ANÁLISE DE CURVAS DE CRESCIMENTO

JULIO DA MOTTA SINGER

DISSERTAÇÃO APRESENTADA

AO

INSTITUTO DE MATEMÁTICA E ESTATÍSTICA

DA

UNIVERSIDADE DE SÃO PAULO

PARA OBTENÇÃO DO GRAU DE MESTRE

EM

ESTATÍSTICA

ORIENTADOR:

PROF. DR. ADOLPHO WALTER PIMAZONI CANTON

- SÃO PAULO, NOVEMBRO DE 1977 -

- o -

A minha esposa,
Maria Lúcia,
e a meus pais,
Edith e José.

AGRADECIMENTOS

Queremos agradecer a todos os que direta ou indiretamente colaboraram para a realização deste trabalho, em particular

ao Professor Doutor Adolpho Walter Pimazoni Canton, orientador e amigo, pela paciência e dedicação com que transmitiu seus conhecimentos;

aos Professores Doutores Clóvis de Araujo Peres e Euclides Custódio de Lima Filho, pelas sugestões apresentadas durante a elaboração deste trabalho;

à Professora Maria Takishita, pela colaboração na parte de computação eletrônica;

ao Professor Carlos Alberto de Bragança Pereira, pela amizade e incentivo constante;

aos amigos do Instituto de Matemática e Estatística da Universidade de São Paulo, e

ao Senhor João Baptista Esteves de Oliveira pela dedicação no trabalho de datilografia.

Julio da Motta Singer

ÍNDICE

1. INTRODUÇÃO. 1

1.1. O Problema - Modelos Matemáticos 1
1.2. Exemplos de Aplicação do Modelo Linear Multivariado de Crescimento. 11
1.3. Considerações Gerais 16
1.4. Desenvolvimento Histórico. 22

2. O PROBLEMA DE ESTIMAÇÃO EM MODELOS LINEARES MULTIVARIADOS DE CRESCIMENTO 25

2.1. A Solução da Proposta por Potthoff e Roy 26
2.2. Estimação com Ajuste por Covariância 28
2.3. Uma Aplicação da Técnica de Ajuste por Covariância ao Modelo Linear Multivariado 33
2.4. A Solução Proposta por Rao 36
2.5. A Escolha de Variáveis Auxiliares. 42

3. O PROBLEMA DE TESTES DE HIPÓTESES EM MODELOS LINEARES MULTIVARIADOS DE CRESCIMENTO 46

3.1. Um Teste de Ajuste do Modelo 46
3.2. Testes de Hipóteses Lineares Sobre os Parâmetros do Modelo 48
3.3. Intervalos de Confiança. 50

4. APLICAÇÕES. 53

4.1. Sugestão para Análise de um Modelo Linear Multivariado de Crescimento utilizando um Programa Convencional de Análise de Variância Multivariada. 53
4.2. Exemplos 57

5. CONCLUSÃO 71

APÊNDICES

A. O PROBLEMA DE ESTIMAÇÃO EM MODELOS LINEARES MULTIVARIADOS . . . 73

A.1. Uma Aplicação do Critério de Mínimos Quadrados ao Modelo Linear Multivariado 74

A.2. Uma Aplicação do Critério de Máxima Verossimilhança ao Modelo Linear Multivariado 77

A.3. Determinação de Estimadores Lineares não Viciados de Mínima Variância 83

A.4. Distribuição dos Estimadores Obtidos 86

B. O PROBLEMA DE TESTES DE HIPÓTESES EM MODELOS LINEARES MULTIVARIADOS. 88

B.1. O Princípio da União-Intersecção de Roy. 89

B.2. O Princípio da Razão de Verossimilhança Generalizada de Wilks. 99

B.3. Outros Critérios de Teste; Considerações Gerais. 106

BIBLIOGRAFIA 110

REFERÊNCIAS. 113

JB

CAPÍTULO I

INTRODUÇÃO

A finalidade deste capítulo é familiarizar o leitor com o tipo de problema analisado neste trabalho, com os modelos matemáticos adotados e com a notação empregada.

Na seção 1.1 introduzimos os modelos matemáticos fundamentais considerados na análise de curvas de crescimento através da discussão de um problema prático.

Na seção 1.2 generalizamos a situação prática apresentada em 1.1 em diversos sentidos, salientando a flexibilidade da aplicação da técnica que pretendemos descrever.

Na seção 1.3 fazemos algumas suposições adicionais sobre os modelos apresentados em 1.1 visando uma simplificação formal dos capítulos subsequentes.

Finalmente na seção 1.4 apresentamos um resumo do desenvolvimento histórico do estudo de ajuste de curvas de crescimento, citando os principais trabalhos nesta área.

1.1 - O PROBLEMA — MODELOS MATEMÁTICOS

Em muitos campos de pesquisa é comum querermos estudar o comportamento no tempo, de uma ou mais variáveis, me

didas em unidades experimentais submetidas a um certo tratamento. Também é frequente a necessidade de comparar a influência de diferentes tratamentos no comportamento dessas variáveis.

O modelo matemático mais usado nessas situações consiste na descrição do comportamento da variável em função do tempo, através de uma curva, geralmente de forma polinomial. Comparações entre efeitos de tratamentos diferentes podem ser estudadas através de comparações dos parâmetros das respectivas curvas. A estimação dos parâmetros determinantes das curvas envolvidas e testes de hipóteses estatísticas sobre esses parâmetros, ou comparações de parâmetros relativos a curvas diferentes são obtidos a partir de observações realizadas em grupos de unidades experimentais.

Como exemplo, consideremos a seguinte situação: um pesquisador deseja estudar a influência de três tipos de ração no peso de uma certa espécie de animal durante um determinado período. Com essa finalidade, 30 animais são divididos em 3 grupos e cada grupo é alimentado com um tipo diferente de ração. O peso de cada animal é medido diariamente durante o período pré-determinado. Com base nessas observações, o pesquisador deseja determinar, por exemplo, que tipo de curva explica melhor o comportamento da variável peso para cada grupo, se as curvas são diferentes para cada grupo e quais as estimativas para os parâmetros envolvidos.

Como na situação acima, na maioria das vezes é vantajoso ou até mesmo necessário que as observações sejam realizadas em uma mesma unidade experimental nos diversos instantes considerados. O mesmo acontece quando essas medidas são observadas após a aplicação de doses crescentes de uma certa droga em uma mesma unidade experimental, ou quando as observações são realizadas a diversas distâncias de um certo ponto, como observações de intensidade luminosa ou sono-

ra a diversas distâncias de uma fonte.

O fato de as medidas serem realizadas em uma mesma unidade experimental implica, geralmente, a correlação entre as observações, negando dessa forma, a suposição de independência necessária para a plicação dos modelos clássicos de Análise de Variância e Regressão na avaliação estatística dos resultados. Uma possível correlação entre as observações realizadas em uma mesma unidades experimental deve ser considerada na análise. Modelos apropriados para essas situações podem ser obtidos através de uma generalização da teoria de Análise de Variância Multivariada, sob o título geral de Análise de Curvas de Crescimento.

O problema considerado pode ser mais explicitamente compreendido através da discussão de um exemplo apresentado por Potthoff e Roy, (1964).

EXEMPLO 1.1.1 - Num estudo realizado na Faculdade de Odontologia da Universidade da Carolina do Norte (E.U.A.), foram medidas as distâncias (em mm) entre o centro da pituitária e a fissura pterio-maxilar em 11 meninas e 16 meninos aos 8, 10, 12 e 14 anos. Em Ortodontia é importante conhecer a variação dessa distância com a idade e seria conveniente ajustar curvas polinomiais (reta ou parábola, por exemplo) às observações, além de comparar as curvas obtidas para os dois grupos (meninas e meninos). As medidas observadas são apresentadas na Tabela 1.1.1 e as médias para as diversas idades estão representadas graficamente na Figura 1.1.1.

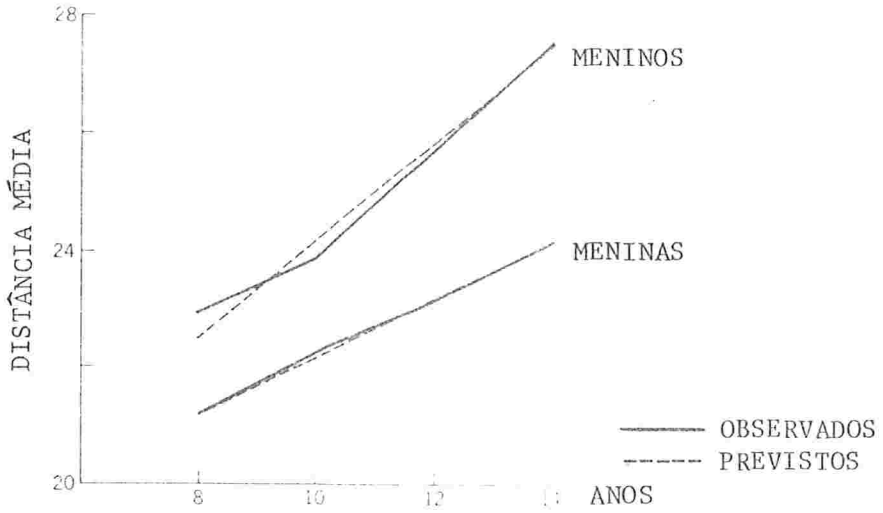
Quando temos, a priori, informação sobre a forma das curvas a serem ajustadas às observações, podemos considerar diretamente o problema de estimar os coeficientes envolvidos, e testar as hipóteses estatísticas de interesse. Essa informação pode ser obtida tanto através de considerações teóricas (quando conhecemos a equação que rege o fenômeno estudado) ou através de uma análise do gráfico de dis-

TABELA 1.1.1

DADOS OBTIDOS POR PESQUISADORES NA "UNIVERSITY OF NORTH CAROLINA DENTAL SCHOOL" — CADA MEDIDA REPRESENTA A DISTÂNCIA EM MILÍMETROS ENTRE O CENTRO DA PITUITÁRIA E FISSURA PTERIO-MAXILAR

	MENINAS					MENINOS			
	8	10	12	14		8	10	12	14
1	21	20	21,5	23	1	26	25	29	31
2	21	21,5	24	25,5	2	21,5	22,5	23	26,5
3	20,5	24	24,5	26	3	23	22,5	24	27,5
4	23,5	24,5	25	26,5	4	25,5	27,5	26,5	27
5	21,5	23	22,5	23,5	5	20	23,5	22,5	26
6	20	21	21	22,5	6	24,5	25,5	27	28,5
7	21,5	22,5	23	25	7	22	22	24,5	26,5
8	23	23	23,5	24	8	24	21,5	24,5	25,5
9	20	21	22	21,5	9	23	20,5	31	26
10	16,5	19	19	19,5	10	27,5	28	31	31,5
11	24,5	25	28	28	11	23	23	23,5	25
					12	21,5	23,5	24	28
					13	17	24,5	26	29,5
					14	22,5	25,5	25,5	26
					15	23	24,5	26	30
					16	22	21,5	23,5	25
MÉDIAS	21,18	22,23	23,09	24,09		22,87	23,81	25,72	27,47
PREVISTAS	21,24	22,19	23,14	24,09		22,46	24,11	25,76	27,41

FIGURA 1.1.1



persão das observações, quanto através do resultado de análises passadas, pois no caso mais geral, um estudo desse tipo raramente envolve um único experimento isolado. Quando, por outro lado, essa informação sobre a forma das curvas a serem ajustadas inexiste, podemos obtê-la através de uma análise preliminar das observações, por exemplo estudando o problema do ponto de vista de Análise de Variância Multivariada.

Embora uma análise do gráfico de dispersão das observações do Exemplo 1.1.1 mostre que o comportamento no tempo da variável possa ser razoavelmente explicado através de uma reta, procuraremos obter essa informação através de uma análise preliminar, com o objetivo de introduzir o modelo clássico de Análise de Variância Multivariada.

Consideremos r grupos cada qual com n_i unidades experimentais independentes, $n_1+n_2+\ldots+n_r=n$. Suponhamos que em cada unidade experimental são realizadas p medidas de uma certa variável, nos instantes $t_1, t_2, \ldots, t_p$. Então o modelo clássico de Análise de Variância Multivariada, pode ser escrito como:

$$\underset{(n\times p)}{Y} = \underset{(n\times r)}{X} \quad \underset{(r\times p)}{\beta} + \underset{(n\times p)}{\varepsilon} \qquad (1.1.1)$$

onde

- Y $(n\times p)$ é a matriz das observações (p medidas para cada uma das n unidades experimentais).
- X $(n\times r)$ onde $r<n$ é a matriz de planejamento (constituída por 0's e 1's, de forma a associar cada unidade experimental ao respectivo grupo).
- β $(r\times p)$ é a matriz de parâmetros desconhecidos (cada linha é constituída por p elementos que representam os valores esperados da variável observada em cada um dos instantes considerados para um

determinado grupo).

ε $(n \times p)$ é a matriz dos componentes aleatórios das observações (cada elemento representa o desvio entre o valor observado e o respectivo valor esperado).

β e ε são matrizes não-observáveis.

Uma suposição fundamental do modelo é que as linhas da matriz ε sejam não correlacionadas e obedeçam a uma distribuição normal multivariada com vetor de valores esperados $0(1 \times p)$ e matriz de variância-covariância não singular Σ $(p \times p)$, comum a todas as linhas. Então:

$$E(Y) = X\beta \qquad (1.1.2)$$

O modelo descrito dessa forma será denominado "Modelo Linear Multivariado" e utilizaremos a notação

$$MLM(Y, X\beta, \Sigma)$$

quando quisermos indicar que as definições e suposições apresentadas acima são válidas para um determinado conjunto de dados.

Tomemos os dados apresentados no Exemplo 1.1.1 e consideremos o $MLM(Y, X\beta, \Sigma)$: Nesse caso temos

$$n = 27 \ (n_1 = 11 \text{ e } n_2 = 16), \ p = 4 \text{ e } r = 2.$$

As matrizes X e β têm as seguintes formas:

$$X = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ \vdots & \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ \vdots & \\ 0 & 1 \end{bmatrix} \begin{matrix} \left.\vphantom{\begin{matrix}1\\1\\\vdots\\1\end{matrix}}\right\} \text{11 meninas} \\ \left.\vphantom{\begin{matrix}0\\0\\\vdots\\0\end{matrix}}\right\} \text{16 meninos} \end{matrix}$$

$$\begin{array}{cc} & \text{idade} \quad 8 \quad 10 \quad 12 \quad 14 \\ \beta = & \begin{bmatrix} \beta_{11} & \beta_{12} & \beta_{13} & \beta_{14} \\ \beta_{21} & \beta_{22} & \beta_{23} & \beta_{24} \end{bmatrix} \begin{matrix} \text{meninas} \\ \text{meninos} \end{matrix} \end{array}$$

Podemos estimar a matriz de valores esperados β aplicando qualquer um dos critérios apresentados no apêndice A e testar a hipótese de que o valor esperado da variável considerada é constante no tempo para cada grupo (hipótese fundamental de Análise de Variância). Se essa hipótese for rejeitada podemos ainda testar hipóteses sobre a tendência de crescimento (linear, quadrática, etc...) da variável para os diversos grupos. As hipóteses citadas podem ser formuladas através de hipóteses equivalentes sobre constrastes da forma $C\beta U$, onde $C(c\times r)$, $c \leq r$, $\rho(C) = c$ e $U(p\times u)$, $u \leq p$, $\rho(U) = u$ são matrizes conhecidas escolhidas convenientemente. Alguns critérios de teste são apresentados no apêndice B.

