.272	71.754	120.383	158.344	177.748	
7	3.774	27.815	63.141	100.770	131.719	147.007	143.577

Tabela 2.15: Dispersão dos resíduos marginais do Modelo 1.

idade	1	2	3	4	5	6	7
1	8.430						
2	11.782	21.871					
3	9.127	19.961	31.577				
4	10.406	23.987	47.959	97.763			
5	8.993	25.753	57.976	118.516	167.090		
6	8.155	26.086	56.338	119.205	161.418	176.831	
7	13.027	28.845	52.864	99.826	133.631	141.212	135.408

Tabela 2.16: Dispersão dos resíduos marginais do Modelo 2.

idade	1	2	3	4	5	6	7
1	8.331						
2	11.908	21.803					
3	9.183	19.817	31.602				
4	10.442	23.781	48.065	98.029			
5	8.962	25.527	58.200	118.986	167.836		
6	8.442	25.751	56.152	119.049	161.419	176.014	
7	13.067	28.603	52.990	100.142	134.187	141.034	135.783

Tabela 2.17: Dispersão dos resíduos marginais do Modelo 3.

idade	1	2	3	4	5	6	7
1	8.266						
2	12.017	21.799					
3	9.181	19.735	31.644				
4	10.390	23.647	48.170	98.262			
5	8.827	25.370	58.382	119.355	168.383		
6	8.613	25.593	56.095	119.027	161.501	175.653	
7	12.997	28.481	53.104	100.384	134.559	141.047	136.031

O intercepto nulo no conjunto de efeitos fixos pode ser testado pela razão de verossimilhança entre os Modelos 2 e 3, não havendo diferença significativa com $p = 0.20$. Com relação aos modelos para os efeitos aleatórios, nota-se uma diferença significativa ($p = 0.001$) entre os Modelos 1 e 2. A razão de verossimilhança entre o Modelo 0 e qualquer um dos outros é significativa com $p = 0.001$, sugerindo haver falta de ajuste do modelo para a dispersão

marginal devido ao modelo de efeitos aleatórios com independência condicional.

Tabela 2.18: Medidas de ajuste e logaritmo das variâncias generalizadas das matrizes de dispersão marginal estimadas e residuais dos Modelos 0, 1, 2 e 3 (reparametrizado).

Medidas de ajuste	Modelo 0	Modelo 1	Modelo 2	Modelo 3
$\log \hat{\mathbf{V}} $	16.401	22.789	19.916	19.918
$\log \mathbf{G} $	16.401	17.897	17.772	17.664
L_f	0	1.496	1.371	1.263
L_v	0	4.892	2.144	2.254
$\log(L)$	-253.866	-298.581	-277.664	-277.680
AIC	-288.866	-303.581	-287.664	-285.680

Figura 2.7: Diagramas paralelos de dispersão das curvas individuais e média estimadas (Modelo 3).

Isto mostra que o intercepto fixo pode ser retirado, sem perda no grau de ajuste, como havia sido diagnosticado através dos desvios padrões robustos (Tabela 2.8), mas mesmo o Modelo 3 ainda tem um ajuste que deixa a

desejar quando comparado ao Modelo 0, fato sugerido pela diferença entre as correções KH e KS das variâncias assintóticas dos efeitos fixos. Quando se associa o critério de ajuste ao da parcimônia no número de parâmetros do modelo, como no AIC, o Modelo 3 pode ser considerado melhor que os demais. Os diagramas paralelos de dispersão das respostas individuais estimadas sob o Modelo 3 se encontram na Figura 2.7.

2.6 Conclusão e discussão

Quando os dados são balanceados com relação ao tempo, completos ou MCAR, um modelo de referência é o modelo parametrizado por médias de caselas e com matriz de dispersão não estruturada (Modelo 0), em relação ao qual outros modelos propostos devem ser comparados, quando a estimação é feita por MV.

Um modelo misto de efeitos aleatórios com erros condicionalmente independentes pode ser identificado a partir dos parâmetros estimados do Modelo 0. O diagrama de dispersão das médias das caselas pode sugerir uma outra parametrização para os efeitos fixos. Um modelo para os efeitos aleatórios pode ser identificado a partir da matriz de dispersão amostral. A identificação de um modelo para os efeitos aleatórios pode, alternativamente, ser feita a partir de um modelo de análise de covariância, parametrizado por médias de caselas (ou por algum outro modelo identificado ou conhecido para os efeitos fixos) e com covariáveis comuns às unidades amostrais (termos polinomiais, por exemplo). Os coeficientes relativos a estas covariáveis podem sugerir um modelo para os efeitos aleatórios. Testes de normalidade com os coeficientes estimados podem confirmar o modelo gaussiano para os efeitos aleatórios.

Quando os dados não são balanceados com relação ao tempo, eles podem ser modelados apenas pelo primeiro estágio (1.2.4) do modelo de efeitos aleatórios, em que todos os parâmetros são considerados fixos. Após uma reparametrização necessária para eliminar os termos redundantes, a estimação é feita por regressão ordinária. Testes sobre os parâmetros podem indicar possíveis modelos para os efeitos fixos e efeitos aleatórios, e a normalidade dos últimos. No Capítulo 4 utilizamos esta metodologia para identificar os efeitos fixos e aleatórios de um modelo não linear ajustado aos dados apresentados no Exemplo 3. Trata-se de um método alternativo também para os dados balanceados com relação ao tempo, mas neste caso os métodos

anteriores são mais simples.

Após o ajuste do modelo identificado, deve-se proceder à análise dos resíduos marginais, avaliando a normalidade dos efeitos aleatórios e dos resíduos condicionais para verificar o ajuste do modelo, o que deixamos de apresentar, por fugir ao problema central, que é a identificação do modelo para os efeitos aleatórios. A seleção de modelos pode ser baseada nos testes da razão de verossimilhança, no AIC e na comparação entre as dispersões assintóticas tradicionais e robustas. Os dois primeiros métodos devem ser preferidos quando for possível ajustar o Modelo 0. Caso contrário, o último método pode mostrar sua utilidade para diagnosticar não só o modelo para os efeitos fixos, como foi proposto por Rutter & Elashoff (1992), como também o modelo para a dispersão marginal, quando complementado por inferências sobre os parâmetros θ .

Se a seleção resultar num modelo misto de efeitos aleatórios, deve-se proceder ao cálculo das correções KH e KS para as variâncias assintóticas dos efeitos fixos. A diferença entre as suas ordens de grandeza pode sugerir possível falta de ajuste do modelo utilizado.

Tanto o diagnóstico de um modelo com base na comparação entre as variâncias assintóticas tradicionais e robustas, ou com base na comparação entre KH e KS, quanto a comparação entre modelos baseada no teste da razão de verossimilhança ou AIC, envolvem a dispersão residual (G) e a dispersão induzida pelo modelo ($\hat{V}$). A comparação entre dois modelos com estrutura de dispersão livre resulta na estatística Λ de Wilks, dada pela razão entre as variâncias generalizadas das dispersões residuais. Propomos considerar a mesma estatística para comparar os modelos de efeitos fixos, pois se seus estimadores são não viciados, as dispersões residuais devem estar próximas da dispersão marginal. Para diagnosticar o modelo para a dispersão marginal, é intuitivo que se faça uma comparação entre G e $\hat{V}$, mas não existem testes desenvolvidos para esta finalidade, podendo-se sugerir o estudo da distribuição de estatísticas como L_v .

O modelo com estrutura livre para a dispersão marginal se ajusta melhor que o modelo polinomial de efeitos aleatórios para os dados do Exemplo 1, como vimos na seção 2.5. Creditamos este fato à limitação da aproximação da trajetória de um processo estocástico por um modelo polinomial com independência condicional. Um polinômio, obviamente, pode se aproximar razoavelmente de qualquer curva, mas não se pode exigir, por exemplo, que os erros condicionais associados à aproximação polinomial sejam independentes.

Torna-se necessário o conhecimento teórico das possíveis estruturas de dispersão que possam ser utilizadas na identificação da dispersão marginal. No caso do modelo de efeitos aleatórios, trata-se da identificação simultânea dos efeitos aleatórios e da estrutura de dispersão condicional, tópico que abordamos no Capítulo 3. O ideal seria poder adotar um modelo mecanístico, como na análise dos dados do Exemplo 3 que apresentamos no Capítulo 4, mas nem sempre esta opção é viável, devido em geral à falta de conhecimento sobre o processo observado.

Capítulo 3

Identificação da estrutura de dispersão condicional

3.1 Introdução

Neste Capítulo, estudamos o comportamento das linhas das matrizes de dispersão (1.3.11) e de correlação correspondente com a finalidade de possibilitar a identificação deste modelo num determinado conjunto de dados. Em particular, nos restringimos a um modelo com dois efeitos aleatórios correspondentes aos coeficientes de uma reta e com uma dispersão intra-individual $\mathbf{R} = (R_{su})$ bem especificada. Nesse caso, a matriz (1.3.11), aqui denotada $\mathbf{V} = (V_{su})$ sem a identificação da unidade amostral é dada por:

$$\begin{aligned}\mathbf{V} &= \begin{pmatrix} \mathbf{1}_T & \mathbf{t} \end{pmatrix} \begin{pmatrix} \sigma_0^2 & \sigma_{01} \\ \sigma_{01} & \sigma_1^2 \end{pmatrix} \begin{pmatrix} \mathbf{1}_T^t \\ \mathbf{t}^t \end{pmatrix} + \sigma^2 \mathbf{R} \\ &= \sigma_0^2 \mathbf{1}_T \mathbf{1}_T^t + \sigma_{01} (\mathbf{t} \mathbf{1}_T^t + \mathbf{1}_T \mathbf{t}^t) + \sigma_1^2 \mathbf{t} \mathbf{t}^t + \sigma^2 \mathbf{R}\end{aligned}$$

onde $\mathbf{t} = (t_1, t_2, \dots, t_T)^t$ é um vetor cujos elementos correspondem aos sucessivos instantes de observação, σ_0^2 e σ_1^2 são as variâncias respectivamente do coeficiente linear e do coeficiente angular e σ_{01} é a covariância entre os dois coeficientes. O elemento V_{su} é dado pela função

$$V(t_s, t_u) = \sigma_0^2 + \sigma_{01}(t_s + t_u) + \sigma_1^2 t_s t_u + \sigma^2 R_{su}. \quad (3.1.1)$$

A variância marginal no instante t é dada por

$$V(t, t) = V(t) \quad (3.1.2)$$

e para t_s fixo, qualquer elemento da linha s da matriz $\mathbf{V}$ é denotado por

$$V_s(t) = V(t_s, t) \quad (3.1.3)$$

fazendo $t = t_u$, $u = 1, \dots, T$. Quando t varia continuamente numa região da reta, $V(t)$ e $V_s(t)$ são funções reais contínuas, que denominamos funções geradoras respectivamente das variâncias marginais e da linha s da matriz $\mathbf{V}$. Conseqüentemente, pode-se dizer que os elementos da linha s da matriz de correlações marginais são geradas pela função

$$\rho_s(t) = V_s(t)/\sqrt{V(t_s)V(t)}. \quad (3.1.4)$$

com t_s fixo. Sua derivada é dada por

$$\dot{\rho}_s(t) = [\dot{V}_s(t)/\sqrt{V(t)} - \dot{V}(t)V_s(t)/2\sqrt{(V(t))^3}]/\sqrt{V(t_s)} \quad (3.1.5)$$

onde $\dot{V}_s(t)$ e $\dot{V}(t)$ são respectivamente as derivadas de V_s e de V avaliadas em t . Para efeito de estudo dos sinais desta função, basta considerar os sinais do segundo membro de

$$\sqrt{V(t_s)(V(t))^3}\dot{\rho}_s(t) = V(t)\dot{V}_s(t) - V_s(t)\dot{V}(t)/2. \quad (3.1.6)$$

Nas duas Seções seguintes, especificamos a matriz $\mathbf{R}$ com a finalidade de avaliar o comportamento de $V(t)$, de $V_s(t)$ e de $\rho_s(t)$. Na Seção 3.2 consideramos o modelo IC para o qual

$$R_{su} = 1_{t_s}(t_u) = \begin{cases} 1 & \text{se } t_u = t_s \\ 0 & \text{caso contrário} \end{cases} \quad (3.1.7)$$

e na Seção 3.3 consideramos o modelo com dispersão intra-individual dada por um MB para o qual $R_{su} = \min(t_s, t_u)$. Os dados do Exemplo 2, com a idade expressa em semestres adicionais à idade de 8 anos, são utilizados na Seção 3.4, com a finalidade de identificação de um dos modelos acima especificados através da análise das linhas e da diagonal principal da matriz de dispersão amostral. Os mesmos dados também são utilizados na Seção 3.5, onde ajustamos os modelos considerados neste capítulo, que são comparados entre si e ao modelo com estrutura livre na matriz de covariâncias, segundo critérios apresentados no Capítulo 2, para averiguar a capacidade de identificação do método proposto. Estes dados (ver diagramas paralelos de dispersão na Figura 3.1) foram analisados por diversos autores, sempre com a utilização de modelos com dois efeitos fixos correspondentes aos coeficientes de uma reta. Componentes aleatórios, também correspondentes aos mesmos coeficientes, foram utilizados por Elston & Grizzle (1962) e Fearn (1975), o

Figura 3.1: Diagramas paralelos de dispersão - Exemplo 2.

primeiro com uma restrição de independência entre os efeitos aleatórios, e ambos com IC. Na Seção 3.6, mencionamos as conclusões deste capítulo e apresentamos sugestões para pesquisas futuras.

3.2 Modelo de efeitos aleatórios com independência condicional

O objetivo desta Seção é caracterizar a matriz de dispersão induzida pelo modelo de retas com coeficientes aleatórios e IC através de $V_s(t)$, $V(t)$ e $\rho_s(t)$, com a finalidade de identificação deste modelo pela inspeção da matriz de dispersão amostral e/ou da matriz de correlações amostrais. A função $V_s(t)$ para o modelo com IC é dada por

$$V_s(t) = \sigma_0^2 - \sigma_{01}^2/\sigma_1^2 + \sigma_1^2(t_s + \sigma_{01}/\sigma_1^2)(t + \sigma_{01}/\sigma_1^2) + \sigma^2 1_{t_s}(t). \quad (3.2.1)$$

Trata-se de uma equação de reta para t_s fixo, com uma descontinuidade em $t = t_s$, quando seu valor representa a variância marginal $V(t)$. Esta por sua vez, é uma parábola com valores positivos, com mínimo em $t = -\sigma_{01}/\sigma_1^2$.

Figura 3.2: $V_s(t)$ e $V(t)$, sob o modelo com IC.

Para este valor do argumento, o valor de V_s é igual a $\sigma_0^2 - \sigma_{01}^2/\sigma_1^2$ qualquer que seja o índice s , o que significa que $(-\sigma_{01}/\sigma_1^2, \sigma_0^2 - \sigma_{01}^2/\sigma_1^2)$ é o ponto de concorrência das retas $V_s(t)$, $s = 1, \dots, T$. Quando $t_s \neq -\sigma_{01}/\sigma_1^2$, $V_s(t)$ muda de sinal em

$$t = -(\sigma_0^2 + \sigma_{01}t_s)/(\sigma_1^2t_s + \sigma_{01}).$$

Caso contrário, ou seja, para $t_s = -\sigma_{01}/\sigma_1^2$, $V_s(t)$ se mantém constantemente igual a $\sigma_0^2 - \sigma_{01}^2/\sigma_1^2$, sendo representada na Figura 3.2 pela reta V_{s3} . As retas V_{s1} e V_{s2} para os quais t_{s1} e t_{s2} são equidistantes de $-\sigma_{01}/\sigma_1^2$ são simétricas em relação à reta $t = -\sigma_{01}/\sigma_1^2$.

A função (3.1.4) é dada por

$$\rho_s(t) = \frac{\sigma_0^2 + \sigma_{01}(t_s + t) + \sigma_1^2 t_s t}{\sqrt{V(t_s)(\sigma_0^2 + \sigma^2 + 2\sigma_{01}t + \sigma_1^2 t^2)}} \quad (3.2.2)$$

que converge para

$$\pm (\sigma_1^2 t_s + \sigma_{01})/(\sigma_1 \sqrt{V(t_s)}) \quad (3.2.3)$$

quando $t \rightarrow \pm\infty$. As funções ρ_{s1} e ρ_{s2} , com t_{s1} e t_{s2} equidistantes de $-\sigma_{01}/\sigma_1^2$, são simétricas em torno de $t = -\sigma_{01}/\sigma_1^2$ (ver Figura 3.3).

Figura 3.3: $\rho_s(t)$, sob o modelo com IC.

Como

$$\begin{aligned}\dot{V}_s(t) &= \sigma_{01} + \sigma_1^2 t_s \\ \dot{V}(t)/2 &= \sigma_{01} + \sigma_1^2 t\end{aligned}$$

segue que o segundo membro de (3.1.6) é

$$\sigma^2 \sigma_{01} + (\sigma^2 \sigma_1^2 + \sigma_0^2 \sigma_1^2 - \sigma_{01}^2) t_s - (\sigma_0^2 \sigma_1^2 - \sigma_{01}^2) t, \quad (3.2.4)$$

que é uma função linear decrescente em t para t_s fixo, cuja raiz,

$$t_s + \frac{\sigma^2 \sigma_1^2}{\sigma_0^2 \sigma_1^2 - \sigma_{01}^2} (t_s + \sigma_{01} / \sigma_1^2),$$

corresponde ao ponto de máximo de $\rho_s(t)$. Sua posição em relação a t_s depende apenas do sinal de $t_s - \sigma_{01} / \sigma_1^2$ (na Figura 3.3 os máximos de $\rho_{s_1}(t)$ e de $\rho_{s_2}(t)$ ocorrem respectivamente em t_{m_1} e t_{m_2}). Trata-se de máximo local, já que o máximo global (= 1) ocorre quando $t = t_s$. O máximo local pode coincidir com o global quando $t_s = -\sigma_{01} / \sigma_1^2$.

A existência de um máximo local ao longo de uma linha da matriz de correlações marginais de um conjunto de dados longitudinais não foi notificada por nenhum dos autores pesquisados neste trabalho, havendo citações

(Elston & Grizzle, 1962 e Kenward, 1987) de que as correlações observadas, em geral, diminuem à medida em que t se distancia de t_s . Isto pode ser uma indicação de que o modelo de efeitos aleatórios com IC, que é o modelo corrente na prática, não seja adequado à maioria dos conjuntos de dados longitudinais, colocando em dúvida a hipótese subjacente simples de que as observações sejam condicionalmente independentes. A hipótese alternativa é composta por outros modelos de processos estocásticos que possam gerar observações condicionalmente dependentes, sendo o MB um caso particular.

3.3 Modelo de efeitos aleatórios com Movimento Browniano

Nesta Seção vamos caracterizar a matriz de dispersão induzida por um modelo de retas com coeficientes aleatórios e com a matriz de dispersão intra-individual dada por um MB com origem igual a zero no eixo dos tempos. $V_s(t)$ pode então ser especificada por

$$V_s(t) = \sigma_0^2 - \sigma_{01}(\sigma_{01} + \sigma^2)/\sigma_1^2 + \begin{cases} \sigma_1^2[t_s + (\sigma_{01} + \sigma^2)/\sigma_1^2](t + \sigma_{01}/\sigma_1^2), & t \leq t_s \\ \sigma_1^2[t + (\sigma_{01} + \sigma^2)/\sigma_1^2](t_s + \sigma_{01}/\sigma_1^2), & t > t_s \end{cases} \quad (3.3.1)$$

Para t_s fixo, trata-se de uma função descrita por duas retas, uma passando por

$$(-\sigma_{01}/\sigma_1^2, \sigma_0^2 - \sigma_{01}(\sigma_{01} - \sigma^2)/\sigma_1^2),$$

e outra por

$$(-(\sigma_{01} + \sigma^2)/\sigma_1^2, \sigma_0^2 - \sigma_{01}(\sigma_{01} - \sigma^2)/\sigma_1^2)$$

(estes pontos estão indicados na Figura 3.4 por $\bullet$) que são concorrentes quando $t = t_s$, situação em que seu valor representa a variância marginal $V(t)$. Esta é uma parábola com mínimo em

$$t = -(2\sigma_{01} + \sigma^2)/(2\sigma_1^2).$$

Figura 3.4: $V(t)$ e $V_s(t)$, sob o modelo com MB.

Os valores de t correspondentes a $-(\sigma_{01} + \sigma^2)/\sigma_1^2$, a $-(2\sigma_{01} + \sigma^2)/(2\sigma_1^2)$ e a $-\sigma_{01}/\sigma_1^2$ estão assinalados na Figura 3.4. $V_s(t) < 0$ quando

$$t < -(\sigma_0^2 + \sigma_{01}t_s)/(\sigma_{01} + \sigma_1^2t_s + \sigma^2) \quad \text{e} \quad t_s > \max\{t, -(\sigma_{01} + \sigma^2)/\sigma_1^2\}$$

ou

$$t > -(\sigma_0^2 + (\sigma_{01} + \sigma^2)t_s)/(\sigma_{01} + \sigma_1^2t_s) \quad \text{e} \quad t_s < \min\{t, -\sigma_{01}/\sigma_1^2\}.$$

Caso contrário, $V_s(t) > 0$. Caso o domínio da função se restrinja a valores não negativos, V_s só pode ser negativo se $t_s > -\sigma_0^2/\sigma_{01}$ com $\sigma_{01} < 0$.

Para ilustração, as curvas da Figura 3.4 foram construídas com $\sigma_0^2 = 8$, $\sigma_1^2 = 1 = \sigma^2$, $\sigma_{01} = -2$, $t_{s1} = 0.5$, $t_{s2} = 2.5$ e $t_{s3} = 4.5$. As funções V_{s1} e V_{s2} para as quais t_{s1} e t_{s2} são equidistantes de $-(2\sigma_{01} + \sigma^2)/(2\sigma_1^2)$, são simétricas em relação à reta $t = -(2\sigma_{01} + \sigma^2)/(2\sigma_1^2)$.

A função (3.1.4) é dada por

$$\rho_s(t) = \frac{\sigma_0^2 + \sigma_{01}(t_s + t) + \sigma_1^2t_st + \sigma^2\min(t_s, t)}{\sqrt{V(t_s)(\sigma_0^2 + (\sigma^2 + 2\sigma_{01})t + \sigma_1^2t^2)}} \quad (3.3.2)$$

que converge para o valor com sinal positivo em (3.2.3) quando $t \rightarrow \infty$. A

Figura 3.5: $\rho_s(t)$ sob o modelo com MB, com $\sigma_{01} < 0$.

Figura 3.5 mostra as curvas $\rho_{s_i}(t)$, correspondentes às curvas $V_{s_i}(t)$ da Figura 3.4, sendo as funções ρ_{s_1} e ρ_{s_2} simétricas em relação a

$$t = -(2\sigma_{01} + \sigma^2)/(2\sigma_1^2).$$

Como

$$\begin{aligned} \dot{V}_s(t) &= \sigma_{01} + \sigma_1^2 t_s + \begin{cases} \sigma^2 & \text{se } t < t_s \\ 0 & \text{caso contrário} \end{cases} \\ \dot{V}(t) &= 2(\sigma_{01} + \sigma_1^2 t) + \sigma^2 \end{aligned}$$

temos que o segundo membro de (3.1.6) é

$$A(\sigma_{01}, t_s) + B(\sigma_{01}, t_s)t \quad (3.3.3)$$

onde

$$\begin{aligned} A(\sigma_{01}, t_s) &= \begin{cases} \sigma_0^2 \sigma^2 + a(\sigma_{01})t_s & \text{se } t < t_s \\ -\sigma_0^2 \sigma^2 - b(\sigma_{01})t_s & \text{caso contrário} \end{cases} \\ B(\sigma_{01}, t_s) &= \begin{cases} [b(\sigma_{01}) + \sigma^2 \sigma_1^2 t_s] & \text{se } t < t_s \\ -[a(\sigma_{01}) + \sigma^2 \sigma_1^2 t_s] & \text{caso contrário} \end{cases} \end{aligned}$$

com

$$\begin{aligned} a(\sigma_{01}) &= 2(\sigma_0^2\sigma_1^2 - \sigma_{01}^2) - \sigma^2\sigma_{01} \\ b(\sigma_{01}) &= \sigma^4 + 3\sigma_{01}\sigma^2 - 2(\sigma_0^2\sigma_1^2 - \sigma_{01}^2) \end{aligned}$$

podendo-se avaliar o comportamento de $\rho_s(t)$ ($t \geq 0$) em relação à variação de t_s (≥ 0) e de σ_{01} com os parâmetros positivos σ_0 , σ_1 e σ^2 fixos. A função $a(\sigma_{01})$ é quadrática, tendo apenas uma das raízes,

$$(\sqrt{\sigma^4 + 16\sigma_0^2\sigma_1^2} - \sigma^2)/4,$$

satisfazendo $|\sigma_{01}| \leq \sigma_0\sigma_1$. Assim, quando

$$\sigma_{01} \leq (\sqrt{\sigma^4 + 16\sigma_0^2\sigma_1^2} - \sigma^2)/4,$$

temos

- (i) $A > 0$ se $t \leq t_s$
- (ii) $B < 0$ caso contrário.

Caso contrário, temos

- (iii) $A > 0$ se $t \leq t_s < \sigma_0^2\sigma^2/[\sigma^2\sigma_{01} - 2(\sigma_0^2\sigma_1^2 - \sigma_{01}^2)]$
- (iv) $A < 0$ se $t_s > \max\{t, \sigma_0^2\sigma^2/[\sigma^2\sigma_{01} - 2(\sigma_0^2\sigma_1^2 - \sigma_{01}^2)]\}$
- (v) $B > 0$ se $t_s < \min\{t, [\sigma^2\sigma_{01} - 2(\sigma_0^2\sigma_1^2 - \sigma_{01}^2)]/(\sigma_1^2\sigma^2)\}$
- (vi) $B < 0$ se $t > t_s > [\sigma^2\sigma_{01} - 2(\sigma_0^2\sigma_1^2 - \sigma_{01}^2)]/(\sigma_1^2\sigma^2)$.

A função $b(\sigma_{01})$ também é quadrática e com apenas uma raiz viável,

$$(\sqrt{\sigma^4 + 16\sigma_0^2\sigma_1^2} - 3\sigma^2)/4,$$

sob a condição de que $\sigma^2 < 3\sigma_0\sigma_1$. Assim, quando

$$\sigma_{01} \leq (\sqrt{\sigma^4 + 16\sigma_0^2\sigma_1^2} - 3\sigma^2)/4,$$

temos

- (vii) $B < 0$ se $t \leq t_s < [2(\sigma_0^2\sigma_1^2 - \sigma_{01}^2) - 3\sigma^2\sigma_{01} - \sigma^4]/(\sigma^2\sigma_1^2)$
- (viii) $B > 0$ se $t_s > \max\{t, [2(\sigma_0^2\sigma_1^2 - \sigma_{01}^2) - 3\sigma^2\sigma_{01} - \sigma^4]/(\sigma^2\sigma_1^2)\}$
- (ix) $A < 0$ se $t_s < \min\{t, \sigma^2\sigma_0^2/[2(\sigma_0^2\sigma_1^2 - \sigma_{01}^2) - 3\sigma^2\sigma_{01} - \sigma^4]\}$
- (x) $A > 0$ se $t > t_s > \sigma^2\sigma_0^2/[2(\sigma_0^2\sigma_1^2 - \sigma_{01}^2) - 3\sigma^2\sigma_{01} - \sigma^4]$.

Caso contrário, temos

- (xi) $B > 0$ se $t \leq t_s$
- (xii) $A < 0$ se $t > t_s$.

Quando A e B são ambos positivos (negativos) pode-se afirmar que $\rho_s(t)$ é (de)crescente. Quando

$$\begin{aligned}\sigma_{01} &\leq (\sqrt{\sigma^4 + 16\sigma_0^2\sigma_1^2} - 3\sigma^2)/4 \\ (\sigma_{01} &\geq (\sqrt{\sigma^4 + 16\sigma_0^2\sigma_1^2} - \sigma^2)/4),\end{aligned}$$

e

$$\begin{aligned}t &\leq t_s < [2(\sigma_0^2\sigma_1^2 - \sigma_{01}^2) - 3\sigma^2\sigma_{01} - \sigma^4]/(\sigma^2\sigma_1^2) \\ (t_s &> \max\{t, \sigma_0^2\sigma^2/[\sigma^2\sigma_{01} - 2(\sigma_0^2\sigma_1^2 - \sigma_{01}^2)]\})\end{aligned}$$

ou

$$\begin{aligned}t &\geq t_s > \sigma^2\sigma_0^2/[2(\sigma_0^2\sigma_1^2 - \sigma_{01}^2) - 3\sigma^2\sigma_{01} - \sigma^4] \\ (t_s &< \min\{t, [\sigma_{01}\sigma^2 - 2(\sigma_0^2\sigma_1^2 - \sigma_{01}^2)]/(\sigma^2\sigma_1^2)\})\end{aligned}$$

verifica-se que $A > (<) 0$ e $B < (>) 0$, indicando que $\rho_s(t)$ pode apresentar um ponto de máximo (mínimo) quando $t = -A/B$. Mas ocorre que

$$A + Bt_s \begin{cases} > 0 & \text{quando } t < t_s \\ < 0 & \text{quando } t > t_s \end{cases} \quad (3.3.4)$$

pois ambas as desigualdades são equivalentes a

$$\sigma_0^2 + (2\sigma_{01} + \sigma^2)t_s + \sigma_1^2 t_s^2 > 0. \quad (3.3.5)$$

Quando $\sigma_{01} \geq (\sqrt{\sigma^4 + 16\sigma_0^2\sigma_1^2} - \sigma^2)/4$, esta inequação é satisfeita pois basta $\sigma_{01} > 0$. O discriminante desta função quadrática em t_s ,

$$\Delta = 4\sigma_{01}^2 + 4\sigma_{01}\sigma^2 + \sigma^4 - 4\sigma_0^2\sigma_1^2 \quad (3.3.6)$$

é negativo quando $\sigma_{01} \in (-\sigma_0\sigma_1, (\sqrt{\sigma^4 + 16\sigma_0^2\sigma_1^2} - 3\sigma^2)/4)$, pois suas raízes estão fora deste intervalo, confirmando as desigualdades (3.3.5) e (3.3.4), o que impede a ocorrência de máximo local.

Em resumo (ver Figura 3.6), $\rho_s(t)$ cresce monòtonamente até t_s e depois decresce monòtonamente com valores positivos, exceto quando:

(I) $t_s < \sigma_0/\sigma_1$ e $\sigma_{01} > (\sqrt{\sigma^4 + 16\sigma_0^2\sigma_1^2 + 8\sigma^2\sigma_1^2 t_s} - \sigma^2)/4$, quando apresenta um mínimo após t_s , ou

(II) $t_s < \sigma_0/\sigma_1$ e $\sigma_{01} < -\sigma_1^2 t_s$, quando tem valores negativos para

$$t > -[\sigma_0^2 + (\sigma^2 + \sigma_{01})t_s]/(\sigma_{01} + \sigma_1^2 t_s), \text{ ou}$$

(III) $t_s > \sigma_0/\sigma_1$ e $\sigma_{01} > (\sqrt{\sigma^4 + 16\sigma_0^2\sigma_1^2 + 8\sigma^2\sigma_0^2/t_s} - \sigma^2)/4$, quando apresenta um mínimo antes de t_s , ou

(IV) $t_s > \sigma_0/\sigma_1$ e $\sigma_{01} < -\sigma_0^2/t_s$, quando tem valores negativos para

$$t < -(\sigma_0^2 + \sigma_{01}t_s)/(\sigma^2 + \sigma_{01} + \sigma_1^2 t_s).$$

A Figura 3.7 mostra os possíveis comportamentos de $\rho_s(t)$ quando $\sigma_{01} > 0$. Para ilustração, as Figuras 3.6 e 3.7 foram construídas usando $\sigma_0^2 = 4$, $\sigma_1^2 = 1$, $\sigma_{01} = 1.5$, $\sigma^2 = 6$, $t_{s1} = 4$ e $t_{s2} = 20$.

Verifica-se assim que as correlações ao longo de uma linha da matriz de dispersão teórica do modelo de retas com os coeficientes aleatórios e erro intra-individual dado por um MB apresentam um único máximo que ocorre sobre a diagonal principal, o que está de acordo com as observações de Elston & Grizzle (1962) e Kenward (1987).

Figura 3.6: Comportamento de $\rho_s(t)$ em função de (σ_{01}, t_s) .

Figura 3.7: $\rho_s(t)$ sob o modelo com MB, quando $\sigma_{01} > 0$.

3.4 Identificação da dispersão condicional

Nesta Seção utilizaremos análises gráficas das linhas da matriz de dispersão amostral para identificar um modelo de retas com coeficientes aleatórios e dispersão intra-individual com IC ou dada por um MB. A Tabela 3.1 mostra a matriz de dispersão amostral dos dados do Exemplo 2. A Figura 3.8 mostra os elementos de cada linha desta matriz de dispersão amostral em função da idade. Os pontos que representam as variâncias estão unidos por segmentos de reta, mostrando que $V(t)$ é crescente. O modelo com IC exige que os pontos que representam as covariâncias de uma mesma linha da matriz de dispersão estejam sobre uma mesma reta; o ponto que representa a variância deve "flutuar" acima dela a uma distância constante em todas as retas. Na Figura 3.8 não é difícil de se imaginar uma reta passando pelos tres pontos que representam as covariâncias de uma mesma linha de $\mathbf{V}$, mas o ponto que representa a variância respectiva nem sempre se situa acima desta reta. Por outro lado, o modelo com MB exige que as covariâncias de uma mesma linha até o elemento da diagonal (que é a variância) inclusive estejam sobre uma mesma reta, e do elemento da diagonal (inclusive) em diante, sobre uma outra reta. Esta conformação teórica parece mais aceitável neste conjunto de dados.

Tabela 3.1: Dispersão amostral dos dados do Exemplo 2.

Idade	0	1	2	3
0	6.32997			
1	6.18908	6.44934		
2	5.77700	6.15342	6.91800	
3	5.54816	5.92342	6.94631	7.46474

Outra forma de identificar um dos dois modelos é através da análise dos incrementos verificados em cada coluna da matriz de dispersão amostral em relação à coluna anterior (ver Tabela 3.2). A função V_s (3.2.1 e 3.3.1) com o argumento t_u é linear em t_u , e sua derivada, sendo independente de t_u , pode ser vista como uma função só de t_s . Na prática, podem-se obter estimativas dos incrementos médios de $V_s(t_u)$ a partir das diferenças entre as colunas adjacentes u e $u - 1$ da matriz de dispersão amostral das observações. Sob o modelo com IC,

Figura 3.8: Covariâncias amostrais e $V(t)$ empírico.

$$\begin{aligned} \Delta V_s(t_u)/\Delta t_u &= (V_s(t_u) - V_s(t_{u-1}))/(\Delta t_u) \\ &= \sigma_{01} + \sigma_1^2 t_s \begin{cases} +0 & \text{se } u \neq s \neq u-1 \\ +\sigma^2/\Delta t_u & \text{se } u = s \\ -\sigma^2/\Delta t_u & \text{se } u-1 = s \end{cases} \quad (3.4.1) \end{aligned}$$

que são retas paralelas (tres, no caso de Δt_u constante). Mas sob o modelo com MB,

$$\Delta V_s(t_u)/\Delta t_u = \sigma_{01} + \sigma_1^2 t_s + \begin{cases} 0 & \text{se } t_{u-1} \geq t_s \\ \sigma^2 & \text{se } t_u \leq t_s \end{cases} \quad (3.4.2)$$

são apenas duas retas paralelas.

A Figura 3.9 mostra os incrementos calculados na Tabela 3.2 em função da idade que aparece na entrada das linhas. O coeficiente angular das retas correspondentes ao modelo com IC (3.4.1) são praticamente determinados pelos 6 pontos representados por $\cdot$, correspondentes à primeira reta de (3.4.1). Os pontos representados por $\bullet$ determinam a segunda reta de (3.4.1), e os pontos representados por $\circ$ a terceira. A separação esperada entre os tres conjuntos de pontos, sob o modelo com IC, não é verificada. Por outro lado,

sob o modelo com MB, os 6 pontos inferiores da Figura 3.10 representados por $\circ$ correspondem à primeira reta de (3.4.2), e os 6 pontos superiores representados por $\bullet$ à segunda reta, mostrando ser este modelo mais aceitável que o primeiro. Assim, pelos dois critérios apresentados, o modelo identificado é o modelo com MB.

Tabela 3.2: Incrementos observados em cada coluna da matriz de dispersão amostral do Exemplo 2.

Idade	1	2	3
0	-0.14088	-0.41208	-0.22884
1	0.26025	-0.29592	-0.23000
2	0.37642	0.76458	0.02832
3	0.37526	1.02290	0.51842

As tres retas paralelas que se ajustariam aos pontos na Figura 3.9 são mais inclinadas que as duas retas paralelas que se ajustariam aos pontos na Figura 3.10. Mas o coeficiente angular teórico, em ambos os casos, é igual a σ_1^2 , de onde se conclui que sua estimativa sob o modelo com IC deva ser

Figura 3.10: Identificação de um modelo com MB.

maior que a estimativa sob o modelo com MB. Por outro lado, a distância entre quaisquer duas retas paralelas adjacentes que se ajustariam aos pontos da Figura 3.9 é menor que a distância entre as duas retas paralelas que se ajustariam aos pontos da Figura 3.10. Quando $\Delta t_u = 1$ como neste exemplo, os valores teóricos destas distâncias correspondem ao parâmetro σ^2 em ambos os modelos, concluindo-se que a estimativa deste parâmetro sob o modelo com IC deva ser menor que a estimativa sob o modelo com MB, atuando de forma oposta à verificada com σ_1^2 . Isto mostra como os diferentes modelos repartem de maneira diversa a dispersão total $\mathbf{V}$ entre a dispersão inter-indivíduos, $\mathbf{Z}\mathbf{I}\mathbf{Z}^t$ e a dispersão intra-individual $\sigma^2\mathbf{R}$. Assim, quando o modelo da dispersão intra-individual é inadequado aos dados, a estimativa de σ_1^2 fica viciada.

3.5 Comparação entre os modelos

Nesta Seção vamos ajustar os dois modelos considerados neste Capítulo aos dados do Exemplo 2, para verificar a validade dos métodos de identificação propostos na Seção anterior. Utilizamos o programa BMDP-5V com a opção de estimação por MVR e constatamos (i) que o programa não fornece os

desvios padrões assintóticos dos parâmetros $\theta = (\sigma_0^2, \sigma_{01}, \sigma_1^2, \sigma^2)$ e (ii) que a matriz de dispersão residual é calculada com a divisão da soma dos quadrados residuais por N , sendo portanto viciada. Apresentamos aqui a dispersão residual corrigida dividindo a soma dos quadrados por $N - 1$. Consideramos também um modelo com a mesma estrutura de efeitos fixos, mas com estrutura de covariâncias livre (modelo 0), em relação ao qual os outros modelos ficam encaixados, como uma referência de comparação dos modelos no caso de estimação por MVR. O AIC que apresentamos é o mesmo que é calculado pelo BMDP-5V.