O resultado dos testes sobre tendência de crescimento pode sugerir a forma da curva que mais se adapta às observações, e então nosso interesse se concentrará na estimação dos coeficientes que determinam essas curvas e em testes de hipóteses estatísticas sobre esses coeficientes. É óbvio que, se tivermos informação *a priori* sobre a forma das curvas de crescimento envolvidas esses procedimentos preliminares de análise devem ser omitidos.

Com a finalidade de continuar a análise após haver decidido sobre a forma das curvas a serem consideradas, vamos introduzir o modelo linear multivariado de crescimento que pode ser escrito na forma:

$$\underset{(n\times p)}{Y} = \underset{(n\times r)}{X} \; \underset{(r\times q)}{\xi} \; \underset{(q\times p)}{G} + \underset{(n\times p)}{\varepsilon} \qquad (1.1.3)$$

onde

Y $(n\times p)$ é a matriz das observações - descrita em (1.1.1).

X $(n\times r)$ é a matriz de planejamento - descrita em (1.1.1).

ξ $(r\times q)$ é a matriz dos coeficientes (desconhecidos) das curvas de crescimento (cada linha tem q elementos que representam os coeficientes da curva correspondente a um determinado grupo).

G $(q\times p)$, $q\leq p$ é a matriz conhecida, geralmente formada por coeficientes de polinômios ortogonais.

ε (n×p) é a matriz dos componentes aleatórios das observações - descrita em (1.1.1).

Também para este modelo supomos que as linhas da matriz ε sejam não correlacionadas e obedeçam a uma distribuição normal multivariada com vetor de média **0** (1×p) e matriz de variância-covariância não singular Σ (p×p), comum a todas as linhas. Então:

$$E(Y) = X \xi G \qquad (1.1.4)$$

Utilizaremos para o modelo linear multivariado de crescimento, a notação MLMC$(Y, X\xi G, \Sigma)$.

Tomemos novamente os dados do Exemplo 1.1.1 e vamos testar a hipótese fundamental de Análise de Variância, ou seja, que o valor esperado da variável para cada grupo seja constante no tempo. Essa hipótese pode ser escrita na forma:

$$H_1: C_1 \beta U_1 = 0 \text{ onde } C_1 = I_2 \text{ e } U_1 = \begin{bmatrix} 1 & 1 & 1 \\ -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

Utilizando um dos critérios apresentados no apêndice B a hipótese H_1 deve ser rejeitada com qualquer nível de significância razoável, ou seja, podemos concluir que existe um efeito do tempo sobre o crescimento da variável.

Podemos então testar a hipótese de que o crescimento da variável considerada é linear no tempo para os dois grupos. Essa hipótese pode ser escrita na forma:

$$H_2: C_2 \beta U_2 = 0, \text{ onde } C_2 = I_2 \text{ e } U_2 = \begin{bmatrix} -3 \\ -1 \\ 1 \\ 3 \end{bmatrix}$$

(como os dados são igualmente espaçados, a matriz U_2 foi obtida a partir de uma tabela de coeficientes de polinômios ortogonais). Também nesse caso a hipótese H_2 deve ser rejeitada, indicando que podemos admitir o modelo linear de crescimento (para os dois grupos). Vamos então considerar o se-

guinte modelo de crescimento:

$$\beta_{it} = \xi_{i0} + \xi_{i1}\, t \qquad (1.1.5)$$

onde:

β_{it} é o valor esperado da variável para o i-ésimo grupo no instante t.
ξ_{i0} é o coeficiente linear da reta relativa ao i-ésimo grupo.
ξ_{i1} é o coeficiente angular da reta relativa ao i-ésimo grupo. i=1,2 (meninas e meninos).

Podemos escrever o modelo (1.1.5) na forma do MLMC(Y,XξG,Σ), fazendo:

$$X = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ \vdots & \\ 1 & 0 \\ 0 & 1 \\ \vdots & \\ 0 & 1 \end{bmatrix} \begin{matrix} \left.\begin{matrix} \\ \\ \\ \\ \end{matrix}\right\} \text{11 meninas} \\ \left.\begin{matrix} \\ \\ \\ \end{matrix}\right\} \text{16 meninos} \end{matrix} \qquad \xi = \begin{array}{c} \begin{matrix} \text{coef.} & \text{coef.} \\ \text{lin.} & \text{ang.} \end{matrix} \\ \begin{bmatrix} \xi_{10} & \xi_{11} \\ \\ \xi_{20} & \xi_{21} \end{bmatrix} \end{array} \begin{matrix} \text{meninas} \\ \\ \text{meninos} \end{matrix}$$

$$G = \begin{array}{c} \begin{matrix} t_1 & t_2 & t_3 & t_4 \end{matrix} \\ \begin{bmatrix} 1 & 1 & 1 & 1 \\ \\ 8 & 10 & 12 & 14 \end{bmatrix} \end{array} \begin{matrix} \text{coef. lin.} \\ \\ \text{coef. ang.} \end{matrix}$$

Outras formas do modelo linear de crescimento podem ser consideradas, por exemplo:

$$\beta_{it} = \xi^*_{io} + \xi^*_{i1}(t-\bar{t}) \qquad (1.1.6)$$

Além disso, há considerável simplificação dos cálculos, quando a matriz G é formada por coeficientes de polinômios ortogonais. Optando pelo modelo (1.1.6) o MLMC(Y,Xξ*G*,Σ) seria definido por:

X idêntica ao caso anterior.

$$\xi^* = \begin{bmatrix} \xi^*_{10} & \xi^*_{11} \\ \xi^*_{20} & \xi^*_{21} \end{bmatrix} \qquad G^* = \begin{bmatrix} 1 & 1 & 1 & 1 \\ -3 & -1 & 1 & 3 \end{bmatrix}$$

Aplicando os resultados obtidos no capítulo 2, obtemos uma estimativa da matriz ξ^*:

$$\hat{\xi}^* = \begin{bmatrix} 22,6477 & 0,4795 \\ 24,9688 & 0,7848 \end{bmatrix}$$

As médias previstas através do modelo linear descrito acima estão apresentadas na Tabela 1.1.2 e representadas na Figura 1.1.1.

TABELA 1.1.1

MÉDIAS PREVISTAS PARA O EXEMPLO 1.1.1

IDADE	MENINAS	MENINOS
8	21,24	22,46
10	22,19	24,11
12	23,14	25,76
14	24,09	27,41

Podemos então testar hipóteses estatísticas sobre os coeficientes das curvas ajustadas. Por exemplo, poderíamos estar interessados em saber se há diferença significativa entre as taxas de crescimento da variável considerada para os dois grupos; isto pode ser verificado através de um teste de igualdade dos coeficientes angulares:

$$H_3: C_3\xi^*U_3 = 0 \text{ onde } C_3 = [1 \ -1] \text{ e } U_3 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Utilizando um dos métodos descritos no apêndice B,

a hipótese H_3 deve ser rejeitada e então concluímos que há evidência experimental de que o crescimento da variável estudada não pode ser descrito pela mesma curva para crianças de sexos diferentes.

Como veremos nos capítulos seguintes, podemos construir intervalos de confiança para cada um dos coeficientes considerados ou para combinações lineares envolvendo esses coeficientes, bem como construir testes de hipóteses para o ajuste do modelo às observações.

1.2 - EXEMPLOS DE APLICAÇÃO DO MODELO LINEAR MULTIVARIADO DE CRESCIMENTO

As principais aplicações do modelo linear multivariado de crescimento, como já vimos na seção 1.1, envolvem problemas em que medidas de uma ou mais variáveis são realizadas numa mesma unidade experimental em diversos instantes. Nesta seção discutiremos algumas situações apresentadas por Potthoff e Roy,(1964) em que o MLMC$(Y, X\xi G, \Sigma)$ é aplicável, descrevendo cada um dos elementos do modelo e formulando algumas hipóteses nas quais podemos estar interessados.

i) Suponhamos um grupo de n animais submetidos às mesmas condições experimentais. Em cada animal são realizadas p medidas de uma certa variável (peso,por exemplo) nos instantes $t_1, t_2, \ldots, t_p$, e desejamos descrever o comportamento desta variável com o tempo através de uma curva polinomial de grau $k \leq p-1$. Podemos então representar o valor esperado da variável através da expressão: $\beta_t = \xi_0 + \xi_1 t + \ldots + \xi_k t^k$. Admitindo que para cada animal as p observações obedeçam a uma distribuição normal multivariada com matriz de variância covariância não singular Σ desconhecida, poderemos considerar o MLMC$(Y, X\beta G, \Sigma)$ onde:

$$\underset{(n\times 1)}{X} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \qquad \underset{(1\times k+1)}{\xi} = [\xi_0 \quad \xi_1 \ldots \xi_k] \qquad \underset{(k+1\times p)}{G} = \begin{bmatrix} 1 & 1 & \ldots & 1 \\ t_1 & t_2 & \ldots & t_p \\ t_1^2 & t_2^2 & \ldots & t_p^2 \\ \vdots & \vdots & & \vdots \\ t_1^k & t_2^k & & t_p^k \end{bmatrix}$$

Quando as p observações de cada unidade experimental são realizadas em instantes igualmente espaçados, a matriz G, acima, pode ser substituída com vantagem por uma matriz formada pelos coeficientes de polinômios ortogonais de graus 0,1,...,k.

A hipótese de que os coeficientes dos termos de grau maior que 2 são nulos pode ser testada através da hipótese equivalente:

$$H_1: C_1 \xi U_1 = 0 \text{ onde } \underset{(1\times 1)}{C_1} = 1 \text{ e } \underset{(k+1\times k+1)}{U_1} = \begin{bmatrix} 0 & 0 & \ldots & 0 \\ 0 & 0 & \ldots & 0 \\ 0 & 0 & \ldots & 0 \\ 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & 1 \end{bmatrix}$$

Podemos ainda construir intervalos de confiança para os parâmetros envolvidos e para a curva considerada.

ii) Suponhamos agora r grupos de animais, cada grupo contendo n_i elementos ($n_1+n_2+\ldots n_r=n$) e submetido a um certo tratamento (tratamentos diferentes para grupos diferentes). Consideremos p medidas de uma certa variável realizadas em cada animal nos instantes $t_1, t_2, \ldots, t_p$. Desejamos descrever o comportamento desta variável com o tempo através de uma curva polinomial de grau $k \leq p-1$ para cada grupo. Neste caso o valor esperado da variável no instante t para animais do

i-ésimo grupo poderá ser expresso por: $\beta_{it} = \xi_{i0} + \xi_{i1}t + \ldots + \xi_{ik}t^k$. Admitindo que as considerações sobre a distribuição da variável de i) sejam válidas para a variável em estudo, podemos considerar o MLMC$(Y, X\xi G, \Sigma)$, onde:

$$\underset{(n\times r)}{X} = \left|\begin{array}{cccc} 1 & 0 & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 1 & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & 1 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & 1 \end{array}\right| \begin{array}{l} \left.\begin{array}{c} \\ \\ \\ \end{array}\right\} n_1 \\ \left.\begin{array}{c} \\ \\ \\ \end{array}\right\} n_2 \\ \\ \\ \left.\begin{array}{c} \\ \\ \\ \end{array}\right\} n_r \end{array} \qquad \underset{(r\times k+1)}{\xi} = \begin{bmatrix} \xi_{10} & \xi_{11} & \cdots & \xi_{1k} \\ \xi_{20} & \xi_{21} & \cdots & \xi_{2k} \\ \vdots & \vdots & & \vdots \\ \xi_{r0} & \xi_{r1} & & \xi_{rk} \end{bmatrix}$$

G definida em i).

A hipótese de que as curvas de crescimento de todos os grupos são iguais pode ser testada através de:

$$H_2: C_2\xi U_2 = 0 \text{ onde: } \underset{(r-1\times r)}{C_2} = \begin{bmatrix} 1 & -1 & 0 & \ldots & 0 \\ 1 & 0 & -1 & \ldots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & 0 & 0 & & -1 \end{bmatrix} \text{ e } U_2 = I_k$$

A hipótese de que os coeficientes das curvas de crescimento de todos os grupos são iguais, com exceção dos coeficientes dos termos de ordem zero, pode ser testada através de:

$$H_3: C_3\xi U_3 = 0 \text{ onde: } C_3 = C_2 \text{ e } \underset{(k\times k)}{U_3} = \begin{bmatrix} 0 & 0 & \ldots & 0 \\ 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & 1 \end{bmatrix}$$

Outras hipóteses relativas aos coeficientes envolvidos podem ser testadas através de hipóteses sobre contrastes da forma $C\xi U$.

iii) A situação apresentada em ii) pode ser generalizada através da consideração de um esquema de planejamento envolvendo dois ou mais fatores cruzados. Suponhamos um problema com dois fatores: tipo de ração (com 3 níveis) e temperatura ambiente (com 2 níveis). Vamos ainda supor que existam n_{ij} animais sujeitos às condições definidas pela combinação da i-ésima ração com o j-ésimo nível do fator temperatura ambiente, e que não exista interação entre esses fatores. Desejamos definir a curva de crescimento da variável observada considerando o valor esperado dessa variável no instante t, para animais do grupo definido pela combinação da i-ésima ração com a j-ésima temperatura, expresso através da equação:

$$\beta_{ijt} = (\mu_{i0}+\mu_{i1}t+\ldots+\mu_{ik}t^k) + (\nu_{j0}+\nu_{j1}t+\ldots+\nu_{jk}t^k).$$

Admitindo as mesmas suposições consideradas nos exemplos anteriores para a distribuição da variável em questão, podemos utilizar o MLMC$(Y, X\xi G, \Sigma)$, onde:

$$\underset{(n\times 5)}{X} = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 1 & 0 & 1 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 1 & 0 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 1 & 0 & 1 \end{bmatrix} \begin{matrix} \left.\vphantom{\begin{matrix}1\\ \vdots\\ 1\end{matrix}}\right\} n_{11} \\ \left.\vphantom{\begin{matrix}1\\ \vdots\\ 1\end{matrix}}\right\} n_{12} \\ \left.\vphantom{\begin{matrix}1\\ \vdots\\ 1\end{matrix}}\right\} n_{21} \\ \\ \left.\vphantom{\begin{matrix}1\\ \vdots\\ 1\end{matrix}}\right\} n_{32} \end{matrix}$$

$$\underset{(5\times k+1)}{\xi} = \begin{bmatrix} \mu_{10} & \mu_{11} & \cdots & \mu_{1k} \\ \mu_{20} & \mu_{21} & \cdots & \mu_{2k} \\ \mu_{30} & \mu_{31} & \cdots & \mu_{3k} \\ \nu_{10} & \nu_{11} & \cdots & \nu_{1k} \\ \nu_{20} & \nu_{21} & \cdots & \nu_{2k} \end{bmatrix}$$

$\underset{(k+1,p)}{G}$ definida em i).