Tabela 3.3: Parâmetros fixos estimados. Entre parênteses os desvios padrões.

Parâmetro	Modelo		
	0	IC	MB
intercepto	48.6534	48.6755	48.6550
d.p.	(0.5598)	(0.5696)	(0.5626)
d.p. robusto	(0.5598)	(0.5552)	(0.5626)
coef. angular	0.9308	0.9330	0.9317
d.p.	(0.1045)	(0.1303)	(0.1224)
d.p. robusto	(0.1045)	(0.1270)	(0.1224)

As estimativas dos parâmetros fixos se encontram na Tabela 3.3. Como era de se esperar, não há diferença marcante entre as estimativas dos efeitos fixos sob os tres modelos, pois sob a correta especificação dos seus valores esperados, os estimadores são não viciados. As estimativas dos desvios padrões assintóticos usuais sob o modelo com IC são um pouco maiores que os robustos, mas sob os outros modelos elas são coincidentes. No modelo 0, isto se deve à estimação da matriz de dispersão marginal pela dispersão residual, mas no modelo com MB isto sugere um bom ajuste do modelo aos dados do Exemplo 2. O desvio padrão do coeficiente angular é, no entanto, menor sob o modelo 0, sugerindo falta de eficiência na estimação deste parâmetro sob o modelo com MB, em relação ao modelo anterior.

Tabela 3.4: Estimativas dos parâmetros da dispersão marginal.

Modelo	Parâmetros			
	σ_0^2	σ_{01}	σ_1^2	σ^2
IC	6.35379	-0.27144	0.30087	0.19345
MB	6.32997	-0.26060	0.13248	0.50200

As estimativas dos parâmetros do vetor θ se encontram na Tabela 3.4 para os modelos com IC e MB. Como destacamos na Seção anterior, as estimativas de σ_1^2 e de σ^2 diferem bastante sob os dois modelos, enquanto aquelas dos demais parâmetros são semelhantes.

As matrizes de dispersão marginal, estimada para o modelo 0, e induzidas pelos outros modelos se encontram nas Tabelas 3.5 (modelo 0), 3.6 (IC) e 3.7 (MB); as matrizes de dispersão residual correspondentes nas Tabelas 3.8 (IC) e 3.9 (MB) e as medidas de comparação entre os modelos na Tabela 3.10.

Tabela 3.5: Dispersão marginal estimada para o modelo 0.

Idade	0	1	2	3
0	6.32998			
1	6.18915	6.45110		
2	5.77709	6.15578	6.92119	
3	5.54816	5.92360	6.94656	7.46475

Tabela 3.6: Dispersão marginal induzida pelo modelo com IC.

Idade	0	1	2	3
0	6.54720			
1	6.08235	6.30524		
2	5.81092	6.14122	6.66498	
3	5.53948	6.17066	6.80184	7.62646

Tabela 3.7: Dispersão marginal induzida pelo modelo com MB.

Idade	0	1	2	3
0	6.32997			
1	6.06937	6.44325		
2	5.80877	6.31512	6.82148	
3	5.54816	6.18700	6.82584	7.46467

Tabela 3.8: Matriz de dispersão residual (IC).

Idade	0	1	2	3
0	6.33042			
1	6.18872	6.44963		
2	5.77638	6.15392	6.91885	
3	5.54869	5.92300	6.94558	7.46537

Tabela 3.9: Matriz de dispersão residual (MB).

Idade	0	1	2	3
0	6.32997			
1	6.18908	6.45089		
2	5.77700	6.15551	6.92081	
3	5.54816	5.92342	6.94632	7.46474

Tabela 3.10: Medidas de ajuste dos modelos.

Medida	Modelo		
	0	IC	MB
$\log \hat{\mathbf{V}} $	0.07662	0.44978	0.32456
$\log \mathbf{G} $	0.07662	0.07675	0.07663
L_v	0	0.37303	0.24793
$\log(L)$	-113.270	-117.801	-116.617
AIC	-123.270	-123.801	-121.617

A variância generalizada dos resíduos marginais dos tres modelos têm valores muito próximos em termos dos seus logaritmos, devido à já mencionada proximidade dos efeitos fixos estimados. Seus valores são também comparáveis ao da variância amostral (0.06611), o que sugere uma boa escolha do modelo de efeitos fixos para os dados do Exemplo 2. L_v é menor para o modelo com MB do que para o modelo com IC, sugerindo melhor ajuste do primeiro, mas o teste da razão de verossimilhança entre o modelo 0 e qualquer um dos outros dois modelos não é significativo ($p > 0.10$). Pelo AIC, entretanto, o modelo com MB é o melhor dentre os tres.

3.6 Conclusão e discussão

O modelo de efeitos aleatórios com independência condicional exibe um máximo local fora da diagonal principal em cada linha da matriz de correlação marginal, o que contradiz a observação usual de que a correlação decresce à medida em que a distância entre os instantes de observações repetidas aumenta. Este padrão de comportamento, no entanto, pode ser obtido através do modelo de efeitos aleatórios com MB. A independência condicional pode ser uma hipótese demasiadamente simplificadora da realidade que dificilmente seja comprovada na prática, ao contrário da dispersão condicional descrita por meio de um processo estocástico.

Modelos diferentes para a dispersão intra-individual podem resultar em estimativas diferentes da variância do coeficiente angular σ_1^2 podendo-se dizer que um modelo incorreto produz vício na sua estimação .

O conhecimento da dispersão marginal induzida por vários modelos de efeitos aleatórios facilita a identificação de um modelo para um conjunto de dados balanceados com relação ao tempo. Sugerimos, portanto, que se pesquisem as dispersões marginais de modelos com efeitos aleatórios com outras dispersões intra-individuais, tanto analiticamente quanto através de simulações nos modelos analiticamente intratáveis. Por outro lado, a formulação de uma dispersão intra-individual adequada pode requerer conhecimentos adicionais sobre modelos estocásticos.

As técnicas de identificação descritas neste Capítulo podem ser estendidas para outros modelos. No Capítulo 4, examinamos a estrutura de covariâncias induzidas por um modelo de efeitos aleatórios com erro condicional dado por um outro processo estocástico.

Capítulo 4

Uma aplicação do modelo normal truncado.

4.1 Introdução

Neste Capítulo, analisamos o conjunto de dados apresentado no Exemplo 3. Para efeito de comparação entre tratamentos, restringimo-nos a análise dos valores médios (y) dos diâmetros das feridas dos dois olhos de cada coelho, já que eles foram submetidos ao mesmo tratamento. Isto simplifica a análise, pois torna o modelo univariado, mas introduz uma censura nas observações finais em que se medem os tamanhos da ferida apenas de um dos olhos, devido à recuperação total da outra ferida, cujo valor então é arbitrado como sendo zero e mantido constante. Em vez de modelar esta mudança de comportamento da curva de recuperação de um dos olhos, é mais fácil considerar estas médias finais como sendo censuradas.

A Figura 4.1 mostra os diagramas paralelos de dispersão de y , podendo-se notar que o processo de cicatrização medido pelo diâmetro da ferida é basicamente linear no tempo, com o coeficiente angular (velocidade de cicatrização) variando entre as unidades amostrais. Como as medidas foram realizadas até completa cicatrização das feridas, os dados se mostram truncados no final do processo, quando são realizadas as últimas medidas de y , antes de se anularem. A ligeira não linearidade em relação ao tempo, mostrada por algumas unidades amostrais no final do processo, pode ser uma consequência do truncamento. A velocidade parece ser menor no início do processo em algumas unidades amostrais, sugerindo haver uma fase de aceleração.

Na Seção 4.2 descrevemos um modelo de efeitos aleatórios em três estágios (Gamerman & Migon, 1993) identificado nas análises apresentadas nas Seções 4.3 e 4.4. Na Seção 4.3 utilizamos os métodos apresentados nos Capítulos anteriores para identificar uma mistura de realizações de processos de OU

Figura 4.1: Diagramas paralelos de dispersão - Exemplo 3.

integrado ou de posição descrito por (1.2.12) a (1.2.14) com realizações de MB (caso particular do processo anterior quando $\lambda \rightarrow 0$) com efeitos aleatórios gaussianos nos parâmetros lineares. A mistura destes processos é modelada no terceiro estágio através de um efeito aleatório dicotômico no parâmetro não linear λ do modelo de OU integrado. A necessidade de se considerar uma distribuição normal multivariada truncada (as observações em cada instante são estritamente positivas) torna a estimação deste modelo algebricamente intratável. Consideramos então um modelo simplificado com a estrutura de erros condicionais descrito apenas por um MB. Análises individuais com base neste modelo são apresentadas na Seção 4.4, onde constatamos a mistura dos dois modelos. Outro aspecto que exploramos é a possibilidade de se identificar uma reparametrização do modelo para estabilizar as variâncias das estimativas individuais dos desvios padrões das observações. Em particular, a reparametrização $\log \sigma_j$ mostrou ser útil para detectar diferenças entre tratamentos. Para estimar os parâmetros populacionais do modelo de efeitos aleatórios descrito acima, utilizamos na Seção 4.5 a mesma metodologia de Hui & Berger (1983) (ver Seção 1.3), que consiste em modelar as estimativas obtidas nas análises individuais.

4.2 Descrição do modelo

O modelo que consideramos neste Capítulo tem o seu primeiro estágio descrito por (1.2.12), conhecidos os parâmetros y_0 , β e $\lambda = \exp(-\alpha)$. Para a unidade amostral j , estes parâmetros são correspondentes a β_{0j} , β_{1j} (que compõem o vetor β_j) e λ_j . Dados estes parâmetros, o vetor de observações y_j é modelado no primeiro estágio por

$$y_j | (\lambda_j, \beta_j) \sim N(\mathbf{X}_{j\lambda_j} \beta_j, \text{Var}(\kappa_{j\lambda_j})) \quad (4.2.1)$$

onde o valor esperado tem a matriz de especificação dada por

$$\mathbf{X}_{j\lambda_j} = (\mathbf{1}, -\mathbf{x}_{j\lambda_j})$$

com $\mathbf{x}_{j\lambda_j} = \mathbf{t}_j + (\mathbf{1} - \lambda_j^{\mathbf{t}_j}) / \log \lambda_j$ em que $\lambda_j^{\mathbf{t}_j}$ é uma coluna onde cada elemento é dado pela potenciação de λ_j pelo elemento correspondente de $\mathbf{t}_j$, e $\kappa_{j\lambda_j}$ é o vetor de erros condicionais, cujo elemento correspondente ao instante t é descrito por (1.2.13) com $\exp(-\alpha)$ substituído pelo valor individual correspondente, λ_j . Conseqüentemente, $\text{Var}(\kappa_{j\lambda_j})$ tem os elementos dados por (1.2.14) quando $\lambda_j = \lambda$. Quando $\lambda_j \rightarrow 0$ (4.2.1) uma equação de reta com os erros condicionais dados por um MB, em que β_{1j} representa a velocidade de cicatrização constante. Para finalidade de comparação entre tratamentos, consideramos a velocidade média $\bar{\beta}_{1j} = \beta_{0j}/t_{0j}$ para as unidades amostrais com $\lambda_j \neq 0$, onde t_{0j} é obtido de $E[y_{jt_{0j}} | (\beta_j, \lambda_j)] = 0$, com $y_{jt_{0j}}$ sendo o valor correspondente ao instante t_{0j} da unidade amostral j , resultando que

$$t_{0j} = \beta_{0j}/\beta_{1j} - (1 - \lambda_j^{t_{0j}}) / \log \lambda_j. \quad (4.2.2)$$

Notando que $\bar{\beta}_{1j} = \beta_{1j}$ quando $\lambda_j = 0$ e definindo $\bar{\beta}_j = (\beta_{0j}, \bar{\beta}_{1j})^t$, consideramos no segundo estágio que

$$\bar{\beta}_j | \lambda_j \sim N(\bar{\beta}, \mathbf{\Gamma}_0) \quad (4.2.3)$$

onde $\bar{\beta} = (\beta_0, \bar{\beta}_1)^t$ depende do tratamento e $\mathbf{\Gamma}_0$ é comum aos tratamentos, sendo as variâncias denotadas por σ_0^2 e $\bar{\sigma}_1^2$ e a covariância por $\bar{\sigma}_{01}$, o que determina a distribuição de $\beta_j = (\beta_{0j}, \beta_{1j})^t$ condicionada a λ_j .

Assim,

$$\beta_j | (\lambda_j = 0) \sim N(\bar{\beta}, \mathbf{\Gamma}_0). \quad (4.2.4)$$

Mas quando $\lambda_j = \lambda$, a normalidade de $\bar{\beta}_{1j}$ implica a não normalidade de $\beta_{1j}|(\lambda_j = \lambda)$ definido por

$$\beta_{1j}|(\lambda_j = \lambda) = \beta_{0j}/[t_{0j} + (1 - \lambda^{t_{0j}})/\log \lambda], \quad (4.2.5)$$

a partir de (4.2.2). Denotamos seu valor esperado por β_1 . Utilizando o Método Delta, temos (ver Apêndice B)

$$\begin{aligned} \sigma_{01} &= Cov(\beta_{0j}, \beta_{1j}|\lambda_j = \lambda) \cong \\ &\quad \{\sigma_0^2[t_0\lambda^{t_0} + (1 - \lambda^{t_0})/\log \lambda] + \bar{\sigma}_{01}t_0^2(1 - \lambda^{t_0})\}/[t_0 + (1 - \lambda^{t_0})/\log \lambda]^2 \end{aligned} \quad (4.2.6)$$

$$\begin{aligned} \sigma_1^2 &= Var(\beta_{1j}|\lambda_j = \lambda) \cong \\ &\quad \{\sigma_0^2[t_0\lambda^{t_0} + (1 - \lambda^{t_0})/\log \lambda]^2 + \bar{\sigma}_1^2t_0^4(1 - \lambda^{t_0})^2 + \\ &\quad 2\bar{\sigma}_{01}[t_0\lambda^{t_0} + (1 - \lambda^{t_0})/\log \lambda]t_0^2(1 - \lambda^{t_0})\}[t_0 + (1 - \lambda^{t_0})/\log \lambda]^4 \end{aligned} \quad (4.2.7)$$

onde $t_0 = \beta_0/\bar{\beta}_1$ depende do tratamento, resultando na dispersão Γ_λ de $\beta_j|(\lambda_j = \lambda)$.

Assim, a dispersão dos valores observados condicionados apenas a λ_j resulta em

$$\mathbf{X}_{j\lambda_j}\Gamma_{\lambda_j}\mathbf{X}_{j\lambda_j}^t + \mathbf{Var}(\kappa_{j\lambda_j})$$

onde Γ_{λ_j} é a dispersão de $\beta_j|\lambda_j$ e $\mathbf{Var}(\kappa_{j\lambda_j})$ depende de λ_j .

Ao considerar o terceiro estágio do modelo dado por

$$\lambda_j \sim \lambda Bernoulli(p), \quad (4.2.8)$$

onde p é a proporção de casos com $\lambda_j \neq 0$, a distribuição marginal de $\mathbf{y}_j$ resulta na distribuição da mistura de dois processos, com o valor esperado dado por $p\mathbf{X}_{j\lambda}\boldsymbol{\beta} + (1 - p)\mathbf{X}_{j0}\bar{\boldsymbol{\beta}}$ onde $\boldsymbol{\beta} = \mathbf{E}[\boldsymbol{\beta}_j|(\lambda_j = \lambda)]$ e a matriz de dispersão por

$$\mathbf{E}[\mathbf{Var}(\mathbf{y}_j|\lambda_j)] + \mathbf{Var}[\mathbf{E}(\mathbf{y}_j|\lambda_j)] \quad (4.2.9)$$

onde

$$\begin{aligned} \mathbf{E}[\mathbf{Var}(\mathbf{y}_j|\lambda_j)] &= p\{\mathbf{X}_{j\lambda}\mathbf{\Gamma}_\lambda\mathbf{X}_{j\lambda}^t + \mathbf{Var}(\boldsymbol{\kappa}_{j\lambda})\} + \\ &\quad (1-p)\{\mathbf{X}_{j0}\mathbf{\Gamma}_0\mathbf{X}_{j0}^t + \mathbf{Var}(\boldsymbol{\kappa}_{j0})\} \end{aligned} \quad (4.2. 10)$$

$$\mathbf{Var}[\mathbf{E}(\mathbf{y}_j|\lambda_j)] = p(1-p)(\mathbf{X}_{j\lambda}\boldsymbol{\beta} - \mathbf{X}_{j0}\bar{\boldsymbol{\beta}})(\mathbf{X}_{j\lambda}\boldsymbol{\beta} - \mathbf{X}_{j0}\bar{\boldsymbol{\beta}})^t \quad (4.2. 11)$$

(ver detalhes no Apêndice C).

4.3 Identificação do modelo

Apesar de haver pequenas diferenças nos tempos de medida, eles estão razoavelmente próximos em cada sessão de avaliação das feridas, fazendo sentido considerá-los comuns para se calcularem as médias e dispersões amostrais entre os valores iniciais antes de se observarem os truncamentos. A Tabela 4.1 mostra os valores médios de y por tratamento em função dos tempos médios. Inspeção detalhada das Tabelas 1.3, a 1.6 mostra que a quinta avaliação parece ter sido realizada em duas sessões, tendo havido um intervalo médio

Tabela 4.1: Valores médios de y em função dos tempos médios.

Tempo médio (h)	Tratamento			
	1	2	3	4
0.000	46.341	45.356	46.032	48.851
11.862	40.743	39.198	39.251	41.189
19.206	36.108	33.431	33.421	34.277
28.645	30.511	25.674	28.362	29.049
34.823	-	-	24.082	22.474
37.165	25.757	20.976	21.767	-
41.419	22.008	15.936	18.988	17.096
47.399	18.625	11.821	13.961	11.757
54.957	14.955	7.598	-	-
59.434	<u>11.386</u>	-	-	-
66.872	<u>8.305</u>	-	-	-

de 2.34 horas entre elas, adiantando-se as avaliações das unidades amostrais do tratamento 4 e de uma parte (5 animais) do tratamento 3 em relação às outras. Os valores sublinhados indicam a existência de pelo menos um valor de y censurado, como foi explicado no Exemplo 3 e no início da Seção 4.1. A Figura 4.2 mostra os diagramas paralelos de dispersão dos valores médios de y por tratamento, onde podemos notar um comportamento linear, com uma pequena curvatura no início do processo.

Figura 4.2: Diagramas paralelos de dispersão de y médio por tratamento em função dos tempos médios.

O valor médio inicial da ferida do tratamento 4 difere dos valores dos outros tratamentos. A distribuição dos valores individuais das feridas iniciais y_{0j} é explorada na Figura 4.3 que mostra o diagrama de dispersão em função dos escores normais padrão, sugerindo uma mistura de duas normais com mesma variância, dado que o diagrama poderia ser sintetizado através de duas retas paralelas. Como a distribuição com a média maior representa basicamente as observações do tratamento 4, este gráfico é refeito na Figura 4.4 separadamente para as observações do tratamento 4 e as dos demais tratamentos, onde o paralelismo observado anteriormente fica menos evidente

Figura 4.3: Gráfico em papel de probabilidade normal de y_{0j} .

Figura 4.4: Gráficos em papel de probabilidade normal de y_{0j} para o tratamento 4 e para outros tratamentos.

mas ainda é aceitável. A observação 19, que aparenta ser um “outlier” em relação a outras observações do tratamento 3, parece ficar menos discrepante em relação a outras observações dos tratamentos 1 a 3. Estas constatações serão relevadas na formulação do modelo para o efeito aleatório na Seção 4.5.

Tabela 4.2: Matriz de dispersão amostral das observações iniciais de y .

tempo médio	00.000	11.862	19.206	28.645	41.419	47.399
00.000	3.4789					
11.862	3.1650	4.3402				
19.206	3.1231	5.3360	8.9834			
28.645	2.3714	5.4622	9.9532	13.3871		
41.419	1.9088	5.0753	10.4691	14.6259	19.5836	
47.399	1.1016	4.3120	9.8483	14.3584	20.0576	21.7723

Figura 4.5: Diagramas paralelos de dispersão das linhas da matriz de dispersão amostral de y .

A Tabela 4.2 mostra a matriz de dispersão amostral das sete primeiras observações, com exceção da quinta. As médias em cada instante de avaliação

foram calculadas por tratamento, exceto no instante inicial, em que foi calculada uma única média para os tres primeiros grupos, e uma outra para o grupo 4. Os diagramas paralelos de dispersão correspondentes às suas linhas estão na Figura 4.5. Os pontos correspondentes às variâncias parecem ser pontos de continuidade das curvas, que também não são retas concorrentes num mesmo ponto, descartando-se o modelo de efeitos aleatórios com IC. As curvas são ligeiramente sigmoidais até os pontos correspondentes às variâncias, e depois são razoavelmente lineares e até possivelmente concorrentes num ponto à esquerda do gráfico. As variâncias parecem se comportar de acordo com uma parábola com mínimo entre as duas avaliações iniciais. Com exceção da curvatura inicial, estas características apontam para um modelo de efeitos aleatórios com MB conforme a discussão apresentada no Capítulo 3.

A Tabela 4.3 mostra os incrementos por unidade de tempo das linhas da matriz de dispersão amostral da Tabela 4.2, cujos valores estão dispostos em função do tempo nas Figuras 4.6 e 4.7, a primeira para tentar identificar um modelo com IC e a segunda um modelo com MB. Na Figura 4.6, o modelo com IC se caracteriza por uma disposição dos pontos sugerindo uma curva média (pelos pontos representados por $\cdot$), curvas paralelas acima passando pelos pontos representados por $\bullet$ e curvas paralelas abaixo passando pelos pontos representados por $\circ$. Esta disposição não se confirma, pois apesar de os pontos correspondentes às curvas superiores se encontrarem em sua maioria acima dos pontos correspondentes às curvas inferiores, a maioria daqueles correspondentes à curva média está mais espalhada que os anteriores. Por outro lado, o modelo com MB pode ser identificado na Figura 4.7, em que existem dois grupos separados de pontos que poderiam ser modelados por duas curvas paralelas.

Tabela 4.3: Incrementos por unidade de tempo observados em cada coluna da matriz de dispersão amostral de y .

Tempo médio	11.862	19.206	28.645	41.419	47.399
00.000	-0.0265	-0.0057	-0.0796	-0.0362	-0.1350
11.862	0.0786	0.1038	0.0407	0.0117	-0.0146
19.206	0.1749	0.4958	0.1508	0.0908	0.0457
28.645	0.2673	0.6373	0.4566	0.1545	0.1089
41.419	0.2880	0.7749	0.5427	0.4695	0.2161
47.399	0.3102	0.8240	0.5827	0.5197	0.3957

Figura 4.6: Incrementos por unidade de tempo verificados em V_s em função de t_s . Identificação do modelo com IC.

Figura 4.7: Incrementos por unidade de tempo verificados em V_s em função de t_s . Identificação do modelo com MB.

O modelo de efeitos aleatórios com erros condicionais OU integrado é capaz de explicar a curvatura inicial apresentada tanto pelas curvas médias na Figura 4.2 quanto pelas linhas da matriz de dispersão amostral, como mostramos a seguir. A dispersão marginal resultante do modelo em tres estágios (4.2.1, 4.2.3 e 4.2.8) para y_j é dada por 4.2.9. Ao fixar o valor t_s correspondente à linha s desta matriz de dispersão teórica obtemos a função geradora da s -ésima linha (no espírito da discussão apresentada na Seção 2.4), definida para $t \geq 0$ por (ver detalhes no Apêndice C)

$$V_s(t) = a_s + b_s t + c_s(1 - \lambda^t)/\log \lambda + V_s^*(t) \quad (4.3.1)$$

onde

$$\begin{aligned} a_s &= \sigma_0^2 - p\sigma_{01}[t_s + (1 - \lambda^{t_s})/\log \lambda] - (1 - p)\bar{\sigma}_{01}t_s \\ b_s &= p\sigma_1^2[t_s + (1 - \lambda^{t_s})/\log \lambda] - [p\sigma_{01} + (1 - p)\bar{\sigma}_{01}] + (1 - p)\bar{\sigma}_1^2 t_s + \\ &\quad p(1 - p)(\beta_1 - \bar{\beta}_1)[(\beta_1 - \bar{\beta}_1)t_s + \beta_1(1 - \lambda^{t_s})/\log \lambda] \\ c_s &= p\sigma_1^2[t_s + (1 - \lambda^{t_s})/\log \lambda] - p\sigma_{01} + \beta_1 p(1 - p)[(\beta_1 - \bar{\beta}_1)t_s + \\ &\quad \beta_1(1 - \lambda^{t_s})/\log \lambda] \end{aligned}$$

e

$$V_s^*(t) = \sigma^2[\min(t_s, t) - p(2\lambda^t + 2\lambda^{t_s} - 2 - \lambda^{|t-t_s|} - \lambda^{t+t_s})/(2\log \lambda)] \quad (4.3.2)$$

é a dispersão da mistura do processo OU integrado com MB. A parcela $a_s + b_s t + c_s(1 - \lambda^t)/\log \lambda$ se assemelha ao valor esperado do modelo (4.2.1) e $\sigma^2 \min(t_s, t)$ introduz uma mudança na direção da curva quando $t = t_s$. $V_s(t)$ se aproxima de uma função linear quando t cresce, pois λ^t , $\lambda^{|t-t_s|}$ e λ^{t+t_s} se aproximam de 0. Em particular, quando $t_s = 0$,

$$V_0(t) = \sigma_0^2 - [p\sigma_{01} + (1 - p)\bar{\sigma}_{01}]t - p\sigma_{01}(1 - \lambda^t)/\log \lambda$$

é decrescente para $p\sigma_{01} + (1 - p)\bar{\sigma}_{01} > 0$ e quando $t = t_s$ a função que descreve a variância se aproxima de uma função quadrática quando t cresce. A primeira derivada

$$\dot{V}_s(t) = b_s - c_s \lambda^t + \dot{V}_s^*(t) \quad (4.3.3)$$

onde

$$\dot{V}_s^*(t) = \begin{cases} \sigma^2[2\lambda^t - p(2 - \lambda^{t_s})\lambda^{2t} - p\lambda^{t_s}]/2\lambda^t & \text{para } t < t_s \\ p\sigma^2\lambda^{t-t_s}(1 - \lambda^{t_s})^2/2 & \text{para } t > t_s \end{cases} \quad (4.3.4)$$

crece com t_s . $\dot{V}_s^*(t) > 0$, obviamente para $t > t_s$, e quando $t < t_s$, as raízes da função quadrática em λ^t no numerador,

$$[1 \pm \sqrt{1 - 2p^2\lambda^{t_s} + p^2\lambda^{2t_s}}]/[p(2 - \lambda^{t_s})]$$

são tais que a menor está aquém de λ^{t_s} e a maior está além de 1, fazendo com que o valor da derivada para $0 \leq t < t_s$ seja positivo, pois o coeficiente do termo quadrático é negativo. Assim, temos $\dot{V}_s(t) > 0$ e para t_s suficientemente grande, $c_s > 0$ pois $\beta_1 > \bar{\beta}_1$. Nestas condições, a derivada segunda

$$\ddot{V}_s(t) = \begin{cases} -\log \lambda \{ [2c_s + (2 - \lambda^{t_s})p\sigma^2]\lambda^{2t} - p\sigma^2\lambda^{t_s} \} / 2\lambda^t & \text{para } t < t_s \\ -\lambda^t \log \lambda [c_s - p\sigma^2(1 - \lambda^{t_s})^2 / 2\lambda^{t_s}] & \text{para } t > t_s \end{cases} \quad (4.3.5)$$

é negativa para t_s suficientemente grande e $t > t_s$ (fazendo com que o crescimento de $V_s(t)$ seja amortecido) e é quadrática em λ^t quando $t < t_s$ com uma única raiz positiva

$$\lambda^{t_0} = p\sigma\lambda^{t_s/2} / \sqrt{2c_s + (2 - \lambda^{t_s})p\sigma^2} < 1$$

pois $\lambda^{t_s} \leq 1 \leq 2 - \lambda^{t_s}$ fazendo com que $\ddot{V}_s(t)$ seja positiva para $0 \leq t < t_0$ e negativa caso contrário. Neste caso, o crescimento de $V_s(t)$ é sigmoidal.

O incremento por unidade de tempo entre dois elementos adjacentes na linha s da matriz de dispersão induzida por este modelo é dado por

$$\Delta V_s(t_u) / \Delta t_u = [V_s(t_u) - V_s(t_{u-1})] / (t_u - t_{u-1}) = b_s + c_s k(t_{u-1}, \Delta t_u) +$$

$$\sigma^2 \begin{cases} pk(t_{u-1}, \Delta t_u)(2 - \lambda^{t_s} + \lambda^{t_s - t_u - t_{u-1}}) / 2 + 1 & \text{para } t_u \leq t_s \\ pk(t_{u-1}, \Delta t_u)(2 - \lambda^{t_s} - \lambda^{-t_s}) & \text{para } t_s \leq t_{u-1} \end{cases} \quad (4.3. 6)$$

onde $k(t_{u-1}, \Delta t_u) = \lambda^{t_{u-1}}(1 - \lambda^{\Delta t_u}) / (\Delta t_u \log \lambda)$. Quando vistas como funções

de t_s , as duas curvas exibem um componente linear comum através de b_s , separadas por $\sigma^2[1 + pk(t_{u-1}, \Delta t_u)(2\lambda^{-t_s} + \lambda^{t_s - t_u - t_{u-1}}) / 2]$. Para Δt_u fixo, $k(t_{u-1}, \Delta t_u)$ tende a ser mais variável quanto menor for t_{u-1} , devido ao mesmo comportamento de $\lambda^{t_{u-1}}$. Este comportamento também pode ser notado na Figura 4.7 onde os pontos $\bullet$ correspondentes a $t_u \leq t_s$ estão mais espalhados

que os outros pontos para t_s fixo. Assim, a matriz de dispersão amostral dos dados do Exemplo 3 (ver Figura 4.5) possui todas as características citadas para a dispersão do modelo de efeitos aleatórios com erro condicional OU integrado.

4.4 Análises individuais

Nesta Seção apresentamos as análises individuais utilizando todos os dados disponíveis para $y_j, j = 1, \dots, N$. Dada a discussão apresentada na Seção 1.5, isto implica a necessidade de considerar o modelo (4.2.1) truncado, o que torna a estimação algebricamente intratável. Para torná-la viável, simplificamos o modelo para o erro condicional, considerando um MB, pois sob este modelo, os incrementos

$$\Delta y_{ju} = y_{ju} - y_{ju-1}, \quad u = 1, \dots, n_j$$

definidos através da aplicação do operador Δ a y_{ju} que é a u -ésima observação da unidade amostral j , correspondente ao instante t_{ju} , são considerados independentes. Simplificação na especificação da estrutura da dispersão das observações, em geral, não introduz nenhum vício na estimação, influenciando apenas a dispersão dos estimadores. O modelo simplificado elimina dos elementos da matriz de dispersão do modelo original apenas uma parcela com variação limitada, a saber

$$-\sigma^2(2\lambda^{t_u} + 2\lambda^{t_s} - 2 - \lambda^{|t_u - t_s|} - \lambda^{t_u + t_s})/(2 \log \lambda),$$

constituindo uma aproximação linear de $V_s^*(t)$.

Também consideramos censura de m_j observações finais, onde m_j varia de acordo com a unidade amostral, podendo ser igual a zero quando não houver dado censurado. Vamos, também, considerar inicialmente $0 \leq \lambda_j < 1$, a ser estimado. Assim, a verossimilhança individual é

$$L_j = \prod_{u=1}^{n_j} \{\phi(\Delta z_{ju})/[\sigma\sqrt{\Delta t_{ju}}\Phi(z_{ju-1})]\} \prod_{k=1}^{m_j} [1 - \Phi(\Delta z_{jn_j+k})] \quad (4.4.1)$$

onde ϕ e Φ são respectivamente as funções de densidade e de distribuição acumulada de uma normal padrão, $\Delta z_{ju} = (\Delta y_{ju} - \beta_{1j}\Delta x_{j\lambda_j u})/(\sigma\sqrt{\Delta t_{ju}})$ e

$z_{ju-1} = (y_{ju-1} - \beta_{1j}\Delta x_{j\lambda,u})/(\sigma\sqrt{\Delta t_{ju}})$. O logaritmo da verossimilhança é dado por

$$l_j = K_j - n_j \log \sigma - \sum_{u=1}^{n_j} [(\Delta z_{ju})^2/2 + \log \Phi(z_{ju-1})] + \sum_{k=1}^{m_j} \log[1 - \Phi(\Delta z_{jn_j+k})] \quad (4.4.2)$$

onde K_j é uma constante conhecida. As derivadas primeiras de l_j são dadas por

$$\begin{aligned} \partial l_j / \partial \beta_{1j} &= \sum_{u=1}^{n_j} (\Delta z_{ju}) x_{ju} + \sum_{u=1}^{n_j} x_{ju} \varphi(\Delta z_{ju-1}) + \sum_{k=1}^{m_j} x_{jn_j+k} \psi(z_{ju-1}) \\ \partial l_j / \partial \lambda_j &= \beta_{1j} \left[\sum_{u=1}^{n_j} \Delta z_{ju} \dot{x}_{ju} + \sum_{u=1}^{n_j} \varphi(\Delta z_{ju-1}) \dot{x}_{ju} + \right. \\ &\quad \left. \sum_{k=1}^{m_j} \psi(\Delta z_{jn_j+k}) \dot{x}_{jn_j+k} \right] \\ \partial l_j / \partial \sigma &= \left[\sum_{u=1}^{n_j} (\Delta z_{ju})^2 + \sum_{u=1}^{n_j} \varphi(z_{ju-1}) z_{ju-1} + \right. \\ &\quad \left. \sum_{k=1}^{m_j} \psi(\Delta z_{jn_j+k}) \Delta z_{jn_j+k} - n_j \right] / \sigma \end{aligned} \quad (4.4.3)$$

onde $x_{ju} = \Delta x_{j\lambda,u} / [\sigma\sqrt{\Delta t_{ju}}]$, $\varphi(z) = \phi(z)/\Phi(z)$, $\psi(z) = \phi(z)/(1 - \Phi(z))$ e $\dot{x}_{ju} = \partial x_{ju} / \partial \lambda_j$. As derivadas segundas são dadas por

$$\begin{aligned} \partial^2 l_j / \partial \beta_{1j}^2 &= \sum_{u=1}^{n_j} x_{ju}^2 [z_{ju-1} \varphi(z_{ju-1}) + \varphi(z_{ju-1})^2 - 1] + \\ &\quad \sum_{k=1}^{m_j} x_{jn_j+k}^2 \psi(\Delta z_{jn_j+k}) [\Delta z_{jn_j+k} - \psi(\Delta z_{jn_j+k})] \\ \partial^2 l_j / \partial \beta_{1j} \partial \sigma &= \left\{ \sum_{u=1}^{n_j} x_{ju} \{ \varphi(z_{ju-1}) [z_{ju-1} (z_{ju-1} + \varphi(z_{ju-1})) - 1] - \right. \\ &\quad \left. 2\Delta z_{ju} \} + \sum_{k=1}^{m_j} \{ \Delta z_{jn_j+k} [\Delta z_{jn_j+k} - \psi(\Delta z_{jn_j+k})] - 1 \} \right\} \end{aligned}$$

$$\begin{aligned}
& x_{jn_j+k} \psi(\Delta z_{jn_j+k}) / \sigma \\
\partial^2 l_j / \partial \sigma^2 &= \{n_j - \sum_{u=1}^{n_j} \{z_{ju-1} \varphi(z_{ju-1}) [z_{ju-1} (z_{ju-1} + \varphi(z_{ju-1})) - 2] \\
& - 3\Delta z_{ju}\} + \sum_{k=1}^{m_j} [\Delta z_{jn_j+k} (\Delta z_{jn_j+k} - \psi(\Delta z_{jn_j+k})) - 2] \\
& \Delta z_{jn_j+k} \psi(\Delta z_{jn_j+k})\} / \sigma^2 \\
\partial^2 l_j / \partial \beta_{1j} \partial \lambda_j &= \sum_{k=1}^{m_j} \dot{x}_{ju} \psi(\Delta z_{jn_j+k}) \{1 + \beta_{1j} x_{jn_j+k} [\Delta z_{jn_j+k} - \psi(\Delta z_{jn_j+k})]\} \\
& + \sum_{u=1}^{n_j} \{\Delta z_{ju} + \varphi(z_{ju-1}) [1 + \beta_{1j} x_{ju} (z_{ju-1} + \varphi(z_{ju-1}))] - \\
& \beta_{1j} x_{ju}\} \dot{x}_{ju} \\
\partial^2 l_j / \partial \sigma \partial \lambda_j &= \{\sum_{u=1}^{n_j} \dot{x}_{ju} [\varphi(z_{ju-1}) (z_{ju-1} (z_{ju-1} + \varphi(z_{ju-1})) - 1) - 2\Delta z_{ju}] \\
& + \sum_{k=1}^{m_j} \psi(\Delta z_{jn_j+k}) [\Delta z_{jn_j+k} (\Delta z_{jn_j+k} - \psi(\Delta z_{jn_j+k})) - 1] \\
& \dot{x}_{jn_j+k}\} \beta_{1j} / \sigma \\
\partial^2 l_j / \partial \lambda_j^2 &= \{\sum_{k=1}^{m_j} \psi(\Delta z_{jn_j+k}) [\ddot{x}_{jn_j+k} + \beta_{1j} \dot{x}_{jn_j+k}^2 (\Delta z_{jn_j+k} - \\
& \psi(\Delta z_{jn_j+k}))] + \sum_{u=1}^{n_j} \ddot{x}_{ju} [\Delta z_{ju} + \varphi(z_{ju-1})] + \beta_{1j} \sum_{u=1}^{n_j} \dot{x}_{ju}^2 \\
& [\varphi(z_{ju-1}) (z_{ju-1} + \varphi(z_{ju-1})) - 1]\} \beta_{1j} \quad (4.4. 4)
\end{aligned}$$

onde $\ddot{x}_{ju} = \partial^2 x_{ju} / \partial \lambda_j^2$.