Hipóteses análogas às testadas em ii) podem ser consideradas. Por exemplo, a hipótese de que não há diferença entre os efeitos das temperaturas ambientes, pode ser testada através de:

$$H_4: C_4 \xi U_4 = 0 \text{ onde } \underset{(1\times 5)}{C_4} = [0 \quad 0 \quad 0 \quad 1 \quad -1] \text{ e } U_4 = I_{k+1}$$

iv) Finalmente consideremos outra generalização da situação apresentada em ii). Novamente sejam r grupos de animais, onde cada animal é submetido a medidas nos instantes $t_1, t_2, \ldots, t_p$. Neste caso, porém, realizamos observações de duas ou mais variáveis em cada instante. Suponhamos, por exemplo, observações de peso e altura de cada animal, e vamos definir as curvas de crescimento através das seguintes expressões para os valores esperados do peso e altura dos animais do i-ésimo grupo no instante t:

$$\beta_{it} = \delta_{i0} + \delta_{i1} t + \ldots + \delta_{ik_i} t^{k_1-1} \qquad (k_1 \leq p)$$

$$\beta'_{it} = \gamma_{i0} + \gamma_{i1} t + \ldots + \gamma_{ik_2} t^{k_2-1} \qquad (k_2 \leq p),\ k_1+k_2=k.$$

Admitindo que as 2p observações realizadas em cada animal obedeçam a uma distribuição normal miltivariada com matriz de variância-covariância não singular Σ (2p×2p), desconhecida, podemos considerar o MLMC($Y, X\xi G, \Sigma$), onde a matriz das observações é construída de forma que as primeiras p colunas contenham as medidas da variável peso e as últimas p colunas, as medidas da variável altura. Então:

$\underset{(n\times r)}{X}$ é definida como em ii).

$$\underset{(r\times k)}{\xi} = \begin{bmatrix} \delta_{10} & \delta_{11} & \cdots & \delta_{1,k_1-1} & \gamma_{10} & \gamma_{11} & \cdots & \gamma_{1,k_2-1} \\ \delta_{20} & \delta_{21} & \cdots & \delta_{2,k_1-1} & \gamma_{20} & \gamma_{21} & \cdots & \gamma_{2,k_2-1} \\ \vdots & \vdots & & \vdots & \vdots & & & \vdots \\ \delta_{r0} & \delta_{r1} & \cdots & \delta_{r,k_1-1} & \gamma_{r0} & \gamma_{r1} & \cdots & \gamma_{r,k_2-1} \end{bmatrix}$$

$$\underset{(k\times 2p)}{G} = \begin{bmatrix} 1 & 1 & \dots & 1 & 0 & 0 & \dots & 0 \\ t_1 & t_2 & \dots & t_p & 0 & 0 & \dots & 0 \\ \vdots & \vdots & & & \vdots & \vdots & & \\ t_1^{k_1-1} & t_2^{k_1-1} & \dots & t_p^{k_1-1} & 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 1 & 1 & \dots & 1 \\ 0 & 0 & \dots & 0 & t_1 & t_2 & \dots & t_p \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & & 0 & t_1^{k_2-1} & t_2^{k_2-1} & \dots & t_p^{k_2-1} \end{bmatrix}$$

A hipótese de que as curvas de crescimento da variável altura são iguais para todos os grupos pode ser formulada através de:

$$H_5: C_5 \xi U_5 = 0 \text{ onde } \underset{(r-1\times r)}{C_5} = \begin{bmatrix} 1 & -1 & 0 & \dots & 0 \\ 1 & 0 & -1 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & 0 & 0 & \dots & -1 \end{bmatrix} \quad \underset{(k\times k)}{U_5} = \begin{bmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 0 \\ 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix}$$

O modelo acima pode ser facilmente modificado para ser aplicado a uma situação em que as observações da variável peso são efetuadas nos instantes $t_{11}, t_{12}, \dots, t_{1p}$ e as observações da variável altura nos instantes $t_{21}, t_{22}, \dots, t_{2p}$, não havendo necessidade de que os dois conjuntos de pontos sejam iguais.

1.3 - CONSIDERAÇÕES GERAIS

Apresentamos nesta seção algumas suposições adicio

nais sobre os elementos dos modelos

$$MLM(Y, X\beta, \Sigma) \quad e \quad MLMC(Y, X\beta Y, \Sigma),$$

além de algumas considerações sobre a utilização dos mesmos. As suposições adicionais são consideradas sem que haja perda de generalidade dos modelos, e têm a finalidade de simplificar formalmente as deduções. Quando admitirmos válidos os modelos acima para um conjunto de dados, estaremos considerando válidas as suposições apresentadas nesta seção, além da suposição fundamental citada na seção 1.1.

A primeira suposição a ser introduzida é que a matriz de planejamento X tenha posto-coluna completo, ou seja $\rho(X) = r$. Se o modelo sugerido inicialmente for tal que a matriz de planejamento corresponde não tenha posto-coluna completo, é possível obter uma reparametrização de forma que a matriz de planejamento do novo modelo apresente a propriedade desejada. A reparametrização pode ser realizada segundo diversos métodos, sendo desejável que os novos parâmetros obtidos sejam interpretáveis. Consideremos um exemplo.

Suponhamos um experimento com dois fatores completamente cruzados: A, com 3 níveis e B, com 2 níveis. Queremos testar hipóteses estatísticas relativas ao efeito de cada um desses fatores em uma variável a partir de um conjunto de observações realizadas nos instantes $t_1, t_2, \ldots, t_p$, e hipóteses estatísticas sobre a interação entre esses fatores. O modelo paramétrico usual para esse tipo de situação é dado por:

$$y_{ijk} = \mu + a_i + b_j + ab_{ij} + \varepsilon_{ijk},$$

onde $i=1,2,3$; $j=1,2$; $k=1,2,\ldots,n_{ij}$; y_{ijk} representa o vetor de p observações correspondente ao k-ésimo elemento submetido ao i-ésimo nível do fator A e j-ésimo nível do fator B; μ, a_i, b_i e ab_{ij} têm os significados usuais. O $MLM(Y, X\beta, \Sigma)$

considerado para este problema seria definido por:

$$\underset{(n\times 12)}{X} = \begin{bmatrix} 1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{matrix} \left.\vphantom{\begin{matrix}1\\ \vdots\\ 1\end{matrix}}\right\} n_{11} \\ \left.\vphantom{\begin{matrix}1\\ \vdots\\ 1\end{matrix}}\right\} n_{12} \\ \vphantom{\begin{matrix}1\\ \vdots\\ 1\end{matrix}} \\ \left.\vphantom{\begin{matrix}1\\ \vdots\\ 1\end{matrix}}\right\} n_{32} \end{matrix}$$

$$\underset{(12\times p)}{\beta} = \begin{bmatrix} \mu_1 & \mu_2 & \cdots & \mu_p \\ a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ a_{31} & a_{32} & \cdots & a_{3p} \\ b_{11} & b_{12} & \cdots & b_{1p} \\ b_{21} & b_{22} & \cdots & b_{2p} \\ ab_{111} & ab_{112} & \cdots & ab_{11p} \\ \vdots & \vdots & & \vdots \\ ab_{321} & ab_{322} & \cdots & ab_{32p} \end{bmatrix}$$

X não tem posto coluna completo neste caso, pois a primeira coluna é uma combinação linear das demais. Então os parâmetros envolvidos não são estimáveis, apesar de algumas combinações lineares dos mesmos o serem. Condições para estimabilidade de combinações lineares dos parâmetros do modelo são apresentadas por Searle (1971), por exemplo. Satisfeitas essas condições, o problema pode ser resolvido através da utilização da Teoria de Matrizes Inversas Generalizadas. Outras soluções formalmente mais simples podem ser obtidas através

da reparametrização do modelo inicial. Consideremos, por exemplo o modelo reparametrizado: $y_{\ell k} = \beta_\ell + \varepsilon_{\ell k}$, onde $\ell=1,2,\ldots,6$ e β_ℓ representa o vetor de valores esperados da variável considerada para as unidades amostrais submetidas à ℓ-ésima combinação dos níveis dos fatores A e B. Então o

$$MLM(Y, X\beta, \Sigma)$$

poderia ser definido por:

$$\underset{(n\times 6)}{X} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 1 & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{matrix} \left.\vphantom{\begin{matrix}1\\1\\\vdots\\1\end{matrix}}\right\} n_{11} \\ \left.\vphantom{\begin{matrix}1\\1\\\vdots\\1\end{matrix}}\right\} n_{12} \\ \\ \left.\vphantom{\begin{matrix}1\\1\\\vdots\\1\end{matrix}}\right\} n_{32} \end{matrix} \; ; \; \underset{(6\times p)}{\beta} = \begin{bmatrix} \beta_{11} & \beta_{12} & \cdots & \beta_{1p} \\ \beta_{21} & \beta_{22} & \cdots & \beta_{2p} \\ \vdots & \vdots & & \vdots \\ \beta_{61} & \beta_{62} & \cdots & \beta_{6p} \end{bmatrix} \begin{matrix} A_1B_1 \\ A_1B_2 \\ \vdots \\ A_3B_2 \end{matrix}$$

Às custas de uma redução do número de parâmetros do modelo (e de uma consequente perda de detalhe paramétrico) conseguimos um novo modelo no qual a matriz de planejamento tem posto-coluna completo.

Vejamos agora como formular as hipóteses sugeridas na definição do problema utilizando o modelo reparametrizado.

A hipótese de que o efeito médio do fator A é constante para todos os seus níveis pode ser formulada como:

$$H_1: C_1\beta U_1 = 0 \text{ onde } C_1 = \begin{bmatrix} 1 & 1 & -1 & -1 & 0 & 0 \\ 1 & 1 & 0 & 0 & -1 & -1 \end{bmatrix} \text{ e } U_1 = I_p.$$

Analogamente podemos formular a hipótese de que o

efeito médio do fator B é constante para todos os seus níveis, através de:

$$H_2: C_2\beta U_2 = 0 \text{ onde } C_2 = [1 \quad -1 \quad 1 \quad -1 \quad 1 \quad -1] \text{ e } U_2 = I_p.$$

Finalmente a hipótese de que não existe interação entre os fatores A e B pode ser formulada como:

$$H_3: C_3\beta U_3 = 0 \text{ onde } C_3 = \begin{bmatrix} 1 & -1 & -1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 & -1 & 1 \end{bmatrix} \text{ e } U_3 = I_p.$$

Desse modo conseguimos atingir os objetivos desejados na proposição do problema, conservando a característica de uma fácil interpretação dos resultados e sem necessitar o emprego da Teoria de Matrizes Inversas Generalizadas, que apesar de gerar resultados equivalentes, envolve dificuldades teóricas e computacionais.

A construção das matrizes C (c×r) e U (p×u) utilizadas nos testes de hipóteses sobre os parâmetros do modelo é relativamente simples, bastando lembrar que a matriz C se refere a hipóteses envolvendo as linhas da matriz β (tratamentos) e a matriz U a hipóteses sobre as colunas de β (respostas). Em geral, para hipóteses de interesse prático, podemos sempre considerar a matriz C de posto-linha completo, $\rho(C)=c$, e a matriz U de posto-coluna completo, $\rho(U) = u$ pois assumimos X de posto-coluna completo.

Existem outros métodos de reparametrização que possibilitam a obtenção de modelos nos quais a matriz de planejamento tem posto-coluna completo. Um desses métodos considera a restrição de que a soma dos efeitos médios de cada um dos fatores é nula

$$\left(\sum_{i=1}^{k} a_i = 0 \right).$$

Um outro método é conhecido por "Técnica da Casela de Referência", no qual os parâmetros são definidos em relação ao ao valor esperado da variável para uma determinada casela, tomada como referência. Como o estudo de reparametrização de modelos utilizados em planejamento de experimentos não é um dos objetivos deste trabalho, não nos aprofundaremos mais nesse tópico. Convém observar que o método apresentado nesta seção é aconselhável tanto pela sua simplicidade, como pelo fato de ser facilmente aplicável à maioria dos problemas usuais. A aplicação do método ao MLMC$(Y, X\beta G, \Sigma)$ é feita de maneira análoga.

Outra suposição a ser introduzida no MLMC$(Y, X\beta G, \Sigma)$ é que a matriz G tenha posto-linha completo, $\rho(G) = q$. Novamente lembramos que se o modelo inicial de planejamento for tal que a matriz G não tenha posto-linha completo, podemos redefinir a matriz de parâmetros ξ de tal forma que essa suposição seja satisfeita.

Apresentamos em seguida um resumo das suposições assumidas quando consideramos o MLM$(Y, X\beta, \Sigma)$ válido para um determinado conjunto de observações:

i) o modelo é escrito na forma: $\underset{(n\times p)}{Y} = \underset{(n\times r)}{X} \underset{(r\times p)}{\beta} + \underset{(n\times p)}{\varepsilon}$.

ii) as linhas da matriz ε são não correlacionadas entre si, e cada uma obedece a uma distribuição normal multivariada com vetor de valores esperados 0 $(1\times p)$ e matriz de variância-covariância não singular Σ $(p\times p)$.

iii) a matriz de planejamento X $(n\times r)$, $r < n$ tem posto-coluna completo, $\rho(X) = r$.

iv) hipóteses estatísticas sobre os parâmetros do modelo são formuladas na forma:

H: $C\beta U = \Gamma$ onde C $(c\times r)$, $c \leq r$ é tal que $\rho(C) = c$

e U $(p\times u)$, $u \leq p$ é tal que $\rho(U) = u$.

v) $n \geq p+r$ (ver seção B.1).

Analogamente para MLMC$(Y, X\beta G, \Sigma)$ temos:

vi) o modelo é escrito na forma: $\underset{(n\times p)}{Y} = \underset{(n\times r)}{X} \; \underset{(r\times q)}{\xi} \; \underset{(q\times p)}{G} + \underset{(n\times p)}{\varepsilon}$

vii) vale a suposição ii).

viii) vale a suposição iii).

ix) hipóteses estatísticas sobre os parâmetros do modelo são formuladas na forma H: $C\beta U = \Gamma$ onde C $(c\times r)$, $c \leq r$ é tal que $\rho(C) = c$ e U $(q\times u)$, $u \leq q$ é tal que $\rho(U) = u$.

x) $q \leq p$.

Uma restrição do MLMC$(Y, X\beta G, \Sigma)$ é que as curvas ajustadas aos diversos grupos devem pertencer à mesma família. Por exemplo, se o valor esperado da variável considerada para um determinado grupo for obtido através de uma curva polinomial de grau k, para os demais grupos deveremos também considerar curvas polinomiais de grau k, apesar de os coeficientes não precisarem ser iguais.

Por outro lado, a matriz de planejamento X não precisa ser necessariamente constituída por 0's e 1's. Na realidade a matriz X pode ser tal que seus elementos correspondam a valores assumidos por variáveis contínuas como no problema analisado sob o título de Regressão Múltipla (univariada e multivariada). Podemos ainda considerar casos em que a matriz X é tal que algumas colunas são constituídas por 0's e 1's, e outras são formadas por valores assumidos por variáveis auxiliares como no problema analisado sob o título de Análise de Covariância (univariada e multivariada). Como veremos no capítulo 2, a solução proposta por Rao para Análise de Curvas de Crescimento envolve esse tipo de matriz.

1.4 - DESENVOLVIMENTO HISTÓRICO — UM RESUMO

Trabalhos envolvendo aspectos específicos ou casos

particulares do problema de estimação e obtenção de testes de significância sobre curvas polinomiais de crescimento foram desenvolvidos por autores como: Wishart (1938), Box (1950), Rao (1958), Leech e Healy (1959), Rao (1959), Healy (1961), Elston e Grizzle (1962) e Bock (1963).