A maximização da verossimilhança foi feita através do método de Newton-Raphson. Em seis casos não houve convergência (animais 2,3,5,17,22 e 25), e em seis outros o parâmetro λ_j não foi significativo (animais 10,16, 20,21,26 e 28), indicando a possibilidade de superparametrização. Assim, para estes casos, os parâmetros foram reestimados com $\lambda_j = 0$ fixo, tornando (4.2.1) um modelo linear em β com a média dada por uma reta e a estrutura de dispersão por um MB, o que fica de acordo com a estrutura adotada nesta Seção. Para averiguar a possibilidade de superparametrização para os casos onde não houve convergência, analisamos seus diagramas de dispersão mostrados na Figura 4.11 (pg. 108), verificando-se a inexistência de cur-

vatura no início do processo. Com a eliminação de λ do modelo, não houve problema de convergência. As Tabelas 4.4 a 4.7 mostram as estimativas obtidas para cada tratamento e a dispersão assintótica dos parâmetros estimados dada pelos valores numéricos assumidos por (4.4.4), substituindo os parâmetros desconhecidos por suas estimativas, e em forma matricial, com os desvios padrões na diagonal principal e as correlações nas outras posições. Foram também calculadas as médias ponderadas dos parâmetros para cada tratamento (no caso de λ as médias foram calculadas apenas entre os valores não nulos), com os pesos iguais aos inversos das variâncias estimadas, em relação às quais os parâmetros individuais foram comparados, sendo sublinhados quando apresentaram diferença significativa, o que pode ser considerado como uma indicação de um valor discrepante ou de efeito aleatório. A estimativa de λ da unidade amostral 19 é discrepante em relação às demais unidades amostrais, mas uma inspeção da dispersão assintótica indica um desvio padrão estimado pequeno para $\hat{\lambda}$ e grande para $\hat{\beta}_1$, em relação a outras unidades amostrais; além disso, o coeficiente de correlação é próximo de 1, indicando haver mal condicionamento dos dados para estimar estes parâmetros. Esta unidade amostral mostrou ser influente para estimar o λ médio, sendo excluída do cálculo de sua estimativa. Muitos valores estimados de β_1 são significativamente diferentes da média, sugerindo haver efeito aleatório. Em relação a σ não se constatou nenhum valor significativamente diferente da média de cada tratamento.

A média global (entre todos os tratamentos) de λ foi estimada em 0.801 (com erro padrão de 0.034) e de σ em 0.480 (erro padrão de 0.023). Com relação à última, alguns tratamentos se mostraram significativamente diferentes (médias sublinhadas). Assim, o parâmetro λ , quando diferente de zero, deve ser considerado fixo, mesmo que tenha um valor distinto para o animal 19. Globalmente, ele deve ser considerado dicotômico, assumindo valor zero com probabilidade estimada de $12/28 = 0.43$. A incidência observada de casos com $\lambda = 0$ diferiu entre os tratamentos, mas não significativamente, pois com $n = 7$ um intervalo de confiança com coeficiente de confiança de aproximadamente 85.5% é dado por $[1; 4]$, o que inclui todas as incidências observadas.

Tabela 4.4: Estimativas dos parâmetros do modelo individual - Tratamento 1.

Unidade amostral	Parâmetro			Dispersão dos estimadores		
	λ	β_1	σ			
1	0.893	0.678	0.368	(0.083) 0.784 0.000	(0.078) 0.000	(0.082)
2	0	<u>0.438</u>	0.516	- -	(0.052) 0.066	(0.101)
3	0	<u>0.480</u>	0.446	- -	(0.047) 0.005	(0.097)
4	0.714	<u>0.790</u>	0.375	(0.148) 0.379 0.023	(0.056) 0.061	(0.104)
5	0	0.497	0.510	- -	(0.056) 0.096	(0.112)
6	0.801	<u>0.791</u>	0.317	(0.071) 0.415 0.008	(0.048) 0.0203	(0.078)
7	0.717	0.498	0.496	(0.276) 0.299 0.036	(0.060) 0.121	(0.102)
Média	0.822	<u>0.592</u>	<u>0.417</u>	-	-	-
D.p.	(0.050)	(0.021)	(0.036)	-	-	-

Tabela 4.5: Estimativas dos parâmetros do modelo individual - Tratamento 2.

Unidade amostral	Parâmetro			Dispersão dos estimadores		
	λ	β_1	σ			
8	0.729	0.772	0.489	(0.168) 0.375 0.000	(0.075) 0.121	(0.131)
9	0.757	0.833	0.489	(0.130) 0.390 0.033	(0.079) 0.085	(0.134)
10	0	<u>0.611</u>	0.413	- -	(0.049) 0.001	(0.093)
11	0.788	0.751	0.517	(0.127) 0.387 0.014	(0.077) 0.037	(0.128)
12	0.857	0.681	0.875	(0.206) 0.618 0.058	(0.155) 0.094	(0.209)
13	0.845	0.680	0.630	(0.139) 0.507 0.034	(0.094) 0.068	(0.152)
14	0.768	<u>0.891</u>	0.507	(0.125) 0.411 0.029	(0.082) 0.071	(0.139)
Média	0.786	<u>0.725</u>	0.510	-	-	-
D.p.	(0.058)	(0.029)	(0.049)	-	-	-

Tabela 4.6: Estimativas dos parâmetros do modelo individual - Tratamento 3.

Unidade amostral	Parâmetro			Dispersão dos estimadores		
	λ	β_1	σ			
15	0.814	0.765	0.579	(0.200) 0.624 0.082	(0.106) 0.132	(0.155)
16	0	0.661	0.602	- -	(0.076) 0.061	(0.144)
17	0	<u>0.547</u>	0.463	- -	(0.054) 0.061	(0.097)
18	0.713	0.604	0.415	(0.223) 0.354 0.006	(0.055) 0.017	(0.092)
19	<u>0.954</u>	1.114	0.653	(0.036) 0.939 0.085	(0.342) 0.099	(0.161)
20	0	0.725	0.412	- -	(0.054) 0.000	(0.103)
21	0	<u>0.823</u>	0.459	- -	(0.067) 0.000	(0.133)
Média	0.769	0.670	0.479	-	-	-
D.p.	(0.149)	(0.026)	(0.045)	-	-	-

Tabela 4.7: Estimativas dos parâmetros do modelo individual - Tratamento 4.

Unidade amostral	Parâmetro			Dispersão dos estimadores		
	λ	β_1	σ			
22	0	<u>0.573</u>	0.633	-	(0.072)	
				-	0.109	(0.134)
				(0.167)		
23	0.771	0.962	0.687	0.465	(0.122)	
				0.000	0.000	(0.198)
				(0.156)		
24	0.768	0.684	0.492	0.356	(0.066)	
				0.016	0.045	(0.120)
25	0	0.817	0.916	-	(0.126)	
				-	0.147	(0.221)
26	0	0.831	0.733	-	(0.097)	
				-	0.059	(0.206)
				(0.157)		
27	0.806	0.957	0.764	0.504	(0.135)	
				0.054	0.107	(0.207)
28	0	0.832	0.509	-	(0.067)	
				-	0.051	(0.141)
Média	0.782	<u>0.758</u>	<u>0.620</u>	-	-	-
D.p.	(0.092)	(0.033)	(0.061)	-	-	-

A velocidade média de cicatrização pode ser aproximada por $\bar{\beta}_{1j} = y_{j0}/t_{0j}$ onde t_{0j} é o tempo esperado para a recuperação total da ferida, estimado através de (4.2.2) iterativamente, com y_{j0} no lugar de β_{0j} e as estimativas individuais substituindo os respectivos parâmetros. A variância do estimador $\hat{t}_{0j}$ calculada através do Método Delta (ver detalhes no Apêndice D) é

$$\begin{aligned}
Var(\hat{t}_{0j}) \cong & \{t_{0j}\lambda_j^{t_{0j}-1}/[\log \lambda_j(1 - \lambda_j^{t_{0j}})] + 1/[\lambda_j(\log \lambda_j)^2]\}^2 Var(\hat{\lambda}_j) \\
& + \{y_{j0}/[\beta_{1j}^2(1 - \lambda_j^{t_{0j}})]\}^2 Var(\hat{\beta}_{1j}) - 2\{y_{j0}/[\beta_{1j}^2(1 - \lambda_j^{t_{0j}})]\} \\
& \{t_{0j}\lambda_j^{t_{0j}-1}/[\log \lambda_j(1 - \lambda_j^{t_{0j}})] + 1/[\lambda_j(\log \lambda_j)^2]\} Cov(\hat{\lambda}_j, \hat{\beta}_{1j}).
\end{aligned}
\tag{4.4. 5}$$

Conseqüentemente, ainda pelo Método Delta (ver Apêndice D) obtem-se

$$Var(\hat{\beta}_{1j}) \cong \bar{\beta}_{1j}^4 Var(\hat{t}_{0j})/y_{j0}^2. \tag{4.4.6}$$

Com a finalidade de estabilizar a variância de $\hat{\sigma}$ uma outra transformação é sugerida através da análise do diagrama de dispersão dos seus desvios padrões estimados em função das estimativas (Figura 4.8). A Figura 4.9 mostra que $-\log \hat{\sigma}$ dá maior estabilidade aos seus desvios padrões estimados. As Tabelas 4.8 e 4.9 mostram os valores transformados dos parâmetros estimados e os desvios padronizados das estimativas em relação à média de cada grupo (estatísticas t). As velocidades médias estimadas são sempre menores que as velocidades assintóticas estimadas (ver Tabelas 4.4 a 4.7), e a unidade amostral 19 não mais se apresenta discrepante. Os parâmetros transformados indicam mais unidades amostrais significativamente diferentes das médias dos grupos de tratamento do que os parâmetros originais, indicando maior sensibilidade aos testes realizados. No caso de $\bar{\beta}_1$ os valores adicionais discrepantes são das unidades amostrais 7, 18 e 24, reforçando a idéia de efeito aleatório, enquanto para $-\log \sigma$ a unidade amostral 12 se mostra discrepante. As médias e os desvios padronizados (indicados por t^* nas Tabelas 4.8 e 4.9) foram recalculados sem este último valor. Os desvios padronizados das médias dos grupos de tratamento em relação à média global indicam haver diferença significativa entre os tratamentos.

Para verificar a normalidade do efeito aleatório, a Figura 4.10 relaciona os desvios em relação à média dos valores estimados de $\bar{\beta}_1$ em função dos escores normais padrão, que exibem um coeficiente de correlação linear de 0.966 ($p < 0.10$), não se podendo refutar sua normalidade. Estes valores estão sujeitos também aos erros de estimação, o que, a rigor, os torna heterocedásticos, mas com menor importância devido à presença de efeito aleatório, o que justifica o uso deste gráfico.

Figura 4.8: Desvio padrão estimado em função da estimativa de σ .

Figura 4.9: Desvio padrão estimado em função da estimativa de $\log \sigma$.

Tabela 4.9: Estimativa de $\bar{\beta}_1$, de $-\log \sigma$ e os valores padronizados em relação à média de cada grupo (continua).

Trata- mento	unidade amostral	$\bar{\beta}_1$			$-\log \sigma$			
		$\bar{\beta}_1$	d.p.	t	$-\log \hat{\sigma}$	d.p.	t	t^*
1	1	0.600	(0.039)	0.203	1.001	(0.224)	0.845	-
	2	0.438	(0.052)	<u>-3.127</u>	0.661	(0.196)	-0.933	-
	3	0.480	(0.047)	<u>-2.577</u>	0.807	(0.218)	-0.094	-
	4	0.751	(0.047)	<u>3.664</u>	0.981	(0.276)	0.589	-
	5	0.497	(0.056)	-1.820	0.673	(0.219)	-0.758	-
	6	0.737	(0.038)	<u>4.244</u>	1.149	(0.247)	1.394	-
	7	0.481	(0.054)	<u>-2.179</u>	0.700	(0.206)	-0.666	-
	média	0.593	(0.017)	<u>-4.702</u>	0.826	(0.084)	<u>-2.329</u>	-
2	8	0.731	(0.063)	0.560	0.716	(0.268)	0.451	0.105
	9	0.780	(0.064)	1.370	0.715	(0.274)	0.438	0.100
	10	0.611	(0.049)	<u>-2.063</u>	0.883	(0.224)	1.382	0.973
	11	0.702	(0.062)	0.042	0.659	(0.247)	0.248	-0.136
	12	0.617	(0.101)	-0.837	0.134	(0.239)	<u>-2.134</u>	<u>-2.575</u>
	13	0.624	(0.069)	-1.154	0.462	(0.241)	<u>-0.637</u>	<u>-1.048</u>
	14	0.834	(0.065)	<u>2.224</u>	0.680	(0.275)	0.299	-0.038
	média	0.699	(0.024)	<u>2.014</u>	0.600	(0.095)	-0.601	-
	s/ o 12	-	-	-	0.690	(0.103)	-0.227	-
3	15	0.705	(0.071)	0.472	0.546	(0.267)	-0.585	-
	16	0.661	(0.076)	-0.172	0.507	(0.239)	-0.840	-
	17	0.547	(0.054)	<u>-2.554</u>	0.770	(0.210)	0.409	-
	18	0.582	(0.047)	<u>-2.173</u>	0.880	(0.221)	0.929	-
	19	0.771	(0.060)	1.766	0.426	(0.246)	-1.168	-
	20	0.725	(0.054)	1.060	0.887	(0.250)	0.836	-
	21	0.823	(0.067)	<u>2.387</u>	0.779	(0.289)	0.314	-
	média	0.673	(0.022)	0.888	0.693	(0.092)	-0.304	-

Tabela 4.9: Estimativa de $\bar{\beta}_1$, de $-\log \sigma$ e os valores padronizados em relação à média de cada grupo (continuação).

Tratamento	unidade amostral	β_1			$-\log \sigma$			
		$\hat{\beta}_1$	d.p.	t	$-\log \hat{\sigma}$	d.p.	t	t^*
4	22	0.573	(0.072)	<u>-2.612</u>	0.458	(0.212)	0.232	-
	23	0.893	(0.093)	1.691	0.376	(0.289)	-0.141	-
	24	0.647	(0.056)	<u>-2.055</u>	0.709	(0.243)	1.319	-
	25	0.817	(0.124)	0.601	0.088	(0.241)	-1.475	-
	26	0.831	(0.097)	0.935	0.310	(0.281)	-0.394	-
	27	0.879	(0.099)	1.424	0.269	(0.271)	-0.572	-
	28	0.832	(0.067)	1.446	0.676	(0.277)	1.005	-
	média	0.745	(0.030)	<u>3.200</u>	0.414	(0.096)	<u>3.040</u>	-
média global		0.655	(0.011)	-	0.650	(0.046)	-	-
sem o 12		-	-	-	0.670	(0.047)	-	-

A Tabela 4.10 mostra as estatísticas t relativas às diferenças entre os tratamentos com relação a σ e $-\log \sigma$. Enquanto há sugestão de diferença significativa ($p < 0.05$) apenas entre os tratamentos 1 e 4 relativamente a σ , a consideração do parâmetro transformado sugere uma diferença significativa do tratamento 4 com todos os outros tratamentos ($p < 0.05$ em relação aos tratamentos 1 e 3 e $p < 0.10$ em relação ao tratamento 2).

Tabela 4.10: Estatísticas t para as diferenças em σ e $-\log \sigma$ entre os tratamentos.

Tratamento	σ			$-\log \sigma$		
	2	3	4	2	3	4
1	1.542	1.104	<u>2.856</u>	1.063	1.117	<u>3.277</u>
2	-	-0.466	1.374	-	-0.022	<u>1.958</u>
3	-	-	1.839	-	-	<u>2.107</u>

A análise das estimativas individuais dos parâmetros do modelo com erros truncados e censura das últimas observações mostrou que λ deve ser considerado dicotômico, com valor 0.801 ou 0 (caso em que a resposta sem truncamento é linear), tendo-se constatado um valor discrepante (0.954) para

Figura 4.10: Gráfico em papel de probabilidade normal dos desvios em relação às médias de $\bar{\beta}_1$.

a unidade amostral 19. O parâmetro σ mostrou ser sensível ao tratamento, com valor 0.475 para os três primeiros grupos de tratamento e 0.661 para o último, tendo-se constatado um valor discrepante (0.875) para a unidade amostral 12. O parâmetro $\bar{\beta}_1$ pode ser considerado aleatório, não havendo evidência de violação da suposição de normalidade. Na Seção 4.5 damos prosseguimento à análise, modelando as estimativas das velocidades médias obtidas nesta Seção através de um modelo de efeitos aleatórios, obtemos novas estimativas sob este modelo e ao final verificamos a qualidade de ajuste do modelo resultante às observações originais.

4.5 Estimação dos parâmetros do modelo para efeitos aleatórios (4.2.3)

No espírito de Hui & Berger (1983), vamos considerar nesta Seção um modelo de efeitos aleatórios para o vetor $\mathbf{b}_{ij} = (y_{0ij}, \hat{\beta}_{1ij})^t$ onde o índice $i = 1, \dots, 4$ se refere ao tratamento, sendo y_{0ij} o valor inicial do diâmetro da ferida e $\hat{\beta}_{1ij}$, a velocidade média de cicatrização estimada na análise individual, lembrando

que nas Seções anteriores já verificamos a normalidade destes termos.

No primeiro estágio do modelo de efeitos aleatórios para $\mathbf{b}_{ij}$ consideramos um modelo com erro de medida apenas no segundo termo, cuja variância, supostamente conhecida, é dada pela variância estimada de $\hat{\beta}_{1ij}$ que vamos denotar por σ_{ij}^2 , resultando em

$$\mathbf{b}_{ij}|\bar{\beta}_{ij} \sim N(\bar{\beta}_{ij}, \text{Diag}(0, \sigma_{ij}^2)). \quad (4.5.1)$$

onde $\bar{\beta}_{ij} = (\beta_{0ij}, \bar{\beta}_{1ij})$ é o valor esperado condicional de $\mathbf{b}_{ij}$. O segundo estágio do modelo é dado por

$$\bar{\beta}_{ij} \sim N(\mathbf{X}_i\bar{\beta}, \Gamma_0) \quad (4.5.2)$$

onde $\mathbf{X}_1 = (\mathbf{e}_1, \mathbf{e}_2)^t$, $\mathbf{X}_2 = (\mathbf{e}_1, \mathbf{e}_3)^t$, $\mathbf{X}_3 = (\mathbf{e}_1, \mathbf{e}_4)^t$, $\mathbf{X}_4 = (\mathbf{e}_5, \mathbf{e}_6)^t$ e $\mathbf{e}_i$ é um vetor de zeros, exceto na posição i em que o elemento correspondente vale 1, com a mesma dimensão de $\bar{\beta} = (\beta_{01}, \bar{\beta}_{11}, \bar{\beta}_{12}, \bar{\beta}_{13}, \beta_{04}, \bar{\beta}_{14})^t$, e Γ_0 está definido em (4.2.3), resultando na distribuição marginal de $\mathbf{b}_{ij}$ dada por

$$\mathbf{b}_{ij} \sim N(\mathbf{X}_i\bar{\beta}, \mathbf{V}_{ij}) \quad (4.5.3)$$

onde $\mathbf{V}_{ij} = \Gamma_0 + \text{Diag}(0, \sigma_{ij}^2)$.

Conseqüentemente, o logaritmo da verossimilhança da unidade amostral j do tratamento i é dado por

$$l_{ij} = -\log 2\pi - \log |\mathbf{V}_{ij}|/2 - (\mathbf{b}_{ij} - \mathbf{X}_i\bar{\beta})^t \mathbf{V}_{ij}^{-1} (\mathbf{b}_{ij} - \mathbf{X}_i\bar{\beta}) \quad (4.5.4)$$

e o estimador de MV de $\bar{\beta}$ é o estimador de Aitken (1935)

$$\hat{\bar{\beta}} = \sum_{i=1}^4 \left(\sum_{j=1}^7 \mathbf{X}_i^t \mathbf{V}_{ij}^{-1} \mathbf{X}_i \right)^{-1} \sum_{i=1}^4 \sum_{j=1}^7 \mathbf{X}_i^t \mathbf{V}_{ij}^{-1} \mathbf{b}_{ij}. \quad (4.5.5)$$

Para estimar $\theta = (\sigma_0^2, \bar{\sigma}_{01}, \bar{\sigma}_1^2)^t$ cujos elementos são os elementos distintos de Γ_0 , utilizamos o método do score de Fisher. A primeira derivada de l_{ij} em relação a θ é dada por

$$\begin{pmatrix} (\mathbf{b}_{ij} - \mathbf{X}_i \bar{\boldsymbol{\beta}})' \mathbf{V}_{ij}^{-1} \partial \Gamma_0 / \partial \sigma_0^2 \mathbf{V}_{ij}^{-1} (\mathbf{b}_{ij} - \mathbf{X}_i \bar{\boldsymbol{\beta}}) - (\bar{\sigma}_1^2 + \sigma_{ij}^2) / (2|\mathbf{V}_{ij}|) \\ (\mathbf{b}_{ij} - \mathbf{X}_i \bar{\boldsymbol{\beta}})' \mathbf{V}_{ij}^{-1} \partial \Gamma_0 / \partial \bar{\sigma}_{01} \mathbf{V}_{ij}^{-1} (\mathbf{b}_{ij} - \mathbf{X}_i \bar{\boldsymbol{\beta}}) + \bar{\sigma}_{01} / |\mathbf{V}_{ij}| \\ (\mathbf{b}_{ij} - \mathbf{X}_i \bar{\boldsymbol{\beta}})' \mathbf{V}_{ij}^{-1} \partial \Gamma_0 / \partial \bar{\sigma}_1^2 \mathbf{V}_{ij}^{-1} (\mathbf{b}_{ij} - \mathbf{X}_i \bar{\boldsymbol{\beta}}) - \sigma_0^2 / (2|\mathbf{V}_{ij}|) \end{pmatrix} \quad (4.5.6)$$

cuja soma em relação a i e j resulta na derivada da verossimilhança total que denotamos por $\boldsymbol{\eta}$ e a informação de Fisher denotada por $\mathbf{J}$ é dada por

$$\begin{pmatrix} (\bar{\sigma}_1^2 + \sigma_{ij}^2)^2 / (2|V_{ij}|^2) \\ -\bar{\sigma}_{01}(\bar{\sigma}_1^2 + \sigma_{ij}^2) / |V_{ij}|^2 & [\sigma_0^2(\bar{\sigma}_1^2 + \sigma_{ij}^2) + \bar{\sigma}_{01}^2] / |V_{ij}|^2 \\ -2\bar{\sigma}_{01}^2 / |V_{ij}|^2 & -\sigma_0^2 \bar{\sigma}_{01} / |V_{ij}|^2 & \sigma_0^4 / (2|V_{ij}|^2) \end{pmatrix}. \quad (4.5.7)$$

Tabela 4.11: Estimativas dos parâmetros populacionais.

Parâmetro	estimativa	d.p.	d.p. robusto
β_{01}	<u>45.9095</u>	0.4070	0.4347
$\bar{\beta}_{11}$	<u>0.5685</u>	0.0367	0.0446
$\bar{\beta}_{12}$	<u>0.7075</u>	0.0407	0.0307
$\bar{\beta}_{13}$	<u>0.6804</u>	0.0394	0.0359
β_{04}	<u>48.8512</u>	0.7050	0.5362
$\bar{\beta}_{14}$	<u>0.7676</u>	0.0457	0.0466
σ_0^2	<u>3.4789</u>	0.9887	-
$\bar{\sigma}_{01}$	0.0456	0.0407	-
$\bar{\sigma}_1^2$	<u>0.0076</u>	0.0033	-

Partindo dos valores iniciais $\theta_0 = (3.43, 0.042, 0.0085)^t$ e utilizando o método do escore de Fisher (incremento < 0.001) obtêm-se as estimativas de MV dos parâmetros. As estimativas finais encontram-se na Tabela 4.11. Apenas o parâmetro $\bar{\sigma}_{01}$ não é significativamente diferente de zero. O desvio padrão robusto não coincide com o desvio padrão assintótico usual, o que pode denotar falta de ajuste do modelo proposto (o que será discutido na Seção 5), razão pela qual deixamos de calcular as correções das variâncias assintóticas.

A diferença entre β_{01} e β_{04} é de -2.9417 com erro padrão de 0.8140 usando o d.p. assintótico usual ou de 0.6903 usando o d.p. assintótico robusto, o que é significativo ($p < 0.001$). As diferenças entre as velocidades médias de cicatrização se encontram na Tabela 4.12, com ambos os desvios padrões, mostrando haver diferenças significativas entre o tratamento 1 e os demais tratamentos relativamente a essas velocidades.

Utilizando as equações (4.1.5) a (4.1.7) obtêm-se as estimativas de β_1 , σ_{01} e σ_1^2 , relativas às unidades amostrais com $\lambda \neq 0$, apresentadas na Tabela 4.13.

Tabela 4.12: Diferenças entre tratamentos relativamente a $\bar{\beta}_1$.

Tratamento	2	3	4
1	$\frac{-0.1390}{(0.0543 \text{ ou } 0.0534)}$	$\frac{-0.1119}{(0.0533 \text{ ou } 0.0568)}$	$\frac{-0.1991}{(0.0586 \text{ ou } 0.0645)}$
2	-	$\frac{0.0271}{(0.0562 \text{ ou } 0.0471)}$	$\frac{-0.0601}{(0.0612 \text{ ou } 0.0558)}$
3	-	-	$\frac{-0.0872}{(0.0604 \text{ ou } 0.0589)}$

Tabela 4.13: Estimativas de β_1 , σ_{01} e σ_1^2 .					
Parâmetro	Tratamento				Animal 19
	1	2	3	4	
β_{1i}	0.6022	0.7604	0.7292	0.8262	1.0950
σ_{01}	0.0512	0.0527	0.0524	0.0529	0.0940
σ_1^2	0.0095	0.0101	0.0100	0.0101	0.0416

Utilizando as estimativas apresentadas nas Tabelas 4.11 e 4.13, pode-se calcular a matriz de dispersão induzida pelo modelo estimado para cada tratamento, com proporções diferentes de casos com $\lambda \neq 0$ e com parâmetros diferenciados para as unidades amostrais 12 e 19 e para o tratamento 4. A matriz de dispersão média entre os quatro tratamentos é apresentada na Tabela 4.14, podendo-se notar que seus valores, com exceção da primeira coluna, são sempre maiores que os verificados na matriz de dispersão amostral (Tabela 4.2), o que pode ser creditado ao truncamento que neste cálculo não foi relevado. Mesmo assim, pode-se notar que o erro relativo, que deveria crescer dos valores da esquerda para os da direita, devido ao crescente truncamento, é demasiadamente grande para os valores da segunda coluna, sugerindo haver possibilidade de melhorar o modelo.

Tabela 4.14: Matriz de dispersão média induzida pelo modelo.

tempo	00.000	11.862	19.206	28.645	41.419	47.399
00.000	3.4789					
11.862	3.0226	7.5700				
19.206	2.6679	8.0957	10.6590			
28.645	2.1992	8.2006	11.6476	15.5986		
41.419	1.5573	8.0161	12.3031	17.7770	24.7925	
47.399	1.2553	7.8655	12.5202	18.5869	26.7499	30.3467

As estimativas de $\bar{\beta}_{1j}$ resultantes do modelo de efeitos aleatórios se encontram na terceira coluna da Tabela 4.15, e na última coluna estão as estimativas correspondentes de β_{1j} obtidas através da equação (4.2.5). Nas duas primeiras colunas se encontram as estimativas correspondentes obtidas dos modelos individuais (Seção 4.3).

Tabela 4.15: Estimativa do coeficiente angular (médio) sob o modelo individual e sob o modelo hierárquico.

Unidade amostral	Modelo individual		Modelo hierárquico	
	β_1	$\bar{\beta}_1$	$\bar{\beta}_1$	β_1
1	0.678	0.600	0.594	0.631
2	0.438	0.438	0.485	0.485
3	0.480	0.480	0.500	0.500
4	0.790	0.751	0.709	0.761
5	0.497	0.497	0.515	0.515
6	0.791	0.737	0.714	0.765
7	0.498	0.481	0.503	0.530
8	0.772	0.731	0.712	0.769
9	0.833	0.780	0.745	0.806
10	0.611	0.611	0.640	0.640
11	0.751	0.702	0.698	0.751
12	0.681	0.617	0.647	0.694
13	0.680	0.624	0.657	0.702
14	0.891	0.834	0.804	0.867
15	0.765	0.705	0.682	0.734
16	0.661	0.661	0.664	0.664
17	0.547	0.547	0.584	0.584
18	0.604	0.582	0.610	0.648
19	1.114	0.771	0.761	1.095
20	0.725	0.725	0.712	0.712
21	0.823	0.823	0.761	0.761
22	0.573	0.573	0.662	0.662
23	0.962	0.893	0.817	0.885
24	0.684	0.647	0.673	0.720
25	0.817	0.817	0.798	0.798
26	0.831	0.831	0.797	0.797
27	0.957	0.879	0.820	0.886
28	0.832	0.832	0.807	0.807

Para verificar a qualidade de ajuste do modelo proposto aos dados do Exemplo 3, calculamos as previsões um passo à frente de cada observação , dadas por

$$E(y_{ju} | -\Delta y_{ju} < y_{ju-1}) = y_{ju-1} - \beta_{1j} \Delta x_{j\lambda, u} + \sigma \sqrt{\Delta t_{ju}} \varphi(z_{ju-1}). \quad (4.5.8)$$

As curvas individuais ajustadas encontram-se na Figura 4.11 juntamente com os valores observados, podendo-se notar uma boa qualidade de ajuste. A Figura 4.12 contém os resíduos condicionais, que se mostram menos variáveis no início do processo (o que pode ser devido ao processo de OU integrado), e com um aparente vício positivo no final devido ao truncamento nesta fase do processo. Algumas unidades amostrais apresentam resíduos negativos de grande magnitude em relação aos demais em alguns instantes entre 20 e 40 horas, o que poderia ser investigado posteriormente, assim como a aparente tendência cíclica na metade final do processo.

4.6 Conclusão e sugestões para análises complementares

O conjunto de dados do Exemplo 3 é surpreendentemente rico em detalhes cuja observação fundamentou os passos da análise, que enumeramos a seguir:

1) O experimento é realizado identicamente nos dois olhos de cada animal. Para simplificar a análise, utilizamos o valor médio dos dois olhos em cada instante de avaliação .

2) O tamanho da ferida inicial não é homogêneo entre os animais, havendo uma diferença entre os animais do tratamento 4 e os demais, o que sugere que a aleatorização tenha sido feita antes, e não após a indução da ferida inicial, como seria desejável.

3) Apesar de os dados serem não balanceados com relação ao tempo, as observações iniciais (com exceção da quinta) estão razoavelmente balanceadas, o que permite avaliar tanto o comportamento médio em cada tratamento, quanto a sua matriz de dispersão média entre os tratamentos. A análise destas estatísticas leva à identificação do modelo de OU integrado.

4) As observações em cada unidade amostral são realizadas até completar a cicatrização das duas feridas, o que nos leva a modelar o mecanismo de não resposta através de uma distribuição normal truncada, e a um modelo de censura, ao trabalharmos com a média entre os olhos de cada animal.

Figura 4.11: Valores individuais observados e respectivas previsões um passo à frente.

Figura 4.12: Resíduos condicionais.

Isto dificulta a estimação do modelo de OU integrado, devido à intratabilidade da matriz de dispersão de uma normal multivariada truncada, tendo-se simplificado o modelo para um MB e a estimação para dois estágios.

5) No primeiro estágio da estimação, as análises individuais revelam dois tipos de curvas, uma assintoticamente linear, e a outra linear, podendo a primeira resultar de um processo de OU, e a segunda do seu caso limite, MB. Para comparação entre as curvas propõe-se utilizar a velocidade média. No segundo estágio de estimação ajustamos um modelo de efeitos aleatórios ao vetor formado pelo tamanho da ferida inicial e pela velocidade média estimada.

Do modelo proposto resulta uma distribuição não gaussiana para a velocidade assintótica, dispersões diferentes entre os grupos de tratamento, devido à diferença na composição da mistura dos dois tipos de curva entre os grupos, e uma covariância não nula entre a primeira observação e as demais, que decresce com o tempo, independentemente de haver covariância não nula entre a primeira observação e a velocidade média.

As previsões um passo à frente de cada observação individual mostram um bom ajuste do modelo aos dados, mas a verificação do ajuste através da comparação entre a dispersão observada e a dispersão induzida pelo modelo não foi concretizada devido à dificuldade no cálculo. Mesmo sendo calculada sem a hipótese de truncamento a comparação entre as duas matrizes indica haver necessidade de melhorar o modelo para explicar melhor o comportamento no início do processo.

Os desvios padrões robustos dos parâmetros populacionais mostraram ser discrepantes em relação aos desvios padrões assintóticos tradicionais, o que pode sugerir a existência de outro modelo mais adequado. Uma análise mais detalhada dos resultados do primeiro estágio de estimação pode sugerir uma diferença entre as velocidades médias dos dois tipos de curvas, no primeiro grupo de tratamento com maior evidência que no segundo, mas não nos outros. Levando em conta este fato, o coeficiente de correlação entre o valor estimado e o escore normal padrão aumenta para 0.982. Mas esta evidência não é muito convincente, por envolver poucas unidades amostrais e a comparação entre os tratamentos não ficou comprometida pelo tipo de desvio padrão assintótico utilizado.

A análise das variâncias das estimativas individuais dos parâmetros fixos pode sugerir uma transformação paramétrica que normalize o seu estimador. No contexto de dados longitudinais, esta análise é bastante simples, po-

dendo ser feita através de diagramas de dispersão, não envolvendo cálculos de estatísticas ou simulações . O que está por trás disto é a existência de várias unidades amostrais com curvas descritas através de uma mesma família paramétrica, o que substitui as simulações .

É interessante notar que neste conjunto de dados, todas as formas de identificação de um modelo de efeitos aleatórios, a saber, a análise das curvas médias, a análise da matriz de dispersão e análise das estimativas individuais, trouxeram contribuições à modelagem.

Para aperfeiçoamento do modelo, podem-se citar as seguintes sugestões:

1) considerar um atraso no início do processo, o que deve contribuir para melhorar o comportamento do modelo no início do processo;

2) reparametrizar o modelo (4.2.1) explicitando a velocidade média $\bar{\beta}_1$ através da equação (4.2.4);

3) estimar o modelo de OU com truncamento, através de integração numérica da verossimilhança (1.3.22) ou da verossimilhança restrita (1.3.32);

4) estender o modelo proposto para as observações multivariadas originais, através da consideração de um estágio inicial que descreva o comportamento dos dois olhos, através de uma média comum igual a do modelo (4.2.1) mais um outro efeito aleatório com sinais trocados entre os olhos, sendo o erro bivariado modelado através de OU integrado bivariado (Taylor & Cumberland, 1991), MB bivariado, ou ainda cada olho com um modelo univariado diferente, MB ou OU integrado;

5) modelar o tempo decorrido em função do tamanho da ferida através da distribuição gaussiana inversa, eliminando a necessidade de um modelo com truncamento;

6) analisar o processo espacial de cicatrização , utilizando as informações espaciais contidas no trabalho de Araújo (1987), em relação às quais o tamanho da ferida é apenas uma medida resumo, o que pode contribuir para explicar os pontos que não ficaram esclarecidos na análise aqui apresentada.

Capítulo 5

Conclusões e Sugestões

A identificação de um modelo de efeitos aleatórios para dados longitudinais tem sido feita através da observação das curvas individuais como foi sugerido por Fearn (1977). Weiss & Lazaro (1992) sugerem a utilização de diagramas paralelos de dispersão enquanto Rutter & Elashoff (1992) sugerem uma análise das curvas individuais por regressão ordinária seguida de uma análise da distribuição amostral dos coeficientes individuais para se verificar a adequação do valor esperado do modelo condicional

$$E(\mathbf{y}_j|\boldsymbol{\gamma}_j) = \mathbf{X}_j\boldsymbol{\beta} + \mathbf{Z}_j\boldsymbol{\gamma}_j, j = 1, \dots, N.$$

Enquanto a metodologia de Weiss & Lazaro (1992) pode dar margem a diferentes interpretações subjetivas dos gráficos, o método proposto por Rutter & Elashoff (1992) pode-se tornar mais objetivo através da utilização de estatísticas adequadas para selecionar os efeitos fixos e aleatórios. Para esta finalidade utilizamos estatísticas t no Capítulo 2, mas pode-se sugerir o desenvolvimento de estatísticas F , como é usual em modelos mistos de efeitos aleatórios para análise de variância. Um método gráfico alternativo se baseia no fato de que a dispersão entre indivíduos representada na dispersão do modelo marginal para dados completos

$$\mathbf{V} = \mathbf{Z}\boldsymbol{\Gamma}\mathbf{Z}^t + \sigma^2\mathbf{R}, j = 1, \dots, N$$

por $\mathbf{Z}\boldsymbol{\Gamma}\mathbf{Z}^t$ tem as suas linhas descritas pela mesma forma funcional do efeito aleatório $\mathbf{Z}\boldsymbol{\gamma}$. Assim, propomos analisar os diagramas paralelos de dispersão das linhas da matriz de dispersão amostral para identificar os efeitos aleatórios.

Todos os métodos acima propostos ignoram a modelagem da dispersão intra-individual $\sigma^2\mathbf{R}$, considerando-a diagonal (modelo de independência condicional). Grady (1992) propõe analisar as linhas da matriz de correlação amostral para identificar uma estrutura intra-individual AR(1). No Capítulo

3 estudamos o comportamento das linhas da matriz de dispersão marginal induzida pelo modelo de reta com intercepto e coeficiente angular aleatórios e com erros condicionalmente independentes ou sob MB, e no Capítulo 4, do modelo de reta assintótica com o intercepto e o coeficiente angular assintótico aleatórios com erros OU integrado. Assim, uma análise dos diagramas paralelos de dispersão das linhas da matriz de dispersão amostral pode identificar um destes modelos. O diagrama de dispersão dos incrementos unitários dos elementos da matriz de dispersão amostral também é útil para identificação da dispersão intra-individual em modelos de retas aleatórias.

Na prática, a análise tanto das curvas individuais quanto das linhas da matriz de dispersão amostral podem trazer informações complementares úteis à modelagem dos dados observados. Nos dados do Exemplo 3, a análise das linhas da matriz de dispersão amostral identificam um modelo de OU integrado, enquanto a análise das curvas individuais identificam o truncamento das observações finais e a mistura dos processos OU integrado e o MB.

O estudo da dispersão induzida pelo modelo de retas aleatórias com erros condicionalmente independentes indica que cada linha da matriz de correlações apresenta um máximo local fora da diagonal principal, o que pode significar que este modelo não é adequado para a maioria dos dados longitudinais, uma vez que nenhum autor consultado neste trabalho menciona este fato em suas análises. Por outro lado, o modelo de retas aleatórias com erros sob MB exibe em cada linha da matriz de correlações apenas um máximo na diagonal principal, com os outros valores podendo decrescer à medida em que se afastam da diagonal principal, o que está de acordo com algumas descrições citadas (Elston & Grizzle, 1962; Kenward, 1987). Além disso, ainda no Capítulo 3 mostramos que a utilização inadequada do modelo de independência condicional pode levar a um vício positivo na estimação da variância do coeficiente angular.

A estimação separada das curvas individuais, além de úteis na identificação de um modelo para os efeitos fixos e aleatórios, fornece os elementos necessários para a estimação em dois estágios do modelo de efeitos aleatórios para dados longitudinais (Hui & Berger, 1983), onde no segundo estágio de estimação utiliza-se adequadamente um modelo de efeitos aleatórios com erro de medida (não necessariamente com independência condicional nem homocedástico) para os coeficientes individuais estimados. Apesar de não utilizar todas as informações contidas nos dados, trata-se de um método de fácil implementação mesmo em casos problemáticos como no Exemplo 3, po-

dendo complementar as análises individuais fornecendo estimativas iniciais para a maximização da verossimilhança. Além disso, a análise das variâncias das estimativas individuais dos parâmetros fixos podem sugerir uma transformação paramétrica que normalize o seu estimador.