A atual linha de pesquisas nessa área tem como fundamento o trabalho de Potthoff e Roy (1964), no qual os autores obtiveram resultados aplicáveis a uma considerável classe de situações através de uma generalização do modelo clássico de Análise de Variância Multivariada. O método sugerido introduz na análise uma matriz de ponderação arbitrária Δ^{-1}, e os autores observam que apesar de o estimador $\hat{\xi}$ obtido ser não viciado independentemente da escolha da matriz Δ^{-1}, sua variância aumenta conforme Δ^{-1} "se afasta" de Σ^{-1} (que é desconhecida). Embora algumas sugestões para a escolha de Δ^{-1} tenham sido apresentadas, não foi considerado o caso em que $\Delta^{-1} = S^{-1}$, onde S é uma das estimativas usuais de Σ calculada a partir das observações. Khatri (1966) determinou o estimador de máxima verossimilhança de ξ, observando que sua expressão é semelhante àquela obtida por Potthoff e Roy com S^{-1} no lugar de Δ^{-1}. Rao (1965) analisou a solução de Potthoff e Roy, mostrando que a mesma não utiliza toda a informação contida nas observações (a menos que $\Delta^{-1} = \Sigma^{-1}$), e sugerindo outra solução, na qual são consideradas como variáveis auxiliares os (p-q) vetores pertencentes ao espaço de erros, introduzindo na análise, dessa forma, a informação anteriormente desprezada. Continuando na mesma linha, Rao, (1966, 1967) mostrou que nem sempre há um aumento de precisão dos resultados ao serem considerados como variáveis auxiliares todos os (p-q) vetores do espaço de erros, e que em função da correlação entre as variáveis auxiliares e as variáveis principais há vantagem em introduzir na análise apenas um subconjunto dos vetores do espaço de erros.

Ainda nesses trabalhos o autor apresenta sugestões relativas à escolha de variáveis auxiliares baseadas em suposições sobre a estrutura de Σ. Grizzle e Allen (1969) discutem a aplicação prática dos resultados desenvolvidos por Rao, fornecendo sugestões para a escolha de variáveis auxiliares a partir de uma análise preliminar das observações. Finalmente esses autores apresentam uma sugestão para a aplicação prática do método através da utilização de um programa convencional de Análise de Variância Multivariada.

- □ -

CAPÍTULO 2

O PROBLEMA DE ESTIMAÇÃO EM MODELOS LINEARES MULTIVARIADOS DE CRESCIMENTO

Consideramos neste capítulo o problema de obtenção de estimadores de ξ, de combinações lineares da forma $C\xi U$ e de Σ a partir de um conjunto de dados para os quais admitimos válidas as suposições do MLMC$(Y, X\xi G, \Sigma)$.

Na seção 2.1 apresentamos um resumo do artigo de Potthoff e Roy (1964) que deu origem à atual linha de pesquisa em Modelos Lineares Multivariados de Crescimento. Esses autores sugerem uma transformação de variável com a finalidade de recair no problema de estimação dos parâmetros do MLM$(Y^*, X\xi, \Sigma^*)$. Essa transformação depende de uma matriz arbitrária Δ^{-1}, e algumas sugestões sobre sua escolha são discutidas.

Na seção 2.2 introduzimos, de uma maneira geral, o conceito de ajuste de um estimador por covariância. Mostramos que, muitas vezes, ao estimar um parâmetro a partir de um conjunto de observações, convém utilizar a informação contida em variáveis auxiliares que têm valor esperado nulo. Mostramos também que o ganho de eficiência (no sentido de variância) do estimador ajustado por covariância depende da co

variância entre as variáveis auxiliares e o estimador considerado.

Na seção 2.3 aplicamos a técnica de ajuste de um estimador por covariância ao problema de estimação em modelos lineares multivariados.

Na seção 2.4 obtemos formalmente os estimadores de ξ, $C\xi U$ e Σ considerando o MLMC$(Y, X\xi G, \Sigma)$ através de uma transformação de variável sugerida por Rao (1965, 1966). Fazemos algumas considerações sobre a distribuição dos estimadores obtidos, e finalmente apresentamos uma crítica (Rao, 1965, 1966) à solução proposta por Potthoff e Roy (1964), em que o autor mostra que não é utilizada toda a informação contida na amostra, a não ser que $\Delta^{-1} = \Sigma^{-1}$. (Do ponto de vista prático isto não é viável pois Σ é desconhecida.)

Na seção 2.5 discutimos o problema da escolha das variáveis auxiliares a serem consideradas, em função de diferentes estruturas da matriz de variância - covariância Σ (Rao, 1967).

2.1 - A SOLUÇÃO PROPOSTA POR POTTHOFF E ROY

Consideremos o MLMC$(Y, X\xi G, \Sigma)$. Seja Δ $(p \times p)$ qualquer matriz simétrica positiva definida, ou qualquer matriz não singular tal que $G\Delta^{-1}G'$ $(q \times q)$ tenha posto completo, e façamos a seguinte transformação:

$$Y^* = Y\Delta^{-1}G'(G\Delta^{-1}G')^{-1} \qquad (2.1.1)$$

Conforme as suposições do MLMC$(Y, X\xi G, \Sigma)$, podemos concluir que as linhas da matriz Y^* $(n \times q)$ definida acima são mutuamente independentes e obedecem a uma distribuição normal multivariada com vetor de valores esperados obtido a partir de:

$$E(Y^*) = X\xi G\Delta^{-1}G'(G\Delta^{-1}G')^{-1} = X\xi \qquad (2.1.2)$$

e matriz de variância-covariância, não singular, dada por:

$$\underset{(q\times q)}{\Sigma^*} = (G\Delta'^{-1}G')^{-1}G\Delta'^{-1}\Sigma\Delta^{-1}G'(G\Delta^{-1}G')^{-1} \qquad (2.1.3)$$

A matriz de observações transformadas Y^* satisfaz então às suposições do $MLM(Y^*, X\xi, \Sigma^*)$ independentemente da escolha da matriz Δ (sujeita às restrições acima). Estimadores de ξ, $C\xi U$ e Σ, e testes de hipóteses estatísticas sobre os parâmetros da matriz ξ podem ser obtidos através dos resultados dos apêndices A e B. Um estimador não viciado de ξ, por exemplo:

$$\hat{\xi} = (X'X)^{-1}X'Y^* = (X'X)^{-1}X'Y\Delta^{-1}G'(G\Delta^{-1}G')^{-1} \qquad (2.1.4)$$

Embora a técnica descrita acima possa ser aplicada qualquer que seja a matriz Δ (sujeita às restrições apresentadas), a escolha dessa matriz *afeta* a variância dos estimadores, a precisão dos intervalos de confiança e o poder dos testes de hipóteses. Algumas considerações sobre a escolha da matriz Δ são apresentadas a seguir.

i) Se $p = q$ então G é inversível, e podemos considerar a transformação $Y^* = YG^{-1}$, não havendo necessidade da introdução da matriz arbitrária Δ.

ii) Uma maneira simples de escolher a matriz Δ é fazer $\Delta = I$. Esta solução tem a vantagem de simplificar consideravelmente os cálculos. Na realidade, as simplificações conseguidas quando a matriz G é formada por coeficientes de polinômios ortogonais só existirão se tomarmos $\Delta = I$.

iii) Se considerarmos o $MLMC(Y, X\xi G, \Sigma)$ com a matriz Σ conhecida, o estimador de $b'C\xi Uf$, não viciado de mínima variância na classe dos estimadores da forma $d'Yw$, onde b, f, d e w são vetores, é dado por:

$$\hat{\gamma} = b'C(X'X)^{-1}X'Y\Sigma^{-1}G'(G\Sigma^{-1}G')^{-1}Uf \qquad (2.1.5)$$

(Potthoff e Roy, 1964, apêndice B).

Comparando (2.1.4) com (2.1.5), parece razoável admitir que $\Delta = \Sigma$ representa a escolha ótima da matriz Δ (segundo o critério de obtenção de estimadores não viciados de variância mínima).

Como na realidade a matriz Σ é desconhecida, não podemos fazer a escolha $\Delta = \Sigma$. No entanto, se dispusermos de uma estimativa razoável de Σ, é melhor considerar Δ igual a essa estimativa do que tomar $\Delta = I$. Devemos observar que, embora a técnica apresentada seja válida para qualquer escolha de Δ (sujeita às restrições apresentadas no início desta seção), podemos suspeitar que menor será o poder dos testes, maior a variância dos estimadores, e menos precisos os intervalos de confiança quanto maior a "diferença" entre Δ e Σ.

A matriz Δ não precisa necessariamente ser não-estocástica. Podemos escolher Δ estocástica desde que seus elementos sejam independentes de Y, pois as distribuições relacionadas com a estimação e testes de hipóteses sobre os parâmetros envolvidos não dependem matematicamente da matriz Δ. Não parece válido considerar Δ baseada numa estimativa de Σ obtida a partir das observações Y. Porém se dispusermos de informação experimental sobre Σ proveniente de experimentos independentes do experimento que gerou Y podemos utilizar essa informação para estimar Σ e tomar Δ igual a essa estimativa.

2.2 - ESTIMAÇÃO COM AJUSTE POR COVARIÂNCIA

Vamos considerar nesta seção, o problema geral de estimação de um vetor de parâmetros a partir de um conjunto de observações, quando temos informação adicional obtida a

partir de variáveis auxiliares que têm valores esperados nulos. A aplicação da técnica ao caso específico de Modelos Lineares Multivariados será discutida na seção 2.3.

Sejam t_1 $(k\times 1)$ e t_2 $(r\times 1)$ duas estatísticas vetoriais tais que $E(t_1) = \tau$ e $E(t_2) = 0$, onde τ é um vetor de k parâmetros desconhecidos, que desejamos estimar. Suponhamos também que a matriz de variância-covariância Σ do vetor:

$$\begin{bmatrix} t_1 \\ t_2 \end{bmatrix}, \quad (k+r\times 1),$$

seja não singular e conhecida. Vamos supor ainda que Σ seja particionada segundo:

$$\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}, \text{ onde } V(t_1) = \Sigma_{11} \text{ (não singular)}, \; V(t_2) = \Sigma_{22}$$

(não singular) e $Cov(t_1, t_2) = \Sigma_{12}$ (2.2.1)

A estatística t_1 é um estimador não viciado de τ e tem matriz de variância-covariância não singular Σ_{11}. Mostremos então que se $Cov(t_1, t_2) \neq 0$, podemos determinar um estimador de τ mais eficiente (no sentido de variância) que t_1.

Consideremos o seguinte estimador de τ:

$$\tau^* = t_1 - \Sigma_{12}\Sigma_{22}^{-1}t \qquad (2.2.2)$$

e calculemos sua matriz de variância-covariância:

$$V(\tau^*) = V(t_1) + \Sigma_{12}\Sigma_{22}^{-1}V(t_2)\Sigma_{22}^{-1}\Sigma_{21} - \Sigma_{12}\Sigma_{22}^{-1}\, Cov(t_2, t_1) -$$

$$- Cov(t_1, t_2)\Sigma_{22}^{-1}\Sigma_{21} \Longrightarrow$$

$$\Longrightarrow V(\tau^*) = \Sigma_{11} + \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} - 2\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} \qquad (2.2.3)$$

Então $V(t_1) - V(\tau^*) = \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$, que é não negativa definida (n.n.d) ou seja:

$$x'[V(t_1)-V(\tau^*)]x \geq 0,\ \forall x \in R^k \Longrightarrow x'V(t_1)x \geq x'V(\tau^*)x,\ \forall x \in R^k.$$

Nesse sentido dizemos que o estimador τ^* é um estimador mais eficiente que t_1.

Na maioria dos problemas reais, a matriz de variância-covariância Σ é desconhecida, porém podemos obter uma estimativa:

$$\hat{\Sigma} = \begin{bmatrix} \hat{\Sigma}_{11} & \hat{\Sigma}_{12} \\ \hat{\Sigma}_{21} & \hat{\Sigma}_{22} \end{bmatrix}$$

Substituindo Σ_{ij} por $\hat{\Sigma}_{ij}$ em (2.2.2), obteremos o estimador "ajustado por covariância":

$$\hat{\tau} = t_1 - \hat{\Sigma}_{12}\hat{\Sigma}_{22}^{-1}t_2 \qquad (2.2.4)$$

Se $\hat{\Sigma}_{ij}$ for distribuída independentemente de t_i $(i,j=1,2)$, teremos:

$$E(\hat{\tau}) = E(t_1) - E(\hat{\Sigma}_{12}\hat{\Sigma}_{22}^{-1})E(t_2) = E(t_1) = \tau$$

Logo o estimador $\hat{\tau}$ é não viciado. Determinemos a matriz de variância-covariância deste estimador:

$$V(\hat{\tau}) = V(t_1) + E[\hat{\Sigma}_{12}\hat{\Sigma}_{22}^{-1}V(t_2)\hat{\Sigma}_{22}^{-1}\hat{\Sigma}_{21}] - E[\hat{\Sigma}_{12}\hat{\Sigma}_{22}^{-1}\,Cov(t_1,t_2)] -$$

$$- E[Cov(t_1,t_2)\hat{\Sigma}_{22}^{-1}\hat{\Sigma}_{21}] \Longrightarrow \qquad (2.2.5)$$

$$\Longrightarrow V(\hat{\tau}) = \Sigma_{11} + E[\hat{\Sigma}_{12}\hat{\Sigma}_{22}^{-1}\Sigma_{22}\hat{\Sigma}_{22}^{-1}\hat{\Sigma}_{21}] - E[\hat{\Sigma}_{12}\hat{\Sigma}_{22}^{-1}\Sigma_{21}] - E[\Sigma_{12}\hat{\Sigma}_{22}^{-1}\hat{\Sigma}_{21}]$$

onde as esperanças são consideradas em relação às variáveis $\hat{\Sigma}_{ij}$.

Não mais podemos garantir, como no caso do estimador τ^*, que $V(t_1) - V(\hat{\tau})$ seja sempre n.n.d. Vamos então considerar duas hipóteses:

i) $\Sigma_{21} = 0$ (ou próxima de 0). Neste caso esperamos que: $V(t_1) - V(\hat{\tau})$ seja negativa definida (n.d) e então t_1 é um estimador mais eficiente que $\hat{\tau}$.

ii) Σ_{21} é "bem diferente de 0". Então esperamos que $V(t_1) - V(\tau)$ seja n.n.d. e $\hat{\tau}$ é um estimador mais eficiente que t_1.

Na seção 4.2 apresentaremos exemplos onde esses aspectos serão discutidos com mais detalhe do ponto de vista prático.

De i) e ii) concluímos que o ajuste por covariância de um estimador pode resultar num decréscimo de eficiência, e que a decisão entre aplicar ou não essa técnica depende da covariância entre as variáveis auxiliares e as variáveis principais. Mesmo que t_2 não se constitua na escolha ótima das variáveis auxiliares para o ajuste por covariância, é possível que uma seleção conveniente de elementos de t_2 como variáveis auxiliares possa fornecer estimadores ajustados mais eficientes. Como a matriz Σ é desconhecida, não podemos determinar a escolha ótima; no entanto, o conhecimento da estrutura da matriz Σ, ou uma estimativa da mesma podem nos fornecer meios para nos aproximarmos dessa escolha ótima. Esse problema será considerado na seção 2.5 (teoricamente) e na seção 4.1 (praticamente).