Para diagnóstico do modelo estimado, muitos autores sugerem analisar os resíduos marginais (Waternaux et al., 1989; Rutter & Elashoff, 1992; Weiss & Lazaro, 1992 e Jones, 1993). O programa BMDP-5V (Schluchter, 1990) fornece para dados completos as matrizes de dispersão dos resíduos marginais e induzida pelo modelo. A comparação entre a primeira e a matriz de dispersão marginal fornece uma estatística L_f equivalente à estatística Λ de Wilks, que serve para testar a adequação do modelo para os efeitos fixos, enquanto a comparação entre as duas fornece uma estatística L_v que permite testar a adequação do modelo para a dispersão marginal. Outro procedimento que nos pareceu útil para esta finalidade é comparar a matriz de dispersão assintótica robusta (White, 1982 e Royall, 1986) com a matriz de dispersão assintótica usual dos efeitos fixos. O mesmo procedimento para os parâmetros desconhecidos da matriz de dispersão marginal deve fornecer mais elementos para diagnóstico do modelo estimado.

Sugestões para pesquisa: estudar as matrizes de dispersões induzidas por modelos de efeitos aleatórios com outras estruturas de dispersão condicional; estudar a distribuição de L_v ; estimar por maximização da verossimilhança os parâmetros do modelo de efeitos aleatórios com mistura de processos OU integrado e MB com distribuição truncada identificado nos dados do Exemplo 3; estimar os parâmetros do modelo bivariado correspondente para as observações dos dois olhos de cada animal; generalizar este modelo para um modelo multivariado; desenvolver a estatística F sugerida no primeiro parágrafo deste Capítulo; desenvolver metodologia para identificar a estrutura de dispersão condicional quando os dados não são balanceados com relação ao tempo.

Apêndice A

Processos de Ornstein-Ühlenbeck

A equação diferencial estocástica (1.2.7) pode ser escrita

$$d\varepsilon_t/dt + \alpha\varepsilon_t = \omega dB_t/dt. \quad (0.0.1)$$

Multiplicando ambos os lados por $\exp(\alpha t)$ obtém-se

$$\exp(\alpha t)d\varepsilon_t/dt + \alpha \exp(\alpha t)\varepsilon_t = \omega \exp(\alpha t)dB_t/dt. \quad (0.0.2)$$

Notando que o lado esquerdo é a derivada de $\exp(\alpha t)\varepsilon_t$, (1.2.7) pode ser escrito

$$d(\exp(\alpha t)\varepsilon_t) = \omega \exp(\alpha t)dB_t. \quad (0.0.3)$$

Da integração entre $t - \Delta t$ e t resulta

$$\exp(\alpha t)\varepsilon_t - \exp(\alpha(t - \Delta t))\varepsilon_{t-\Delta t} = \omega \int_{t-\Delta t}^t \exp(\alpha\tau)dB_\tau. \quad (0.0.4)$$

Isolando ε_t pode-se obter a forma autorregressiva

$$\varepsilon_t = \exp(-\alpha\Delta t)\varepsilon_{t-\Delta t} + \omega \int_{t-\Delta t}^t \exp(-\alpha(t - \tau))dB_\tau. \quad (0.0.5)$$

O último termo tem a variância calculada por

$$\begin{aligned} \omega^2 E\left[\int_{\tau=t-\Delta t}^t \exp(-\alpha(t - \tau))dB_\tau\right]^2 &= \omega^2 \int_{\tau=t-\Delta t}^t \exp(-2\alpha(t - \tau))d\tau \\ &= \frac{\omega^2}{2\alpha}(1 - \exp(-2\alpha\Delta t)). \end{aligned} \quad (0.0.6)$$

Quando $\varepsilon_t = \beta - \beta_t$ (0.0.5) resulta em (1.2.10). Quando $\Delta t = t$

$$\beta_t = \beta(1 - \exp(-\alpha t)) + \exp(-\alpha t)\beta_0 + \omega \int_0^t \exp(-\alpha(t - \tau))dB_\tau \quad (0.0.7)$$

onde β_0 é o valor inicial (quando $t = 0$) de β_t . Sua covariância com β_s é dada por

$$\begin{aligned} Cov(\beta_s, \beta_t) &= \omega^2 E\left[\int_{\tau=0}^t \exp(-\alpha(t - \tau))dB_\tau \int_{\tau=0}^s \exp(-\alpha(s - \tau))dB_\tau\right] \\ &= \omega^2 \int_{\tau=0}^{\min(s,t)} \exp(-\alpha(t + s - 2\tau))d\tau \end{aligned} \quad (0.0. 8)$$

resultando em (1.2.11).

Se a velocidade de um processo é descrita por $-\beta_t$, com $\beta_0 = 0$, sua posição y_t é dada por

$$\begin{aligned} y_t &= y_0 - \int_0^t \beta_\tau d\tau \\ &= y_0 - \beta \int_0^t (1 - \exp(-\alpha\tau))d\tau + \omega \int_0^t \left[\int_0^\tau \exp(-\alpha(\tau - u))dB_u \right] d\tau \\ &= y_0 - \beta \{t + [1 - \exp(-\alpha t)]/\alpha\} + \omega \int_0^t \left[\int_\tau^t \exp(-\alpha(\tau - u))d\tau \right] dB_u \\ &= y_0 - \beta \{t + [1 - \exp(-\alpha t)]/\alpha\} + \frac{\omega}{\alpha} \int_0^t [1 - \exp(-\alpha(t - u))]dB_u \end{aligned} \quad (0.0. 9)$$

onde y_0 é a posição inicial. Sua covariância com y_s , fazendo $\sigma^2 = \omega^2/\alpha^2$, é dada por

$$\begin{aligned} Cov(y_s, y_t) &= \sigma^2 E\left\{ \int_0^t [1 - \exp(-\alpha(t - u))]dB_u \right. \\ &\quad \left. \int_0^s [1 - \exp(-\alpha(s - u))]dB_u \right\} \\ &= \sigma^2 \int_0^{\min(s,t)} [1 - \exp(-\alpha(t - u))][1 - \exp(-\alpha(s - u))]du \\ &= \sigma^2 \int_0^{\min(s,t)} [1 - \exp(-\alpha(t - u)) - \exp(-\alpha(s - u)) \end{aligned}$$

$$\begin{aligned}
& + \exp(-\alpha(t+s-2u))]du \\
& = \sigma^2 \{ \min(s,t) - [\exp(-\alpha(t - \min(t,s))) - \exp(-\alpha t) \\
& + \exp(-\alpha(s - \min(t,s))) - \exp(-\alpha s)]/\alpha \\
& + [\exp(-\alpha(t+s-2\min(t,s))) - \exp(-\alpha(t+s))]/(2\alpha) \} \\
& = \sigma^2 \{ \min(s,t) - [2 - 2\exp(-\alpha t) + 2\exp(-\alpha|s-t|) \\
& - 2\exp(-\alpha s) - \exp(-\alpha|t-s|) + \exp(-\alpha(t+s))]/(2\alpha) \} \\
& = \sigma^2 \{ \min(s,t) - [2 - 2\exp(-\alpha t) + \exp(-\alpha|s-t|) \\
& - 2\exp(-\alpha s) + \exp(-\alpha(t+s))]/(2\alpha) \} \quad (0.0. 10)
\end{aligned}$$

resultando em (1.2.14) fazendo $\lambda = \exp(-\alpha)$.

Apêndice B

Expansão linear de $\beta_{1j}|(\lambda_j = \lambda)$

Lembrando que $t_{0j} = \beta_{0j}/\bar{\beta}_{1j}$ a expansão em série de Taylor do lado esquerdo de (4.2.5) em torno dos parâmetros populacionais β_0 e $\bar{\beta}_1$, com

$$\begin{aligned} t_0 &= \beta_0/\bar{\beta}_1 \\ d\lambda^{t_{0j}}/d\beta_{0j} &= \lambda^{t_{0j}} \log \lambda/\bar{\beta}_{1j} \\ d\lambda^{t_{0j}}/d\bar{\beta}_{1j} &= -\beta_{0j}\lambda^{t_{0j}} \log \lambda/\bar{\beta}_{1j}^2 \end{aligned}$$

resulta na aproximação

$$\begin{aligned} \beta_{1j}|(\lambda_j = \lambda) &\cong \beta_1 + (\beta_{0j} - \beta_0)\{1/[t_0 + (1 - \lambda^{t_0})/\log \lambda] \\ &- \beta_0(1/\bar{\beta}_1 - \lambda^{t_0}/\bar{\beta}_1)/[t_0 + (1 - \lambda^{t_0})/\log \lambda]^2\} \\ &- (\bar{\beta}_{1j} - \bar{\beta}_1)\beta_0(-\beta_0/\bar{\beta}_1^2 + \lambda^{t_0}\beta_0/\bar{\beta}_1^2)/[t_0 + (1 - \lambda^{t_0})/\log \lambda]^2 \\ &= \beta_1 + \{(\beta_{0j} - \beta_0)[t_0 + (1 - \lambda^{t_0})/\log \lambda - t_0(1 - \lambda^{t_0})] \\ &+ (\bar{\beta}_{1j} - \bar{\beta}_1)(\beta_0/\bar{\beta}_1)^2(1 - \lambda^{t_0})\}/[t_0 + (1 - \lambda^{t_0})/\log \lambda]^2 \\ &= \beta_1 + \{(\beta_{0j} - \beta_0)[(1 - \lambda^{t_0})/\log \lambda + t_0\lambda^{t_0}] \\ &+ (\bar{\beta}_{1j} - \bar{\beta}_1)t_0^2(1 - \lambda^{t_0})\}/[t_0 + (1 - \lambda^{t_0})/\log \lambda]^2. \end{aligned}$$

concluindo-se (4.2.6) e 4.2.7).

Apêndice C

Desenvolvimento de $\text{Var}(\mathbf{y}_j)$

Como $\mathbf{E}(\mathbf{y}_j|\lambda_j) = \mathbf{X}_{j\lambda_j}\boldsymbol{\beta}_{\lambda_j}$, onde $\boldsymbol{\beta}_{\lambda_j} = \mathbf{E}(\boldsymbol{\beta}_j|\lambda_j)$,

$$\text{Var}[\mathbf{E}(\mathbf{y}_j|\lambda_j)] = \mathbf{E}[\mathbf{X}_{j\lambda_j}\boldsymbol{\beta}_{\lambda_j}\boldsymbol{\beta}_{\lambda_j}^t\mathbf{X}_{j\lambda_j}^t] - [\mathbf{E}(\mathbf{y}_j)][\mathbf{E}(\mathbf{y}_j)]^t, \quad (0.0.12)$$

com

$$\mathbf{E}[\mathbf{X}_{j\lambda_j}\boldsymbol{\beta}_{\lambda_j}\boldsymbol{\beta}_{\lambda_j}^t\mathbf{X}_{j\lambda_j}^t] = p\mathbf{X}_{j\lambda}\boldsymbol{\beta}\boldsymbol{\beta}^t\mathbf{X}_{j\lambda}^t + (1-p)\mathbf{X}_{j0}\bar{\boldsymbol{\beta}}\bar{\boldsymbol{\beta}}^t\mathbf{X}_{j0}^t \quad (0.0.13)$$

e

$$\begin{aligned} [\mathbf{E}(\mathbf{y}_j)][\mathbf{E}(\mathbf{y}_j)]^t &= p^2\mathbf{X}_{j\lambda}\boldsymbol{\beta}\boldsymbol{\beta}^t\mathbf{X}_{j\lambda}^t + p(1-p)\mathbf{X}_{j0}\bar{\boldsymbol{\beta}}\bar{\boldsymbol{\beta}}^t\mathbf{X}_{j\lambda}^t \\ &\quad + p(1-p)\mathbf{X}_{j\lambda}\boldsymbol{\beta}\bar{\boldsymbol{\beta}}^t\mathbf{X}_{j0}^t + (1-p)^2\mathbf{X}_{j0}\bar{\boldsymbol{\beta}}\bar{\boldsymbol{\beta}}^t\mathbf{X}_{j0}^t. \end{aligned} \quad (0.0.14)$$

Assim,

$$\begin{aligned} \text{Var}[\mathbf{E}(\mathbf{y}_j|\lambda_j)] &= p(1-p)[\mathbf{X}_{j\lambda}\boldsymbol{\beta}\boldsymbol{\beta}^t\mathbf{X}_{j\lambda}^t + \mathbf{X}_{j0}\bar{\boldsymbol{\beta}}\bar{\boldsymbol{\beta}}^t\mathbf{X}_{j0}^t - \mathbf{X}_{j0}\bar{\boldsymbol{\beta}}\boldsymbol{\beta}^t\mathbf{X}_{j\lambda}^t \\ &\quad - \mathbf{X}_{j\lambda}\boldsymbol{\beta}\bar{\boldsymbol{\beta}}^t\mathbf{X}_{j0}^t] \\ &= p(1-p)[\mathbf{X}_{j\lambda}\boldsymbol{\beta}(\mathbf{X}_{j\lambda}\boldsymbol{\beta} - \mathbf{X}_{j0}\bar{\boldsymbol{\beta}})^t + \mathbf{X}_{j0}\bar{\boldsymbol{\beta}}(\mathbf{X}_{j\lambda}\boldsymbol{\beta} - \mathbf{X}_{j0}\bar{\boldsymbol{\beta}})^t] \\ &= p(1-p)(\mathbf{X}_{j\lambda}\boldsymbol{\beta} - \mathbf{X}_{j0}\bar{\boldsymbol{\beta}})(\mathbf{X}_{j\lambda}\boldsymbol{\beta} - \mathbf{X}_{j0}\bar{\boldsymbol{\beta}})^t. \end{aligned} \quad (0.0.16)$$

Como

$$\mathbf{X}_{j\lambda}\boldsymbol{\beta} - \mathbf{X}_{j0}\bar{\boldsymbol{\beta}} = (\beta_1 - \bar{\beta}_1)\mathbf{t}_j + \beta_1(1 - \lambda^{\mathbf{t}_j})/\log \lambda, \quad (0.0.17)$$

segue que

$$\begin{aligned} \text{Var}[\mathbf{E}(\mathbf{y}_j|\lambda_j)] &= p(1-p)[(\bar{\beta}_1 - \beta_1)\mathbf{t}_j - \beta_1(1 - \lambda^{\mathbf{t}_j})/\log \lambda] \\ &\quad [(\bar{\beta}_1 - \beta_1)\mathbf{t}_j - \beta_1(1 - \lambda^{\mathbf{t}_j})/\log \lambda]^t \\ &= p(1-p)\{(\bar{\beta}_1 - \beta_1)^2\mathbf{t}_j\mathbf{t}_j^t + \beta_1(\bar{\beta}_1 - \beta_1)[(1 - \lambda^{\mathbf{t}_j})\mathbf{t}_j^t \\ &\quad + \mathbf{t}_j(1 - \lambda^{\mathbf{t}_j})^t]/\log \lambda + \beta_1^2(1 - \lambda^{\mathbf{t}_j})(1 - \lambda^{\mathbf{t}_j})^t/(\log \lambda)^2\}. \end{aligned} \quad (0.0. 18)$$

Por outro lado,

$$\begin{aligned} \mathbf{E}[\text{Var}(\mathbf{y}_j|\lambda_j)] &= p\{\mathbf{X}_{j\lambda}\boldsymbol{\Gamma}_\lambda\mathbf{X}_{j\lambda}^t + \text{Var}(\boldsymbol{\kappa}_{j\lambda})\} \\ &\quad + (1-p)\{\mathbf{X}_{j0}\boldsymbol{\Gamma}_0\mathbf{X}_{j0}^t + \text{Var}(\boldsymbol{\kappa}_{j0})\} \\ &= p\{\sigma_0^2\mathbf{1}\mathbf{1}^t - \sigma_{01}[\mathbf{t}_j + (1 - \lambda^{\mathbf{t}_j})/\log \lambda]\mathbf{1}^t \\ &\quad - \sigma_{01}\mathbf{1}[\mathbf{t}_j + (1 - \lambda^{\mathbf{t}_j})/\log \lambda]^t \\ &\quad + \sigma_1^2[\mathbf{t}_j + (1 - \lambda^{\mathbf{t}_j})/\log \lambda][\mathbf{t}_j + (1 - \lambda^{\mathbf{t}_j})/\log \lambda]^t\} \\ &\quad + (1-p)\sigma_0^2\mathbf{1}\mathbf{1}^t - \bar{\sigma}_{01}[\mathbf{t}_j\mathbf{1}^t + \mathbf{1}\mathbf{t}_j^t] + \bar{\sigma}_1^2\mathbf{t}_j\mathbf{t}_j^t\} + \mathbf{V}^* \\ &= \sigma_0^2\mathbf{1}\mathbf{1}^t - p\sigma_{01}[\mathbf{t}_j + (1 - \lambda^{\mathbf{t}_j})/\log \lambda]\mathbf{1}^t - (1-p)\bar{\sigma}_{01}\mathbf{t}_j\mathbf{1}^t \\ &\quad - [p\sigma_{01} + (1-p)\bar{\sigma}_{01}]\mathbf{1}\mathbf{t}_j^t + p\sigma_1^2[\mathbf{t}_j + (1 - \lambda^{\mathbf{t}_j})/\log \lambda]\mathbf{t}_j^t \\ &\quad + p\{\sigma_1^2[\mathbf{t}_j + (1 - \lambda^{\mathbf{t}_j})/\log \lambda] - \sigma_{01}\mathbf{1}^t\}(1 - \lambda^{\mathbf{t}_j})^t/\log \lambda \\ &\quad + (1-p)\bar{\sigma}_1^2\mathbf{t}_j\mathbf{t}_j^t + \mathbf{V}^* \end{aligned} \quad (0.0. 19)$$

onde

$$\mathbf{V}^* = p\text{Var}(\boldsymbol{\kappa}_{j\lambda}) + (1-p)\text{Var}(\boldsymbol{\kappa}_{j0}). \quad (0.0.20)$$

Assim,

$$\begin{aligned} \text{Var}(\mathbf{y}_j) &= \mathbf{E}[\text{Var}(\mathbf{y}_j|\lambda_j)] + \text{Var}[\mathbf{E}(\mathbf{y}_j|\lambda_j)] \\ &= \sigma_0^2\mathbf{1}\mathbf{1}^t - p\sigma_{01}[\mathbf{t}_j + (1 - \lambda^{\mathbf{t}_j})/\log \lambda]\mathbf{1}^t - (1-p)\bar{\sigma}_{01}\mathbf{t}_j\mathbf{1}^t \\ &\quad - [p\sigma_{01} + (1-p)\bar{\sigma}_{01}]\mathbf{1}\mathbf{t}_j^t + p\sigma_1^2[\mathbf{t}_j + (1 - \lambda^{\mathbf{t}_j})/\log \lambda]\mathbf{t}_j^t \end{aligned}$$

$$\begin{aligned}
& + p(1-p)[(\bar{\beta}_1 - \beta_1)^2 \mathbf{t}_j \mathbf{t}_j^t + \beta_1(\bar{\beta}_1 - \beta_1)(\mathbf{1} - \lambda^{\mathbf{t}_j}) \mathbf{t}_j^t / \log \lambda] \\
& + p\{\sigma_1^2[\mathbf{t}_j + (\mathbf{1} - \lambda^{\mathbf{t}_j}) / \log \lambda] - \sigma_{01} \mathbf{1}\}(\mathbf{1} - \lambda^{\mathbf{t}_j})^t / \log \lambda \\
& + p(1-p)\beta_1(\bar{\beta}_1 - \beta_1) \mathbf{t}_j (\mathbf{1} - \lambda^{\mathbf{t}_j})^t / \log \lambda + (1-p)\bar{\sigma}_1^2 \mathbf{t}_j \mathbf{t}_j^t + \mathbf{V}^*
\end{aligned}
\tag{0.0. 21}$$

resultando em (4.3.1).