Com a finalidade de estudar as distribuições dos estimadores ajustados por covariância, vamos supor que:

$$\begin{bmatrix} t_1 \\ t_2 \end{bmatrix} \sim N_{k+r} \left[\begin{pmatrix} \tau \\ 0 \end{pmatrix}, \Sigma \right] \text{ e } f\hat{\Sigma} \sim W_{k+r}(f, \Sigma, \cdot) \text{ central.}$$

Então demonstra-se (Rao, 1967) que as distribuições condi-

cionais de

$$\hat{\tau} = t_1 - \hat{\Sigma}_{12}\hat{\Sigma}_{22}^{-1}t_2 \quad e \quad \hat{\Sigma}_{11}^* = \hat{\Sigma}_{11} - \hat{\Sigma}_{12}\hat{\Sigma}_{22}^{-1}\hat{\Sigma}_{21}$$

dados t_2 e $\hat{\Sigma}_{22}$ são independentes e são dadas por:

$$\begin{aligned} \hat{\tau} &\sim N_k[\tau, \Sigma_{11}^*(1+T_r^2)] \\ \hat{\Sigma}_{11}^* &\sim W_k(f-r, \Sigma_{11}^*, \cdot) \text{ central} \end{aligned} \qquad (2.2.6)$$

onde

$$\Sigma_{11}^* = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} \quad e \quad T_r^2 = \frac{1}{f}\, t_2'\hat{\Sigma}_{22}^{-1}t_2$$

Além disso pode-se demonstrar (Rao, 1973, por exemplo) que:

$$\frac{f-r+1}{2}\, T_r^2 \sim F(r, f-r+1)$$

Então

$$E(1+T_r^2) = 1 + \frac{r}{f-r+1} \times \frac{f-r+1}{f-r-1} = 1 + \frac{r}{f-r-1} = \frac{f-1}{f-r-1} \qquad (2.2.7)$$

Com base nos resultados (2.2.6) e (2.2.7) podemos determinar a matriz de variância-covariância do estimador $\hat{\tau}$, apresentada simbolicamente em (2.2.5):

$$V(\hat{\tau}) = E[V(\hat{\tau}|t_2, \hat{\Sigma}_{22})] + V[E(\hat{\tau}|t_2, \hat{\Sigma}_{22})] =$$

$$= E[(1+T_r^2)\Sigma_{11}^*] + V(\tau) = E(1+T_r^2)\Sigma_{11}^* = \frac{f-1}{f-r-1}\,\Sigma_{11}^* \qquad (2.2.8)$$

Quando Σ é conhecida, a matriz de variância-covariância do estimador τ^* é dada por Σ_{11}^*, portanto a perda de eficiência que ocorre quando utilizamos uma estimativa de Σ para o ajuste por covariância é de $\frac{r}{f-r-1}$. Além disso, se $\Sigma_{12} = 0$, teremos $\Sigma_{11}^* = \Sigma_{11}$ e então t_1 será mais eficiente que $\hat{\tau}$, quaisquer que sejam f e r.

2.3 - UMA APLICAÇÃO DA TÉCNICA DE AJUSTE POR COVARIÂNCIA AO MODELO LINEAR MULTIVARIADO

Consideremos agora o problema de estimação com ajuste por covariância em modelos lineares multivariados. Seja uma matriz de observações Y $(n \times q)$ para a qual admitimos as suposições do $MLM(Y, X\beta, \Sigma_{11})$. Seja também uma matriz de observações auxiliares Z $(n \times p-q)$, para a qual admitimos válidas as suposições do $MLM(Z, 0, \Sigma_{22})$.

Se ignorarmos as observações Z, os resultados do apêndice A nos permitem concluir que um estimador não viciado de $C\beta$ é dado por: $C\hat{\beta} = C(X'X)^{-1}X'Y$. A matriz de variância-covariância desse estimador pode ser representada por:

$$V(C\hat{\beta}) = \Sigma_{11} \otimes [C(X'X)^{-1}C'] \qquad (2.3.1)$$

Um estimador não viciado de $C\beta U$ podem ser obtidos através da transfomação $Y^* = YU$.

Vamos agora determinar outro estimador de $C\beta$ considerando a informação contida nas variáveis auxiliares. Tomemos então a matriz de observações $[Y\ Z]$ $(n \times p)$ para a qual podemos admitir válidas as suposições do $MLM([Y\ Z], X[\beta\ 0], \Sigma)$, onde:

$$\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}$$

Consideremos a distribuição condicional de Y dadas as observações da variável auxiliar Z. Utilizando resultados conhecidos sobre a distribuição normal multivariada (Rao, 1973, por exemplo), podemos concluir que as linhas da matriz Y (dadas as observações da variável auxiliar Z) obedecem a distribuições normais multivariadas independentes, com vetor de valores esperados obtidos a partir de $E(Y|Z) = X\beta + Z\Sigma_{22}^{-1}\Sigma_{21}$ e

matriz de variância-covariância dada por:

$$\Sigma_{11}^{*} = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}.$$

Fazendo

$$D = [X \;\; Z], \quad \eta = Z\Sigma_{22}^{-1}\Sigma_{21} \quad e \quad \gamma = \begin{bmatrix} \beta \\ \eta \end{bmatrix},$$

o modelo condicional acima é equivalente ao $MLM(Y, D\gamma, \Sigma_{11}^{*})$, e um estimador não viciado de $C^{*}\gamma$ é dado por:

$$C^{*}\hat{\gamma} = C^{*}(D'D)^{-1}D'Y.$$

Vamos obter a expressão do estimador $C^{*}\hat{\gamma}$ em função de C^{*}, Y X e Z utilizando a técnica de obtenção da inversa de uma matriz particionada (Rao, 1973, por exemplo). Devemos observar que a matriz D terá posto coluna $(r+p-q)$ em q.t.p. pois Z é uma variável aleatória multidimensional contínua.

$$D'D = \begin{bmatrix} X' \\ Z' \end{bmatrix} [X \;\; Z] = \begin{bmatrix} X'X & X'Z \\ Z'X & Z'Z \end{bmatrix}$$

$$(D'D)^{-1} = \begin{bmatrix} (X'X)^{-1} + (X'X)^{-1}X'ZS_{22}^{-1}Z'X(X'X)^{-1} & -(X'X)^{-1}X'ZS_{22}^{-1} \\ -S_{22}^{-1}Z'X(X'X)^{-1} & S_{22}^{-1} \end{bmatrix}$$

onde $S_{22} = Z'[I - X(X'X)^{-1}X']Z$.

Como

$$D'Y = \begin{bmatrix} X' \\ Z' \end{bmatrix} Y,$$

obtemos o estimador

$$C^*\hat{\gamma} = C^*\begin{bmatrix}\hat{\beta}\\ \hat{\eta}\end{bmatrix} = C^*\begin{bmatrix}(X'X)^{-1}X'Y+(X'X)^{-1}X'ZS_{22}^{-1}Z'X(X'X)^{-1}X'Y-(X'X)^{-1}X'ZS_{22}^{-1}Z'Y\\ -S_{22}^{-1}Z'X(X'X)^{-1}X'Y+S_{22}^{-1}Z'Y\end{bmatrix}$$

$$\Longrightarrow C^*\hat{\gamma} = C^*\begin{bmatrix}(X'X)^{-1}X'Y-(X'X)^{-1}X'ZS_{22}^{-1}S_{21}\\ S_{22}^{-1}S_{21}\end{bmatrix} \qquad (2.3.2)$$

onde $S_{21} = Z'[I-X(X'X)^{-1}X']Y$.

Então, para obter um estimador não viciado de $C\beta$, dadas as observações da variável auxiliar Z, basta tomar em (2.3.2), $C^* = [C\ 0]$. O estimador desejado será dado por:

$$C\hat{\hat{\beta}} = C(X'X)^{-1}X'Y-C(X'X)^{-1}X'ZS_{22}^{-1}S_{21} \qquad (2.3.3)$$

Comparando (2.2.4) com (2.3.3), observamos que $C\hat{\hat{\beta}}$ equivale ao estimador $C\hat{\beta}$ ajustado por covariância.

Segundo os resultados obtidos no apêndice A, sabemos que o estimador $C\hat{\hat{\beta}}$ corresponde ao estimador de máxima verossimilhança de $C\beta$ considerando o modelo condicional

$$MLM(Y|Z,[X\ Z]\begin{bmatrix}\beta\\ \eta\end{bmatrix},\Sigma_{11}^*).$$

Como a função densidade de probabilidade marginal de Z não depende de β, o estimador $C\hat{\hat{\beta}}$ também corresponde ao estimador de máxima verossimilhança de $C\beta$ considerando o

$$MLM([Y\ Z],X[\beta\ 0],\Sigma).$$

A matriz de variância-covariância do estimador $C\beta$ pode ser representada por:

$$V(C\hat{\hat{\beta}}) = \frac{(n-r)-1}{(n-r)-(p-q)-1}\,\Sigma_{11}^*\otimes[C(X'X)^{-1}C'] \qquad (2.3.4)$$

Comparando as expressões (2.3.1) e (2.3.4), e lembrando que

$\Sigma^*_{11} = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$, podemos concluir que o estimador $C\hat{\beta}$ é mais eficiente que $C\hat{\hat{\beta}}$ sempre que $\Sigma_{12} = 0$. Por outro lado, observamos que a eficiência de $C\hat{\hat{\beta}}$ aumenta conforme Σ_{12} "se afasta" de 0.

No apêndice A mostramos que um estimador não viciado de $C^*\gamma U^*$ considerando o $MLM(Y, D\gamma, \Sigma^*_{11})$ é dado por $C^*\hat{\gamma}U^* = C^*(D'D)^{-1}D'YU^*$. Então tomando $C^* = [C \;\; 0]$ e $U^* = U$, podemos escrever um estimador não viciado de $C\beta U$ ajustado por covariância, na forma:

$$C\hat{\hat{\beta}}U = C[(X'X)^{-1}X'Y - (X'X)^{-1}X'ZS_{22}^{-1}S_{21}]U \qquad (2.3.5)$$

Sua matriz de variância-covariância pode ser representada por:

$$V(C\hat{\hat{\beta}}U) = \frac{(n-r)-1}{(n-r)-(p-q)-1}[U'\Sigma^*_{11}U] \otimes [C(X'X)^{-1}C'] \qquad (2.3.6)$$

Um estimador não viciado de $T\eta$ pode ser obtido tomando-se $C^* = [0 \;\; T]$ em (2.3.2). Teremos então:

$$T\hat{\eta} = TS_{22}^{-1}S_{21} \qquad (2.3.7)$$

2.4 - A SOLUÇÃO PROPOSTA POR RAO

Inicialmente vamos demonstrar um lema (Khatri, 1966) cujo resultado será empregado na obtenção dos estimadores da matriz ξ e de combinações lineares da forma $C\xi U$ a partir de um conjunto de dados para os quais admitimos válidas as suposições do $MLMC(Y, X\xi G, \Sigma)$.

LEMA 2.4.1 (Khatri) - Sejam as matrizes G $(q \times p)$ e G_2 $(p \times p-q)$ de postos q e $p-q$ respectivamente, tais que $GG_2 = 0$. Então se S $(p \times p)$ for uma matriz simétrica e positiva definida, teremos:

$$G_2(G_2'SG_2)^{-1}G_2' = S^{-1} - S^{-1}G'(GS^{-1}G')^{-1}GS^{-1}$$

Como S é uma matriz simétrica e p.d. existe uma matriz A (p×p) não singular, tal que $S = AA'$. Analogamente podemos concluir que existem matrizes B_1 (q×q) e B_2 (p-q×p-q), não singulares, tais que: $(GS^{-1}G')^{-1} = B_1B_1'$ e $(G_2'SG_2)^{-1} = B_2B_2'$.

Consideremos a matriz $\Delta = [A^{-1}G'B_1 \quad A'G_2B_2]$. Como $S = AA'$, temos $S^{-1} = (AA')^{-1} = A'^{-1}A^{-1}$. Então:

$$\Delta'\Delta = \begin{bmatrix} B_1'GA'^{-1} \\ B_2'G_2'A \end{bmatrix} [A^{-1}G'B_1 \; A'G_2B_2] = \begin{bmatrix} B_1'GA'^{-1}A^{-1}G'B_1 & B_1'GA'^{-1}A'G_2B_2 \\ B_2'G_2'AA^{-1}G'B_1 & B'G_2'AA'G_2B_2 \end{bmatrix} =$$

$$= \begin{bmatrix} B_1'GS^{-1}G'B_1 & B_1'GG_2B_2 \\ B_2'G_2'G'B_1 & B_2'G_2'SG_2B_2 \end{bmatrix} = \begin{bmatrix} I_q & 0 \\ 0 & I_{p-q} \end{bmatrix} = I_p.$$

Então

$$\Delta\Delta' = I = [A^{-1}G'B_1 \; A'G_2B_2] \begin{bmatrix} B_1'GA'^{-1} \\ B_2'G_2'A \end{bmatrix} \Longrightarrow$$

$$\Longrightarrow I = A^{-1}G'B_1B_1'GA'^{-1} + A'G_2B_2B_2'G_2'A \Longrightarrow$$

$$\Longrightarrow I = A^{-1}G'(GS^{-1}G')^{-1}GA'^{-1} + A'G_2(G_2'SG_2)^{-1}G_2'A \Longrightarrow$$

$$\Longrightarrow A'^{-1}A^{-1} = A'^{-1}A^{-1}G'(GS^{-1}G')^{-1}GA'^{-1}A^{-1} + A'^{-1}A'G_2(G_2'SG_2)^{-1}G_2'AA^{-1} \Longrightarrow$$

$$\Longrightarrow S^{-1} - S^{-1}G'(GS^{-1}G')^{-1}GS^{-1} = G_2(G_2'SG_2)^{-1}G_2'$$

Consideremos agora o MLMC$(Y, X\xi G, \Sigma)$ e vamos obter estimadores de ξ, $C\xi U$ e Σ.

Sejam as matrizes G_1 (p×q) e G_2 (p×p-q) de postos q e (p-q) respectivamente, tais que: $GG_1 = I$ (q×q) e $GG_2 = 0$ (q×p-q) e consideremos a transformação:

$$Y[G_1 \; G_2] = [YG_1 \; YG_2] = [Y_1 \; Y_2] \qquad (2.4.1)$$

Vale a pena observar que tomando $G_1 = G'(GG')^{-1}$ e G_2 igual a

uma base da variedade linear gerada pelas colunas de

$$[I - G'(GG')^{-1}G]$$

as condições acima estarão satisfeitas. O problema da determinação de G_2 será discutido na seção 4.1.

A transformação (2.4.1) nos permite concluir que:

$$E(Y_1) = X\xi GG_1 = X\xi \text{ e } E(Y_2) = X\xi GG_2 = 0,$$

ou seja que Y_1 é formada por q vetores pertencentes ao espaço de estimação, e Y_2 por (p-q) vetores pertencentes ao espaço de erros.

O modelo pode ser definido em função de Y_1 e Y_2 através de:

$$[Y_1 \ Y_2] = X[\xi \ \Delta]+\varepsilon_1 \qquad (2.4.2)$$

Essa forma nos permitirá construir um teste de ajuste do modelo inicial, pois $\Delta = 0$ se o MLMC$(Y, X\xi G, \Sigma)$ for adequado. Nesse caso podemos considerar o

$$MLM([Y_1 \ Y_2], X[\xi \ 0], \Sigma_1),$$

onde

$$\Sigma_1 = \begin{bmatrix} G_1' \\ G_2' \end{bmatrix} \Sigma [G_1 \ G_2] = \begin{bmatrix} G_1'\Sigma G_1 & G_1'\Sigma G_2 \\ G_2'\Sigma G_1 & G_2'\Sigma G_2 \end{bmatrix}$$

A matriz Y_2 pode ser tomada como matriz de variáveis auxiliares e então podemos aplicar os resultados da seção 2.3 ao modelo condicional:

$$MLM(Y_1 | Y_2, [X \ Y_2] \begin{bmatrix} \xi \\ \eta \end{bmatrix}, \Sigma_2) \qquad (2.4.3)$$

onde

$$\eta = (G_2'\Sigma G_2)^{-1}(G_2'\Sigma G_1)$$

e

$$\Sigma_2 = G_1'\Sigma G_1 - G_1'\Sigma G_2(G_2'\Sigma G_2)^{-1}G_2'\Sigma G_1 = (G\Sigma^{-1}G')^{-1} \qquad (2.4.4)$$

O resultado (2.4.4) pode ser obtido através de uma aplicação do Lema de Khatri. Lembrando que $GG_1 = I$, $GG_2 = 0$ e Σ é uma matriz simétrica e positiva definida, temos:

$$G_2(G_2'\Sigma G_2)^{-1}G_2' = \Sigma^{-1} - \Sigma^{-1}G'(G\Sigma^{-1}G')^{-1}G\Sigma^{-1}.$$

Então:

$$\begin{aligned}\Sigma_2 &= G_1'\Sigma G_1 - G_1'\Sigma[\Sigma^{-1} - \Sigma^{-1}G'(G\Sigma^{-1}G')^{-1}G\Sigma^{-1}]\Sigma G_1 = \\ &= G_1'\Sigma G_1 - G_1'\Sigma\Sigma^{-1}\Sigma G_1 + G_1'\Sigma\Sigma^{-1}G'(G\Sigma^{-1}G')^{-1}G\Sigma^{-1}\Sigma G_1 = \\ &= G_1'G'(G\Sigma^{-1}G')^{-1}GG_1 = (G\Sigma^{-1}G')^{-1}\end{aligned}$$

Fazendo

$$D = [X \;\; Y_2] \quad e \quad \gamma = \begin{bmatrix}\xi \\ \eta\end{bmatrix}$$

podemos representar o modelo (2.4.3) através do $MLM(Y_1, D\gamma, \Sigma_2)$. Então, utilizando o resultado (2.3.3) concluímos que um estimador não viciado de $C\xi$ pode ser dado por:

$$C\hat{\hat{\xi}} = C[(X'X)^{-1}X'Y_1 - (X'X)^{-1}X'Y_2S_{22}^{-1}S_{21}]$$

Como

$$Y_1 = YG_1;\; Y_2 = YG_2$$

$$S_{22} = Y_2'[I - X(X'X)^{-1}X']Y_2 = G_2'Y'[I - X(X'X)^{-1}X']YG_2 = G_2'SG_2$$

$$S_{21} = G_2'SG_1$$

temos:

$$C\hat{\hat{\xi}} = C[(X'X)^{-1}X'YG_1 - (X'X)^{-1}X'YG_2(G_2'SG_2)^{-1}G_2'SG_1]$$

e aplicando o lema de Khatri, obtemos:

$$C\hat{\hat{\xi}} = C\{(X'X)^{-1}X'YG_1-(X'X)^{-1}X'Y[S^{-1}-S^{-1}G'(GS^{-1}G')^{-1}GS^{-1}]SG_1\} =$$

$$= C\{(X'X)^{-1}X'YG_1-(X'X)^{-1}X'Y[I-S^{-1}G'(GS^{-1}G')^{-1}G]G_1\} =$$

$$= C\{(X'X)^{-1}X'Y[I-I+S^{-1}G'(GS^{-1}G')^{-1}G]G_1\} \Longrightarrow$$

$$\Longrightarrow C\hat{\hat{\xi}} = C(X'X)^{-1}X'YS^{-1}G'(GS^{-1}G')^{-1} \qquad (2.4.5)$$

Utilizando (2.3.4) podemos representar a matriz de variância-covariância desse estimador por:

$$V(C\hat{\hat{\xi}}) = \frac{(n-r)-1}{(n-r)-(p-q)-1}(G\Sigma^{-1}G')^{-1}\otimes[C(X'X)^{-1}C'] \qquad (2.4.6)$$

Utilizando os resultados (2.3.5) e (2.3.6) podemos concluir que um estimador não viciado de $C\xi U$ é dado por:

$$C\hat{\hat{\xi}}U = C[(X'X)^{-1}X'YS^{-1}G'(GS^{-1}G')]U \qquad (2.4.7)$$

e sua matriz de variância-covariância pode ser representada por:

$$V(C\hat{\hat{\xi}}U) = \frac{(n-r)-1}{(n-r)-(p-q)-1}[U'(G\Sigma^{-1}G')^{-1}U]\otimes[C(X'X)^{-1}C'] \qquad (2.4.8)$$

Apesar de conseguirmos obter expressões para o valor esperado e para a matriz de variância-covariância de estimadores da forma $C\hat{\hat{\xi}}$, ainda não foi obtida uma expressão simples para a sua distribuição. Gleser e Olkin, (1966), expressaram a função densidade de probabilidade do estimador $\hat{\hat{\xi}}$ na forma de uma integral.

A distribuição condicional de $C\hat{\hat{\xi}}$ dadas as observações transformadas Y_2 é normal multivariada com vetor de valores esperados obtido a partir de $C\xi$ e a matriz de variância-covariância dada por: $(G\Sigma^{-1}G')^{-1}\otimes(C'RC)$, onde R $(r\times r)$ é o menor principal da matriz $(D'D)^{-1}$.

Os resultados do apêdice A nos permitem concluir que um estimador não viciado de Σ é dado por:

$$\hat{\Sigma} = \frac{1}{n-r} Y'[I-X(X'X)^{-1}X']Y = \frac{1}{n-r} S \qquad (2.4.9)$$

e que sua distribuição é proporcional a $W_p(n-r,\Sigma,\cdot)$ central.

Consideremos novamente o modelo condicional

$MLM(Y_1, D\gamma, \Sigma_2)$, e a matriz $S_1 = Y_1'[I-D(D'D)^{-1}D']Y_1$.

Do apêndice A podemos concluir que a distribuição condicional de S_1 dadas as observações transformadas Y_2 é

$$W_q[(n-r)-(p-q),\Sigma_2\cdot] \text{ central.}$$

Como essa distribuição só depende da matriz D através de seu posto (que é um parâmetro do modelo linear multivariado de crescimento original), a distribuição (incondicional) de S_1 também é $W_q[(n-r)-(p-q),\Sigma_2,\cdot]$ central.

Então $E(S_1) = [(n-r)-(p-q)]\Sigma_2$, e um estimador não viciado de Σ_2 é dado por:

$$\hat{\Sigma}_2 = \frac{1}{(n-r)-(p-q)} S_1$$

Podemos demonstrar, aplicando o lema de Khatri, que $S_1 = (G'S^{-1}G)^{-1}$; então o estimador $\hat{\Sigma}_2$ pode ser escrito na forma:

$$\hat{\Sigma}_2 = \frac{1}{(n-r)-(p-q)} (G'S^{-1}G)^{-1} \qquad (2.4.10)$$

Um estimador não viciado da matriz de variância-covariância do estimador $C\hat{\hat{\xi}}$ (2.4.6), é dado por:

$$\hat{V}(C\hat{\hat{\xi}}) = \frac{(n-r)-1}{(n-r)-(p-q)-1} \hat{\Sigma}_2 \otimes [C(X'X)^{-1}C'] \qquad (2.4.11)$$

Vale a pena observar que os estimadores obtidos são invariantes em relação à escolha das matrizes G_1 e G_2.

Rao (1965, 1966) faz alguns comentários sobre a solução proposta por Potthoff e Roy (1964) para o problema de estimação utilizando o MLMC$(Y, X\xi G, \Sigma)$.

Em primeiro lugar, Rao critica a arbitrariedade na escolha da matriz Δ, e considera que a sugestão de escolher Δ a partir de informação proveniente de outro conjunto de dados não é inteiramente satisfatória. O autor observa, também, que através dessa transformação, a matriz de observações Y $(n\times p)$ é reduzida para uma matriz Y^* $(n\times q)$ onde $q \leq p$, o que geralmente implica uma perda de informação (a não ser que a matriz de variância-covariância Σ seja conhecida e tomemos $\Delta = \Sigma$.)

Em seguida Rao mostra que, ao considerarmos a técnica de estimação com ajuste por covariância, poderíamos fazer a transformação $[Y_1 \;\; Y_2] = Y[G_1 \;\; G_2]$, tomando

$$G_1 = \Delta^{-1}G'(G\Delta^{-1}G')^{-1} \quad \text{e} \quad G_2 \text{ tal que } GG_2 = 0.$$

Concluímos então que a solução proposta por Potthoff e Roy simplesmente despreza a informação contida em Y_2, o que implica uma menor precisão (em relação ao estimador ajustado por covariância), a não ser que essa informação seja desprezível.

2.5 - A ESCOLHA DE VARIÁVEIS AUXILIARES

Na seção 2.4 consideramos o MLMC$(Y, X\xi G, \Sigma)$ e apresentamos um estimador de $C\xi$ obtido através do ajuste por covariância (em relação às variáveis transformadas $Y_2 = YG_2$) do estimador $C\hat{\xi} = C(X'X)^{-1}X'Y_1$ (onde $Y_1 = YG_1$). Foi observado também, que o estimador obtido é invariante em relação a escolha das matrizes arbitrárias G_1 e G_2 (satisfeitas as restrições apresentadas). Por outro lado, vimos na seção 2.3 que a técnica de ajuste por covariância pode não fornecer estimadores mais eficientes, e que a eficiência do estimador a-

justado é função da covariância entre as variáveis auxiliares (Y_2) e as variáveis principais (Y_1). Essa covariância é dada por $G_1'\Sigma G_2$, e portanto depende da escolha de G_1 e G_2.

Discutiremos a seguir o problema da escolha das variáveis auxiliares, apresentando alguns resultados obtidos por Rao, (1967). Consideraremos algumas estruturas especiais da matriz Σ (comuns aos modelos usuais de planejamento de experimentos), determinaremos as matrizes G_1 e G_2 e selecionaremos as variáveis auxiliares que devem ser introduzidas na análise.

i) Seja $\Sigma = G'\Gamma G + H'\theta H + I\sigma^2$, onde Γ $(q\times q)$ e θ $(q\times q)$ são matrizes de variância-covariância desconhecidas e H $(q\times p)$ é uma matriz conhecida tal que $\rho(H) = q$ e $GH' = 0$.

Consideremos $G_1 = G'(GG')^{-1}$ e G_2 como uma base da variedade linear gerada pelas colunas de $I - G'(GG')^{-1}G$. Então temos:

$$HG_1 = HG'(GG')^{-1} = 0$$

$$G_1'G_2 = 0$$

$$GG_2 = 0$$

Façamos a transformação: $[Y_1 \; Y_2] = Y[G_1 \; G_2]$. A covariância entre a i-ésima linha de Y_1 (y_i^1) e a i-ésima linha de Y_2 (y_i^2) será dada por:

$$\mathrm{Cov}(y_i^1, y_i^2) = G_1'\Sigma G_2 = G_1'[G'\Gamma G + H'\theta H + I\sigma^2]G_2 =$$

$$= G_1'G'\Gamma GG_2 + G_1'H'\theta HG_2 + G_1'G_2\sigma^2 = 0$$

Então Y_2 não fornece nenhuma informação sobre Y_1 e pode ser desprezada. Estimadores de $C\xi$ podem ser obtidos considerando o $\mathrm{MLM}(Y_1, X\xi, \Sigma_{11})$ e esta solução é equivalente à solução proposta por Potthoff e Roy, (1964) com $\Delta = I$.

Se $H = 0$ ou $\theta = 0$, as conclusões acima continuam vá-

lidas. Este modelo foi estudado por Rao, (1965), e se adapta ao caso onde a dispersão dos desvios das observações de cada unidade experimental em relação à curva do grupo ao qual pertence é dada por Γ e a dispersão em relação a sua própria curva é dada por $\sigma^2 I$.

ii) Seja $\Sigma = T'\Gamma T+\sigma^2 I$, onde Γ é uma matriz de variância covariância desconhecida e T é uma matriz arbitrária, conhecida, tal que $\rho(T) = t$.

Consideremos $G_1 = G'(GG')^{-1}$; G_2 $(p\times t)$ como uma base da variedade linear gerada pelas colunas de

$$[I-G'(GG')^{-1}G]T', \quad \text{e} \quad G_3 \ (p\times p-t-q)$$

tal que $GG_3 = 0$, $TG_3 = 0$ e $[G_1 \ G_2 \ G_3]$ seja não singular. Façamos a transformação $[Y_1 \ Y_2 \ Y_3] = [G_1 \ G_2 \ G_3]$. A covariância entre a i-ésima linha de Y_1 (y_i^1) e a i-ésima linha de Y_2 (y_i^2) é dada por:

$$\mathrm{Cov}(y_i^1, y_i^2) = G_1'\Sigma G_2 \neq 0$$

Analogamente podemos escrever:

$$\mathrm{Cov}(y_i^1, y_i^3) = G_1'\Sigma G_3 = G_1'[T'\Gamma T+\sigma^2 I]G_3 =$$

$$= G_1'T'\Gamma TG_3+\sigma^2 G_1'G_3 = 0$$

$$\mathrm{Cov}(y_i^2, y_i^3) = G_2'\Sigma G_3 = G_2'[T'\Gamma T+\sigma^2 I]G_3 =$$

$$= G_2'T'\Gamma TG_3+\sigma^2 G_2'G_3 = 0$$

Então Y_3 não fornece nenhuma informação sobre Y_1 ou Y_2, e pode ser desprezada. A matriz Y_2, no entanto, contém informação sobre Y_1 e deve ser considerada como variável auxiliar. Estimadores de $C\xi$ podem ser obtidos considerando o modelo condicional

$$\mathrm{MLM}(Y_1 \mid Y_2, [X \ Y_2]\begin{bmatrix}\xi\\ \eta\end{bmatrix} \ \Sigma_2),$$

descrito anteriormente.

Devemos observar que Y_2 e Y_3 dividem o espaço de erros em duas partes: uma contendo informação sobre Y_1 e outra sem nenhuma informação sobre essa variável.

Outras suposições sobre a estrutura de Σ podem ser consideradas, obtendo-se diferentes escolhas para as variáveis auxiliares. Essas considerações, porém, não apresentam muita utilidade em problemas práticos, porque em geral não temos informação sobre Σ. Esse problema será analisado no seu aspecto prático, na seção 4.1.

- □ -

CAPÍTULO 3

O PROBLEMA DE TESTES DE HIPÓTESES EM MODELOS LINEARES MULTIVARIADOS DE CRESCIMENTO

O objetivo deste capítulo é obter testes de significância para hipóteses estatísticas envolvendo elementos do MLMC$(Y, X\xi G, \Sigma)$, a partir de um conjunto de dados para o qual são admitidas válidas as suposições desse modelo.

Na seção 3.1 apresentamos um teste de ajuste do modelo de crescimento adotado.

Na seção 3.2 apresentamos testes de hipóteses sobre combinações lineares dos parâmetros.

O princípio utilizado nessas duas seções para obtenção dos testes é o de redução das hipóteses originais à hipóteses equivalentes envolvendo um modelo linear multivariado, obtido através da transformação $[Y_1 \ Y_2] = Y[G_1 \ G_2]$, e a aplicação dos resultados do apêndice B.

Finalmente na seção 3.3 fazemos considerações sobre a determinação de intervalos de confiança envolvendo os parâmetros do modelo.