Apêndice D

Expansão linear de $\hat{t}_{0j}$ e de $\hat{\beta}_{1j}$

t_{0j} é estimado por

$$\hat{t}_{0j} = y_{j0}/\hat{\beta}_{1j} - (1 - \hat{\lambda}_j^{t_{0j}})/\log \hat{\lambda}_j. \quad (0.0.22)$$

Expandindo o lado esquerdo em série de Taylor em torno dos parâmetros verdadeiros obtem-se a aproximação

$$\begin{aligned} \hat{t}_{0j} - t_{0j} &\cong (\hat{\lambda}_j - \lambda_j) \{ t_{0j} \lambda_j^{t_{0j}-1} / \log \lambda_j + (1 - \lambda_j^{t_{0j}}) / [\lambda_j (\log \lambda_j)^2] \} \\ &\quad - (\hat{\beta}_{1j} - \beta_{1j}) y_{j0} / \beta_{1j}^2 + (\hat{t}_{0j} - t_{0j}) \lambda_j^{t_{0j}} \\ (1 - \lambda_j^{t_{0j}})(\hat{t}_{0j} - t_{0j}) &\cong (\hat{\lambda}_j - \lambda_j) \{ t_{0j} \lambda_j^{t_{0j}-1} / \log \lambda_j + (1 - \lambda_j^{t_{0j}}) / [\lambda_j (\log \lambda_j)^2] \} \\ &\quad - (\hat{\beta}_{1j} - \beta_{1j}) y_{j0} / \beta_{1j}^2 \\ \hat{t}_{0j} - t_{0j} &\cong (\hat{\lambda}_j - \lambda_j) \{ t_{0j} \lambda_j^{t_{0j}-1} / [\log \lambda_j (1 - \lambda_j^{t_{0j}})] \\ &\quad + 1 / [\lambda_j (\log \lambda_j)^2] \} - (\hat{\beta}_{1j} - \beta_{1j}) y_{j0} / [\beta_{1j}^2 (1 - \lambda_j^{t_{0j}})] \end{aligned} \quad (0.0. 23)$$

resultando em (4.4.5).

Notando que

$$\begin{aligned} \hat{\beta}_{1j} &= y_{j0}/\hat{t}_{0j} \\ &\cong \bar{\beta}_{1j} - (\hat{t}_{0j} - t_{0j}) y_{j0} / t_{0j}^2 \\ &= \bar{\beta}_{1j} - (\hat{t}_{0j} - t_{0j}) \bar{\beta}_{1j} / t_{0j} \end{aligned}$$

conclui-se (4.4.6).

Referências

- [1] Aitken, A.C. (1935). On least squares and linear combination of observations. **Proceedings of the Royal Society of Edinburgh**, 55: 42-8.
- [2] Anderson, S.J. & R.H. Jones (1991). Smoothing splines for longitudinal data. Technical Report, University of Pittsburgh, Department of Biostatistics.
- [3] Anderson, T.W. (1984). **An Introduction to Multivariate Statistical Analysis - Second Edition**. New York: John Wiley & Sons, Inc.
- [4] Andrade, D.F. & J.M. Singer (1986). **Análise de dados longitudinais**. Texto para minicurso apresentado no 7º SINAPE em Campinas, SP. São Paulo: Associação Brasileira de Estatística.
- [5] Araújo, R.G. (1987). **Influência da iododesoxiuridina, iododesoxicitidina e tri-fluorotimidina sobre a regeneração do epitélio corneano em coelhos: estudo clínico e histopatológico**. Belo Horizonte: Departamento de Oftalmologia e Otorrinolaringologia da Faculdade de Medicina da Universidade Federal de Minas Gerais.
- [6] Azzalini, A. (1987). Growth curves analysis for patterned covariance matrices. In **New Perspectives in Theoretical and Applied Statistics**, ed. M.L. Puri, J.P. Vilaprana & W. Wertz: 63-73. New York: Wiley.
- [7] Beckman, R.J., C.J. Nachstein & R.D. Cook (1987). Diagnostics for mixed-model analysis of variance. **Technometrics**, 29: 413-26.
- [8] Berk, R.H. (1966). Limiting behavior of posterior distribution when the model is incorrect. **The Annals of Mathematical Statistics**, 37: 51-8.
- [9] Berk, R.H. (1970). Consistency a posteriori. **The Annals of Mathematical Statistics**, 41: 1894-906.
- [10] Bertalanffy, L. von (1938). A quantitative theory of organic growth. **Human Biology**, 10: 181-243.
- [11] Box, G.E.P. (1950). Problems in the analysis of growth and wear curves. **Biometrics**, 6: 362-89.
- [12] Box, G.E.P. & G.M. Jenkins (1976). **Time series analysis: forecasting and control (revised edn.)**. San Francisco: Holden Day.
- [13] Breslow, N.E. & D.G. Clayton (1993). Approximate inference in generalized linear mixed models. **Journal of the American Statistical Association**, 88: 9-25.
- [14] Chi, E.M. & G.C. Reinsel (1989). Models for longitudinal data with random effects and AR(1) errors. **Journal of the American Statistical Association**, 84: 452-9.

- [15] Cox, D.R. & D.V. Hinkley (1974). **Theoretical Statistics**. London: Chapman and Hall.
- [16] Cramér, H. (1946). **Mathematical Methods of Statistics**. Princeton: Princeton University Press.
- [17] Crowder, M.J. & J. Hand (1990). **Analysis of repeated measures**. London: Chapman and Hall.
- [18] Dempster, A.P., N.M. Laird & D.B. Rubin (1977). Maximum likelihood from incomplete data via the EM algorithm. **Journal of the Royal Statistical Society, series B**, **39**: 1-38.
- [19] Diggle, P.J. (1988). An approach to the analysis of repeated measurements. **Biometrics**, **44**: 959-71.
- [20] Durrett, R. (1984). **Brownian motion and martingales in analysis**. Belmont: Wadsworth Advanced Books and Software.
- [21] Dwyer, J. (1992). Differential equation models for longitudinal data. Application: blood pressure and relative weight. In **Statistical models for longitudinal studies of health**, eds. J.H. Dwyer, M. Feinleib, P. Lippert and H. Hoffmeister. New York: Oxford University Press: 3-48.
- [22] Dwyer, J.H. & M. Feinleib (1992). Introduction to statistical models for longitudinal observations. In **Statistical models for longitudinal studies of health**, eds. J.H. Dwyer, M. Feinleib, P. Lippert and H. Hoffmeister: 3-48. New York: Oxford University Press.
- [23] Elston, R.C. & J.E. Grizzle (1962). Time response curves and their confidence bands. **Biometrics**, **18**: 148-59.
- [24] Fearn, T. (1975). A Bayesian approach to growth curves. **Biometrika**, **62**: 89-100.
- [25] Fearn, T. (1977). A two-stage model for growth curves wich leads to Rao's covariance adjusted estimators. **Biometrika**, **64**: 141-3.
- [26] Fitzmaurice, G.M. & N.M. Laird (1993). A likelihood-based method for analysing longitudinal binary responses. **Biometrika**, **80**: 141-51.
- [27] Freedman, D.A. & S.C. Peters (1984). Bootstrapping a regression equation: some empirical results. **Journal of the American Statistical Association**, **79**: 97- 106.
- [28] Gabriel, K.R. (1962). Ante-dependence analysis of an ordered set of variables. **Annals of Mathematical Statistics**, **33**: 201-12.
- [29] Gamerman & Migon (1993). Dynamic hierarchical models. **Journal of the Royal Statistics Society, Series B**, **55**: 629-42.

- [30] Gelfand, A.E., S.E. Hills, A. Racine-Poon & A.F.M. Smith (1990). Illustration of Bayesian inference in normal data models using Gibbs sampling. **Journal of the American Statistical Association**, 85: 972-85.
- [31] Gerald, C.F. (1970). **Applied Numerical Analysis**, Reading, MA: Addison-Wesley.
- [32] Grady, J.J. (1992). Structural covariance matrices for incomplete longitudinal data. Institute of Statistics Mimeo Series, 2102T. University of North Carolina at Chapel Hill.
- [33] Graybill, F.A. (1976). **Theory and Application of the Linear Model**. Belmont: Duxbury Press.
- [34] Grizzle, J.E. & D.M. Allen (1969). Analysis of dose and response curves. **Biometrics**, 25: 357-82.
- [35] Harville, D.A. (1974). Bayesian inference for variance components using only error contrasts. **Biometrika**, 61: 383-5.
- [36] Harville, D.A. (1976). Extension of the Gauss-Markov theorem to include the estimation of random effects. **Annals of Statistics**, 4: 384-95.
- [37] Harville, D.A. (1977). Maximum likelihood approaches to variance component estimation and to related problems. **Journal of the American Statistical Association**, 72: 320-40.
- [38] Hartley, H.O. & J.N.K. Rao (1967). Maximum-likelihood estimation for the mixed analysis of variance model. **Biometrika**, 54: 93-108.
- [39] Helms, R.W. & K. McCarroll (1991). Comparisons of approximate F statistics and their small sample distributions for fixed effect hypothesis in mixed models. Submitted for publication.
- [40] Hui, S.L. & J.O. Berger (1983). Empirical Bayes estimation of rates in longitudinal studies. **Journal of the American Statistical Association**, 78: 753-60.
- [41] Izenman, A.J. & J.S. Williams (1989). A class of linear spectral models and analyses for the study of longitudinal data. **Biometrics**, 45: 831-49.
- [42] Jones, R.H. (1990). Serial correlation or random subject effects? **Communications in Statistics - Simulation**, 19: 1105-23.
- [43] Jones, R.H. (1993). **Longitudinal data with serial correlation: a state-space approach**. London: Chapman and Hall.
- [44] Jørgensen, B. (1987). Exponential dispersion models (with discussion). **Journal of the Royal Statistics Society, Series B**, 49: 127-62.

- [45] Jørgensen, B., R. Labouriau & S. Lundbye-Christensen (1991). Growth curve analysis based on exponential dispersion models. Technical report, Institute of Electronic Systems, The University of Aalborg, Denmark.
- [46] Kackar, R.N. & D.A. Harville (1981). Unbiasedness of two-stage estimation and prediction procedures for mixed linear models. **Communications in Statistics - Theory and Methods**, **10**: 1249-61.
- [47] Kackar, R.N. & D.A. Harville (1984). Approximations for standard errors of estimators of fixed and random effects in mixed linear models. **Journal of the American Statistical Society**, **79**: 853-62.
- [48] Kass, R.E. & D. Steffey (1989). Approximate Bayesian inference in conditionally independent hierarchical models (parametric empirical Bayes models). **Journal of the American Statistical Association**, **84**: 717-26.
- [49] Kenward, M.G. (1987). A method for comparing profiles of repeated measurements. **Applied Statistics**, **36**: 296-308.
- [50] Khatri, C.G. (1966). A note on a manova model applied to problems in growth curve. **Annals of the Institute of Statistical Mathematics**, **18**: 75-86.
- [51] Laird, N.M. (1982). Computation of variance components using the EM algorithm. **Journal of Statistical Computation and Simulation**, **14**: 295-303.
- [52] Laird, N.M. (1988). Missing data in longitudinal studies. **Statistics in Medicine**, **7**: 305-15.
- [53] Laird, N.M., N. Lange & D. Stram (1987). Maximum likelihood computations with repeated measures: application of the EM algorithm. **Journal of the American Statistical Association**, **82**: 97-105.
- [54] Laird, N.M. & J.H. Ware (1982). Random-effects models for longitudinal data. **Biometrics**, **38**: 963-74.
- [55] Lange, N. & L. Ryan (1989). Assessing normality in random effects models. **The Annals of Statistics**, **17**: 624-42.
- [56] La Vange, L.M. & R.W. Helms (1983). The analysis of incomplete longitudinal data with time series covariance structures. **Joint Statistical Meetings**, Toronto, Canada.
- [57] Lee, J.C. & S. Geisser (1975). Applications of growth curve prediction. **Sankhyà A**, **37**: 239-56.
- [58] Lee, Y. (1991). Jackknife variance estimators of the location estimator in the one-way random-effects model. **Annals of the Institute of Statistical Mathematics**, **43**: 707-14.

- [59] Liang, K. & S.L. Zeger (1986). Longitudinal data analysis using generalized linear models. **Biometrika**, **73**: 13-22.
- [60] Lindley, D.V. & A.F.M. Smith (1972). Bayes estimates for the linear model (with discussion). **Journal of the Royal Statistical Society, Series B**, **34**: 1-41.
- [61] Lindstrom, M.J. & D.M. Bates (1988). Newton-Raphson and EM algorithms for linear mixed effects-models for repeated-measures data. **Journal of the American Statistical Association**, **83**: 1014-22.
- [62] Lindstrom, M.J. & D.M. Bates (1990). Nonlinear mixed effects models for repeated measures data. **Biometrics**, **46**: 673-87.
- [63] Liski, E.P. (1991). Detecting influential measurements in a growth curve model. **Biometrics**, **47**: 659-68.
- [64] Longford, N.T. (1993). **Random coefficient models**. Oxford: Clarendon Press.
- [65] Lundbye-Christensen, S. (1991). A multivariate growth curve model for pregnancy. **Biometrics**, **47**: 637-57.
- [66] Mandel, J. (1957). Fitting a straight line to certain types of cumulative data. **Journal of the American Statistical Association**, **52**: 552-66.
- [67] Marquardt, D.W. (1963). An algorithm for least squares estimation of nonlinear parameters. **SIAM Journal**, **11**: 431-41.
- [68] Murray, S.C. & R.W. Helms (1990). Linear models with generalized AR(1) covariance structure for irregularly-timed data. Institute of Statistics Mimeo Series, 1876. University of North Carolina at Chapel Hill.
- [69] Palmer, M.J., B.F. Phillips & G.T. Smith (1991). Application of nonlinear models with random coefficients to growth data. **Biometrics**, **47**: 623-35.
- [70] Patterson, H.D. & R. Thompson (1971). Recovery of inter-block information when block sizes are unequal. **Biometrika**, **58**: 545-54.
- [71] Potthoff, R.F. & S.N. Roy (1964). A generalized multivariate analysis of variance model useful especially for growth curve problems. **Biometrika**, **51**: 313-26.
- [72] Rao, C.R. (1959). Some problems involving linear hypotheses in multivariate analysis. **Biometrika**, **46**: 49-58.
- [73] Rao, C.R. (1965). The theory of least squares when the parameters are stochastic and its application to the analysis of growth curves. **Biometrika**, **52**: 447-58.

- [74] Rao, C.R. (1967). Least squares theory using an estimated dispersion matrix and its application to measurement of signals. In **Proceedings of Fifth Berkeley Symposium of Mathematical Statistics and Probability**, 1, ed. L.M. LeCam and J. Neyman: 355-72. Berkeley and Los Angeles, California: University of California Press.
- [75] Rao, C.R. (1987). Prediction of future observations in growth curve models. **Statistical Science**, 4: 434-71.
- [76] Rao, M.N. & C.R. Rao (1966). Linked cross-sectional study for determining norms and growth rates - a pilot survey of Indian school-going boys. **Sankhyā B**, 28: 237-58.
- [77] Ratkowsky, D.A. (1983). **Nonlinear regression modelling: a unified practical approach**. New York: Marcel Dekker.
- [78] Robinson, G.K. (1991). That BLUP is a good thing: the estimation of random effects. **Statistical Science**, 6: 15-51.
- [79] Rochon, J. & R.W. Helms (1989). Maximum likelihood estimation for incomplete repeated-measures experiments under an ARMA covariance structure. **Biometrics**, 45: 207-18.
- [80] Rosner, B. & A. Muñoz (1988). Autoregressive modelling for the analysis of longitudinal data with unequally spaced examinations. **Statistics in Medicine**, 7: 59-71.
- [81] Rosner, B. & A. Muñoz (1992). Conditional linear models for longitudinal data. Application: cigarette smoking and respiratory function in adolescents . In **Statistical models for longitudinal studies of health**, eds. J.H. Dwyer, M. Feinleib, P. Lipert and H. Hoffmeister. New York: Oxford University Press.
- [82] Royall, R.M. (1986). Model robust confidence intervals using maximum likelihood estimators. **International Statistical Review**, 54: 221-6.
- [83] Rutter, C.M. & R.M. Elashoff (1992). Analysis of longitudinal data: random coefficient regression modelling. Technical report, UCLA Department of Biomathematics.
- [84] Ryan, Jr., T.A., B.L. Joiner, B.F. Ryan (1981). **Minitab Reference Manual**. Statistics Department, Pennsylvania State University.
- [85] Sandland, R.L. & C.A. McGilchrist (1979). Stochastic growth curve analysis. **Biometrics**, 35: 255-71.
- [86] Schluchter, M.D. (1988). Analysis of incomplete multivariate data using linear models with structured covariance matrices. **Statistics in Medicine**, 7: 317-24.
- [87] Schluchter, M.D. (1990). Unbalanced repeated measures models with structured covariance matrices. In **BMDP statistical software manual, 1990 software release**, vol. 2, eds. Los Angeles: BMDP Statistical Software, Inc.

- [88] Schluchter, M.D. (1992). Likelihood-based methods for the analysis of continuous longitudinal data. 1992 Meeting of the Western North American Region of the Biometrics Society and the Institute of Mathematical Statistics, Corvallis, Oregon.
- [89] Stiratelli, R., N.M. Laird & J.H. Ware (1984). Random-effects models for serial observations with binary response. **Biometrics**, 40: 961-71.
- [90] Suyama, E., R.G. Araújo & S. Aguiar (1986). Efeito dos medicamentos antivirais na recuperação do epitélio corneano. In **Atas do 7º SINAPE, Campinas, SP**. São Paulo: Associação Brasileira de Estatística, 283-96.
- [91] Taylor, J.M.G. & W.G. Cumberland (1991). A stochastic model for analysis of longitudinal data. Technical report, UCLA Department of Biostatistics.
- [92] Vonesh, E.F. & R.L. Carter (1992). Mixed-effects nonlinear regressions for unbalanced repeated measures. **Biometrics**, 48: 1-17.
- [93] Ware, J.H. (1985). Linear models for the analysis of longitudinal studies. **American Statistician**, 39: 95-101.
- [94] Ware, J.H. & V. De Gruttola (1985). Multivariate linear models for longitudinal data: a bootstrap study of the GLS estimator. In **Biostatistics - Statistics in Biomedical, Public Health and Environmental Science (the Bernard Greenberg Volume)**, ed. P.K. Sen. Amsterdam: Elsevier.
- [95] Waternaux, C., N.M. Laird & J.H. Ware (1989). Methods of analysis of longitudinal data: blood-lead concentrations and cognitive development. **Journal of the American Statistical Association**, 84: 33-41.
- [96] Weiss, R.E. & C.G. Lazaro (1992). Residual plots for repeated measures. **Statistics in Medicine**, 11: 115-24.
- [97] White (1982). Maximum likelihood estimation of misspecified models. **Econometrica**, 50: 1-25.
- [98] Wishart, J. (1938). Growth-rate determinations in nutrition studies with the bacon pig, and their analysis. **Biometrika**, 30: 16-28.
- [99] Zeger, S.L. & M.R. Karim (1991). Generalized linear models with random effects; a Gibbs sampling approach. **Journal of the American Statistical Association**, 86: 79-86.
- [100] Zerbe, G.O. (1979). Randomization analysis of the completely randomized design extended to growth and response curves. **Journal of the American Statistical Association**, 74: 215-21.