3.1 - UM TESTE DE AJUSTE DO MODELO

Consideremos o MLMC$(Y, X\xi G, \Sigma)$. Fazendo a transforma

ção $[Y_1 \; Y_2] = Y[G_1 \; G_2]$ apresentada na seção 2.4, podemos escrever o modelo na forma:

$$[Y_1 \; Y_2] = X[\xi \; \Delta] + \varepsilon_1,$$

onde $\Delta = 0$ se o modelo de crescimento $Y = X\xi G + \varepsilon$ for adequado. Obter um teste de significância para o ajuste do modelo corresponde então, a testar a hipótese:

$$H_A: E(Y_2) = X\xi G G_2 = 0 \qquad (3.1.1)$$

Considerando o $MLM(Y, X\xi^*, \Sigma)$ onde $\xi^* = \xi G$, a hipótese (3.1.1) é equivalente à:

$$H_A: C\xi^* U = 0 \text{ onde } C = I \text{ e } U = G_2 \qquad (3.1.2)$$

As matrizes de somas de quadrados e produtos cruzados devidas ao erro (S_A) e ao desvio em relação à hipótese nula (H_A), podem ser obtidas através dos resultados (B.3.1) e (B.3.2), e são dadas por:

$$S_A = G_2' Y' [I - X(X'X)^{-1} X'] Y G_2$$

$$H_A = [(X'X)^{-1} X' Y G_2]' [(X'X)^{-1}]^{-1} [(X'X)^{-1} X' Y G_2] = G_2' Y' X (X'X)^{-1} X' Y G_2$$

Além disso, podemos concluir que as respectivas distribuições são independentes e são dadas por:

$$S_A \sim W_{p-q}[n-r, G_2' \Sigma G_2, \cdot] \text{ central}$$

$$H_A \sim W_{p-q}[r, G_2' \Sigma G_2, G_2' \xi^{*\prime} X' X \xi^* G_2] \text{ não central.}$$

Quando a hipótese H_A: $\xi^* G_2 = 0$ é verdadeira, H_A tem uma distribuição central.

Todos os critérios de teste apresentados no apêndice B dependem das raízes da equação $|H_A - \lambda S_A| = 0$, e qualquer um deles pode ser utilizado para testar a hipótese (3.1.2).

3.2 - TESTES DE HIPÓTESES LINEARES SOBRE OS PARÂMETROS DO MODELO LINEAR MULTIVARIADO DE CRESCIMENTO

Admitindo que o MLMC$(Y, X\xi G, \Sigma)$ é adequado ao fenômeno em estudo, podemos testar hipóteses lineares da forma:

$$H_\xi: \quad C\xi U = \Gamma \tag{3.2.1}$$

Considerando o modelo condicional

$$MLM(Y_1 | Y_2, [X \ Y_2]\begin{bmatrix}\xi \\ \eta\end{bmatrix}, \Sigma_2)$$

que pode ser escrito na forma MLM$(Y_1, D\gamma, \Sigma_2)$, a hipótese (3.2.1) é equivalente à:

$$H_\xi: \ C^*\gamma U = [C \ 0]\begin{bmatrix}\xi \\ \eta\end{bmatrix} U = C\xi U = \Gamma \tag{3.2.2}$$

onde $C^* = [C \ 0]$.

Neste caso, as matrizes de somas de quadrados e produtos cruzados devidos ao erro (S_ξ) e ao desvio em relação à hipótese nula (H_ξ), obtidas através de (B.3.1) e (B.3.2), são dadas por:

$$S_\xi = U'G_1'Y'[I-D(D'D)^{-1}D']YG_1U = U'S_1U = U'(G'S^{-1}G)^{-1}U$$

$$H_\xi = [C\hat{\hat{\xi}}U-\Gamma]'\{[C \ 0](D'D)^{-1}\begin{bmatrix}C' \\ 0'\end{bmatrix}\}^{-1}[C\hat{\hat{\xi}}U-\Gamma] =$$

$$=[C(X'X)^{-1}X'YS^{-1}G'(GS^{-1}G')^{-1}U-\Gamma]'[CRC']^{-1}[C(X'X)^{-1}X'YS^{-1}G'(GS^{-1}G')^{-1}U-\Gamma]$$

onde R $(r \times r)$ é o menor principal de $(D'D)^{-1}$.

Além disso podemos concluir que as respectivas distribuições são independentes e são dadas por:

$$S_\xi \sim W_u[(n-r)-(p-q), U'\Sigma_2U, \cdot] \text{ central}$$

$$H_\xi | Y_2 \sim W_u[c; U'\Sigma_2 U; (C\xi U-\Gamma)'(CRC')^{-1}(C\xi U-\Gamma)] \text{ não central.}$$

Quando a hipótese H_ξ: $C\xi U = \Gamma$ é verdadeira, a distribuição condicional de H_ξ dadas as observações Y_2 é central e não depende da matriz D. Portanto coincide com a distribuição (incondicional) de H_ξ.

Os critérios de teste apresentados no apêndice B dependem das raízes da equação $|H_\xi - \lambda S_\xi| = 0$, e qualquer um deles pode ser usado para testar a hipótese (3.2.2).

Se alguns vetores pertencentes ao espaço de erros não forem considerados como variáveis auxiliares, devemos substituir, na expressão que indica o número de graus de liberdade da distribuição de S_ξ, o parâmetro (p-q) pelo número de vetores realmente utilizados.

A adoção do modelo condicional

$$MLM(Y_1 | Y_2, [X \ Y_2]\begin{bmatrix}\xi \\ \eta\end{bmatrix}, \Sigma_2)$$

também nos permite obter testes de hipóteses sobre os parâmetros η. Por exemplo a hipótese:

$$H_\eta: \ C\eta U = \Gamma \tag{3.2.3}$$

pode ser testada através da hipótese equivalente:

$$H_\eta: \ C^*\gamma U = [0 \ C]\begin{bmatrix}\xi \\ \eta\end{bmatrix} U = C\eta U = \Gamma \tag{3.2.4}$$

onde $C^* = [0 \ C]$.

Analogamente ao caso de testes sobre os parâmetros ξ, podemos obter as matrizes S_η e H_η

$$S_\eta = U'G_1'Y'[I-D(D'D)^{-1}D']YG_1U = U'(G'S^{-1}G)^{-1}U$$

$$H_\eta = [C\hat{\hat{\eta}}U-\Gamma]'\{[0\ C](D'D)^{-1}\begin{bmatrix}0'\\C'\end{bmatrix}\}^{-1}[C\hat{\hat{\eta}}U-\Gamma] =$$

$$= [C\hat{\hat{\eta}}U-\Gamma]'[CR^*C']^{-1}[C\hat{\hat{\eta}}U-\Gamma]$$

onde R^* $(p-q\times p-q)$ é o menor inferior direito de $(D'D)^{-1}$. Podemos também concluir que as distribuições das matrizes S_η e H_η são independentes e são dadas por:

$$S_\eta \sim W_u[(n-r)-(p-q), U'\Sigma_2 U] \text{ central}$$

$$H_\eta | Y_2 \sim W_u[c, U'\Sigma_2 U, (C\eta U-\Gamma)'(CR^*C')^{-1}(C\eta U-\Gamma)] \text{ não central.}$$

Sob a hipótese nula, a distribuição condicional de H_η dadas as observações Y_2 é central e não depende da matriz D, coincidindo com a distribuição (incondicional) de H_η. Qualquer um dos critérios apresentados no apêndice B pode ser usado para testar a hipótese (3.2.3).

Devemos observar que Y_2 só deve ser incluída no modelo como matriz de variáveis auxiliares se $\eta \neq 0$. Portanto podemos verificar a conveniência da utilização de Y_2 como variável auxiliar através de um teste da hipótese (3.2.3) tomando $C = I$, $U = I$ e $\Gamma = 0$.

3.3 - INTERVALOS DE CONFIANÇA

Como no caso univariado, a rejeição de uma hipótese de igualdade de efeitos médios dos diversos tratamentos (grupos) de um problema multivariado não indica quais os tratamentos responsáveis pela rejeição da hipótese ou quais aqueles que podem ser considerados como provenientes de uma mesma população. Nos problemas multivariados, podemos ainda considerar hipóteses de igualdade dos efeitos médios das diversas respostas (tempos, por exemplo), e da mesma forma, não podemos inferir quais as respostas cujos efeitos médios levaram à rejeição da hipóte-

se.

Por outro lado, geralmente a simples rejeição de uma hipótese estatística não é uma conclusão satisfatória para o pesquisador, pois ele estaria interessado em conhecer a magnitude dos efeitos significativos e em determinar intervalos de confiança para as estimativas obtidas.

Esses problemas podem ser resolvidos através de comparações múltiplas. Como vimos no primeiro parágrafo, o problema no caso multivariado deve ser generalizado no sentido de considerar comparações tanto entre linhas (tratamentos, por exemplo) como entre as colunas (tempos, por exemplo) da matriz de parâmetros.

Nesse sentido o critério da União-Intersecção de Roy apresenta vantagens sobre os demais, pois sua própria construção leva diretamente à obtenção de intervalos de confiança para todas as formas bilineares envolvendo a matriz $C\xi U$. Roy, (1957) e Khatri, (1966) mostraram que intervalos de confiança com coeficiente de confiança $(1-\alpha)\times 100\%$ para *todas* as formas bilineares $a'C\xi Ub$, onde a e b são vetores $(c\times 1)$ e $(u\times 1)$ respectivamente, são dados por:

$$a'C\hat{\xi}Ub - \sqrt{\lambda_\alpha (b'S_\xi b)(a'CRC'a)} \leq a'C\xi Ub \leq$$

$$\leq a'C\hat{\xi}Ub + \sqrt{\lambda_\alpha (b'S_\xi b)(a'CRC'a)} \qquad (3.3.1)$$

onde:

S_ξ é a matriz de somas de quadrados e produtos cruzados devida ao erro, associada ao teste da hipótese $H: C\xi U = 0$ considerando o $MLM(Y_1, D\gamma, \Sigma_2)$.

R $(r\times r)$ é o menor principal de $(D'D)^{-1}$, onde D é a matriz de planejamento do $MLM(Y_1, D\gamma, \Sigma_2)$.

$\lambda_\alpha = \frac{\theta_\alpha}{1-\theta_\alpha}$, onde θ_α é o α-ésimo percentil da distribuição

da maior raiz da equação $|H_\xi - \theta(S_\xi + H_\xi)| = 0$.

Quando $s = \min(c,u) = 1$, o valor λ_α pode ser obtido diretamente através da distribuição F. Nesse caso, se:

$$c = 1, \text{ então } \lambda_\alpha = \frac{u}{(n-r)-(p-q)-u+1} F_{[\alpha,u,(n-r)-(p-q)-c+1]}$$

$$u = 1, \text{ então } \lambda_\alpha = \frac{c}{(n-r)-(p-q)} F_{[\alpha,c,(n-r)-(p-q)]}$$

Essa técnica de comparações múltiplas é uma extensão da técnica desenvolvida por Scheffé.

- □ -

CAPÍTULO 4

APLICAÇÕES

A finalidade básica deste capítulo é ilustrar a aplicação da técnica de Análise de Curvas de Crescimento através da discussão de exemplos práticos.

Na seção 4.1 mostramos como a análise de um conjunto de dados, ao qual queremos ajustar o MLMC$(Y, X\xi G, \Sigma)$, pode ser realizada através da utilização de um programa convencional de Análise de Variância Multivariada capaz de testar hipóteses da forma H: $C\beta U = 0$ considerando o MLM$(Y, X\beta, \Sigma)$.

Na seção 4.2 discutimos a aplicação da técnica a dois conjuntos de dados encontrados na bibliografia, Potthoff e Roy, (1964) e Grizzle & Allen, (1969) e a um conjunto de dados submetidos ao SEA (Setor de Estatística Aplicada do Departamento de Estatística do IME-USP) para análise.

4.1 - SUGESTÃO PARA ANÁLISE DE UM CONJUNTO DE DADOS UTILIZANDO UM PROGRAMA CONVENCIONAL DE ANÁLISE DE VARIÂNCIA MULTIVARIADA

A análise estatística de um conjunto de dados através de um modelo linear multivariado de crescimento tem como ponto de partida uma matriz de observações Y $(n \times p)$, onde cada uma das n linhas representa o conjunto p medidas de uma variável, realizadas em uma determinada unidade experi-

mental (por exemplo, nos instantes $t_1, t_2, \ldots, t_p$, ou após a aplicação das doses $d_1, d_2, \ldots, d_p$ de uma certa droga, etc...). Para estudar o comportamento dessa variável em função do tempo, ou das doses da droga aplicada, através do MLMC$(Y, X\xi G, \Sigma)$, devemos admitir que as n linhas da matriz Y sejam não correlacionadas entre si, e que cada uma obedeça a uma distribuição normal multivariada com vetor de valores esperados 0 e matriz de variância-covariância, não singular, Σ (desconhecida).

O procedimento de análise consiste basicamente de quatro etapas:

i) Definição da matriz de planejamento X.

ii) Definição do grau do polinômio que queremos ajustar às observações, e consequentemente da matriz G e da forma da matriz de parâmetros desconhecidos, ξ.

iii) Escolha de possíveis variáveis auxiliares para obtenção de estimadores de ξ ajustados por covariância.

iv) Obtenção de estimadores e testes de hipóteses sobre os parâmetros envolvidos.

A matriz X é função do esquema de planejamento utilizado para o experimento, e sugestões para sua construção estão apresentadas no capítulo 1.

A determinação da matriz G e da forma da matriz ξ podem ser obtidas a partir de testes de ajuste de modelos de crescimento (linear, quadrático, etc...) (seção 3.1).

A escolha de possíveis variáveis auxiliares depende da matriz de correlação entre as variáveis do espaço de estimação e aquelas do espaço de erros. Esse aspecto foi discutido teoricamente no capítulo 2 e será considerado praticamente nesta seção.

O problema de estimação e obtenção de testes de hipóteses sobre os parâmetros do modelo está discutidos nas se

ção 2.4 e 3.2.

Nesta seção mostraremos como as quatro etapas descritas acima podem ser realizadas através de duas análises utilizando um programa convencional de Análise de Variância Multivariada capaz de testar hipóteses da forma $H: C\beta U = 0$, considerando o $MLM(Y, X\beta, \Sigma)$.

Na primeira análise supomos, por exemplo, que um polinômio de grau (q-1) se adapta às observações e verificamos a validade dessa suposição através do teste do ajuste do $MLMC(Y, X\xi G, \Sigma)$ sugerido na seção 3.1. Para isso consideramos o $MLM(Y, X\xi^*, \Sigma)$, onde $\xi^* = \xi G$, e testamos a hipótese $H: I\xi^* G_2 = 0$, onde G_2 $(p \times p-q)$ é uma matriz tal que $GG_2 = 0$. Quando as p observações de cada unidade experimental são realizadas em instantes igualmente espaçados as linhas da matriz G podem ser formadas pelos coeficientes dos polinômios ortogonais de graus $0, 1, \ldots, q-1$, pois nesse caso G_2 pode ser construída de tal forma que suas colunas sejam formadas pelos coeficientes dos polinômios ortogonais de graus $q, q+1, \ldots, p-1$. Quando isto não acontece, temos duas alternativas: obter os coeficientes de polinômios ortogonais para observações não igualmente espaçadas, ou tomar G_2 como uma base da variedade linear gerada pelas colunas de $[I-G'(GG')^{-1}G]$. A utilização de coeficientes de polinômios ortogonais é recomendada sempre que possível, pelas facilidades de cálculo que proporciona. Variando q de (p-1) até 2 e testando as hipóteses de ajuste do modelo correspondente, podemos determinar o grau de polinômio mais conveniente para explicar o comportamento da variável.

Se o programa fornecer, para cada escolha da matrizes C e U, a matriz de correlação correspondente, podemos ainda na primeira análise testar a hipótese $H: I\xi^*[G_1 \; G_2] = 0$, onde as colunas da matriz $[G_1 \; G_2]$ são formadas pelos coeficientes dos polinômios ortogonais de graus $0, 1, \ldots, p-1$. Essa hipótese não apresenta interesse, porém a matriz de cor-

relação obtida pode ser particionada de forma que possamos examinar as estimativas das correlações entre as variáveis pertencentes ao espaço de erros ($Y_2 = YG_2$) e as variáveis pertencentes ao espaço de estimação ($Y_1 = YG_1$). As variáveis do espaço de erros cujas correlações com variáveis do espaço de estimação sejam altas, podem ser consideradas como variáveis auxiliares na segunda análise. Um teste para a hipótese múltipla de que m coeficientes de correlação são significativamente diferentes de zero, pode ser obtido através da transformação de Fisher e da desigualdade de Bonferroni (Morrison, 1976, por exemplo).

Como resultado da primeira análise, obtemos então o grau do polinômio que desejamos ajustar às observações (e consequentemente as matrizes G, G_1 e G_2), além de uma indicação das variáveis do espaço de erros, que podem ser consideradas como variáveis auxiliares.

Se todas as variáveis pertencentes ao espaço de erros forem selecionadas como variáveis auxiliares, fazemos a transformação $Y_1 = YG_1$ e $Y_2 = YG_2$ (através de uma sub-rotina) e consideramos para a segunda análise, o modelo condicional:

$$MLM(Y_1 | Y_2, [X \;\; Y_2]\begin{bmatrix}\xi \\ \eta\end{bmatrix}, \Sigma_2)$$

definido na seção 2.4. Nessa análise obtemos as estimativas e testes de hipóteses desejados. Um estimador não viciado de $C\xi$, por exemplo, é dado por (2.4.5).

Quando nem todos os vetores de Y_2 forem selecionados como variáveis auxiliares, basta eliminar da matriz G_2, as colunas correspondentes àquelas variáveis cujas correlações com as variáveis do espaço de estimação não sejam significativamente diferentes de zero. Isto corresponde a fazer $G_2 = [G_2^* \;\; G_3]$ e considerar as transformações

$Y_1 = YG_1$, $Y_2 = YG_2^*$ e $Y_3 = YG_3$; Y_1 corresponde às variáveis dependentes, Y_2 às variáveis auxiliares e Y_3 pode ser ignorada. A análise prossegue como no caso anterior. Nesse caso, o estimador $C\xi$ obtido, corresponde ao estimador dado por (2.4.5) com S^{-1} substituída por uma outra matriz S^{*-1}, que leva em conta o fato de desprezarmos Y_3.

Se nenhum dos vetores pertencentes ao espaço de erros for selecionado como variável auxiliar, a análise deve prosseguir sem a utilização da matriz Y_2, ou seja considerando o $MLM(Y_1, X\xi, \Sigma_2)$. Isto equivale a adotar a solução proposta por Potthoff e Roy, (1964) com $\Delta = I$. Por exemplo, um estimador não viciado de $C\xi$, neste caso, é dado por:

$$C\hat{\xi} = C(X'X)^{-1}X'YG'(GG')^{-1}.$$

4.2 - EXEMPLOS

Nesta seção apresentamos situações reais para as quais o modelo linear multivariado pode ser adotado. Em cada caso discutimos aspectos práticos relacionados com a aplicação da técnica de análise desenvolvida neste trabalho.

Os cálculos foram realizados com a utilização do programa MGLM (Multivariate General Linear Models) desenvolvido por Starmer e Grizzle em 1968 na Universidade da Carolina do Norte (E.U.A.). Uma das sub-rotinas (READIN) foi modificada para tornar possível a transformação de variáveis exigida na análise de modelos lineares de crescimento.

Convém lembrar que, mesmo nos casos mais simples, a utilização de um programa de computador adequado é necessária, em virtude da grande quantidade de produtos e inversões de matrizes envolvidos. Para atender possíveis usuários da técnica, o programa encontra-se implantado no Centro de Computação Eletrônica da Universidade de São Paulo. Informações

sobre sua utilização podem ser obtidas junto ao Setor de Estatística Aplicada do Departamento de Estatística do IME-USP.

EXEMPLO 4.2.1 - Voltamos a discutir o problema apresentado por Potthoff e Roy, (1964) citado na seção 1.1 (Exemplo 1.1.1).

Examinando o gráfico das médias observadas (Figura 1.1.1) podemos concluir que aparentemente o comportamento da variável estudada (distância entre a pituitária e a fissura plerio-maxilar) no tempo poderia ser explicado por uma função linear para os dois grupos. Vamos então, ajustar o seguinte modelo (já descrito na seção 1.1) às observações:

$$\beta_{it} = \xi_{i0} + \xi_{i1}(t-\bar{t}) \qquad (i=1,2) \qquad (4.2.1)$$

A matriz abaixo (onde as estimativas das variâncias e covariâncias estimadas entre as medidas nos instantes t_1, t_2, t_3 e t_4 estão representadas acima da diagonal, e as estimativas das correlações, abaixo) pode nos fornecer uma indicação da conveniência de utilizarmos uma técnica de análise multivariada.

	(t_1)	(t_2)	(t_3)	(t_4)
(t_1)	5,42	2,72	3,91	2,71
(t_2)	0,57	4,18	2,93	3,32
(t_3)	0,66	0,56	6,56	4,13
(t_4)	0,52	0,73	0,73	4,99

Para aplicação da Técnica de Análise de Curvas de Crescimento, devemos escrever o modelo (4.2.1) na forma do MLMC$(Y, X\xi G, \Sigma)$. Então:

$$X = \begin{bmatrix} 1 & 0 \\ \vdots & \\ 1 & 0 \\ 0 & 1 \\ \vdots & \\ 0 & 1 \end{bmatrix} \begin{matrix} \left.\vphantom{\begin{matrix}1\\ \vdots\\ 1\end{matrix}}\right\} 11 \\ \left.\vphantom{\begin{matrix}1\\ \vdots\\ 1\end{matrix}}\right\} 16 \end{matrix} \; ; \; \xi = \begin{bmatrix} \xi_{10} & \xi_{11} \\ \xi_{20} & \xi_{21} \end{bmatrix} ; \; G = \begin{bmatrix} 1 & 1 & 1 & 1 \\ -3 & -1 & 1 & 3 \end{bmatrix}$$

Na primeira análise, adotamos o $MLM(Y, X\xi^*, \Sigma)$ onde $\xi^* = \xi G$ e testamos duas hipóteses:

i)

$$H_1: C\xi U = 0 \text{ onde } C = I_2 \text{ e } U = G_2 = \begin{bmatrix} 1 & -1 \\ -1 & 3 \\ -1 & -3 \\ 1 & 1 \end{bmatrix}$$

Se o modelo se ajusta às observações, devemos ter $\xi^* G_2 = 0$ e a não rejeição da hipótese H_1 implica a aceitação do modelo.

Utilizamos o critério da Razão de Verossimilhança Generalizada de Wilks, cuja estatística de teste (no caso) pode ser transformada numa estatística com distribuição F exata com (4,48) graus de liberdade sob a hipótese nula;o valor da estatística para os dados apresentados é $F_{obs} = 0,6780$ (nível descritivo P=0,6134), de onde concluímos que não devemos rejeitar a hipótese nula, ou seja, uma função linear poderia explicar o comportamento da variável no tempo.

ii)

$$H_2: C\xi^* U = 0 \text{ onde } C = I_2 \text{ e } U = [G_1 \; G_2] = \begin{bmatrix} 1 & -3 & 1 & -1 \\ 1 & -1 & -1 & 3 \\ 1 & 1 & -1 & -3 \\ 1 & 3 & 1 & 1 \end{bmatrix}$$

Esta hipótese em si não tem interesse, porém obtemos a estimativa da matriz de correlação:

		Y_1		Y_2	
		(0)	(1)	(2)	(3)
Y_1	(0)	1,00			
	(1)	0,10	1,00		
Y_2	(2)	-0,04	-0,29	1,00	
	(3)	-0,23	0,07	-0,03	1,00

As correlações entre colunas de Y_1 e Y_2 significativamente diferentes de zero e nos indicarão quais as variáveis que podem ser consideradas como variáveis auxiliares na segunda análise.

O teste sugerido em Morrison, (1976) com um nível de significância $\alpha = 0,05$ para os $m = 4$ coeficientes de correlação nos quais estamos interessados apresenta um valor crítico igual a 0,47. Como todos os 4 coeficientes de correlação estimados têm valor absoluto menor que 0,47 concluímos que os coeficientes de correlação correspondentes não são significativamente diferentes de zero, e que nenhuma variável do espaço de erros deve ser incluída na segunda análise.

Para a segunda análise, fazemos a transformação $Y_1 = YG_1$, e consideramos o $MLM(Y_1, X\xi, \Sigma_1)$ desprezando a matriz G_2. Um estimador pontual de ξ, por exemplo, é dado por (2.1.4) com $\Delta = I$. A estimativa de ξ obtida neste caso é dada por:

$$\hat{\xi} = \begin{array}{c} \begin{array}{cc} \text{coef.lin.} & \text{coef.ang.} \end{array} \\ \begin{bmatrix} 22,6477 & 0,4795 \\ 24,9688 & 0,7848 \end{bmatrix} \end{array} \begin{array}{l} \\ \text{meninas} \\ \text{meninos} \end{array}$$

Os desvios-padrão estimados dos elementos do estimador $\hat{\xi}$ são são obtidos a partir de:

$$\hat{V}(\hat{\xi}) = \frac{1}{n-r} [(GG')^{-1} G'SG(GG')^{-1}] \otimes (X'X)^{-1}$$

e são apresentados na tabela 4.2.1.

Para efeito de comparação, obtivemos também uma estimativa de ξ considerando as duas colunas de $Y_2 = YG_2$ como variáveis auxiliares (2.4.5, com $C = I$), bem como os desvios-padrão estimados dos elementos do estimador considerado (obtidos de (2.4.6), com $C = I$). Esses dados são apresentados na tabela 4.2.1

TABELA 4.2.1

	VARIÁVEIS AUXILIARES			
	NENHUMA		1ª E 2ª COLUNAS DE Y_2	
ESTIMATIVA DE ξ	22,6477	0,4795	22,6654	0,4764
	24,9688	0,7848	24,9371	0,8268
DESVIOS-PADRÃO ESTIMADOS	0,5861	0,1037	0,5949	0,1034
	0,4860	0,0860	0,5206	0,0905

Examinando os desvios-padrão estimados, podemos observar que neste caso, as estimativas obtidas *sem ajuste por covariância* são ligeiramente mais precisas do que aquelas obtidas utilizando a técnica de ajuste por covariância.

Também na segunda análise podemos testar várias hipóteses de interesse sobre os elementos da matriz ξ, algumas das quais apresentamos a seguir. Utilizamos como critério o critério da Razão de Verossimilhança Generalizada de Wilks.

Por exemplo, podemos testar a hipótese de que os coeficientes lineares das duas retas são iguais, ou seja, que $\xi_{10} = \xi_{20}$. Isto pode ser feito através da hipótese

$$H_3: C\xi U = 0 \text{ onde } C = [1 \quad -1] \text{ e } U = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

A estatística de teste neste caso, tem distribuição F exata com (1,25) graus de liberdade sob a hipótese nula; o valor observado da estatística é $F_{obs} = 9,2921$ (nível descritivo $P = 0,0055$), o que indica que devemos rejeitar a hipótese, ou seja, podemos concluir que as retas relativas aos dois grupos apresentam coeficientes lineares diferentes.

Uma outra hipótese de interesse poderia ser a hipótese de que os coeficientes angulares das duas retas são iguais ($\xi_{11} = \xi_{21}$). A hipótese equivalente em termos de contrastes da forma $C\xi U$ pode ser escrita como:

$$H_4: \ C\xi U = 0 \text{ onde } C = [1 \quad -1] \text{ e } U = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

A estatística de teste tem distribuição F exata com (1,25) graus de liberdade sobre a hipótese nula, e o valor obtido a partir dos dados é $F_{obs} = 5,1186$ (nível descritivo $P=0,0308$), indicando que a hipótese deve também ser rejeitada. Concluímos então que as taxas de crescimento da variável devem ser consideradas diferentes para crianças de sexos diferentes (ver Figura 1.1.1).

EXEMPLO 4.2.2 - Procuramos através deste exemplo ilustrar a aplicação da técnica de Análise de Curvas de Crescimento aos dados da tabela 4.2.2, apresentados por Grizzle e Allen, (1969).

Os dados foram obtidos de um experimento onde foi medida a concentração de potássio na coronária aos 1, 3, 5, 7, 9, 11 e 13 minutos após a oclusão da coronária em cachorros pertencentes a quatro grupos (cada grupo representando um tratamento).

Na primeira análise testamos o ajuste de curvas polinomiais de graus 5, 4, 3, 2 e 1 aos dados, considerando hipóteses da forma $H. \ I\xi^* G_2 = 0$ onde G_2 é formada por colunas

TABELA 4.2.2

Extraída de Grizzle e Allen, (1969)

CORONARY SINUS POTASSIUM—MIL EQUIVALENTS PER LITER

Group 1
Control dogs during the initial 13 minutes after coronary occlusion

Dog	Time-*minutes* after occlusion 1	3	5	7	9	11	13
1	4.0	4.0	4.1	3.6	3.6	3.8	3.1
2	4.2	4.3	3.7	3.7	4.8	5.0	5.2*
3	4.3	4.2	4.3	4.3	4.5	5.8	5.4*
4	4.2	4.4	4.6	4.9	5.3	5.6	4.9
5	4.6	4.4	5.3	5.6	5.9	5.9	5.3
6	3.1	3.6	4.9	5.2	5.3	4.2	4.1
7	3.7	3.9	3.9	4.8	5.2	5.4	4.2
8	4.3	4.2	4.4	5.2	5.6	5.4	4.7
9	4.6	4.6	4.4	4.6	5.4	5.9	5.6

* Denotes animals developing ventricular fibrillation

Group 2
Extrinsic cardiac denervation three weeks prior to coronary occlusion

Dog	Time-*minutes* after occlusion 1	3	5	7	9	11	13
10	3.1	3.4	3.5	3.1	3.1	3.7	3.3
11	3.0	3.2	3.0	3.0	3.1	3.2	3.1
12	3.0	3.1	3.2	3.0	3.3	3.0	3.0
13	3.1	3.2	3.2	3.2	3.3	3.1	3.1
14	3.8	3.9	4.0	2.9	3.5	3.5	3.4
15	3.0	3.6	3.2	3.1	3.0	3.0	3.0
16	3.3	3.3*	3.3	3.4	3.6	3.1	3.1
17	4.2	4.0	4.2	4.1	4.2	4.0	4.0
18	4.1	4.2	4.3	4.3	4.2	4.0	4.2
19	4.5	4.4	4.3	4.5	5.3	4.4	4.4

* Denotes brief period of ectopic ventricular activity

Group 3
Extrinsic cardiac denervation immediately prior to coronary occlusion

Dog	Time-*minutes* after occlusion 1	3	5	7	9	11	13
20	3.2	3.3	3.8	3.8	4.4	4.2	3.7
21	3.3	3.4	3.4	3.7	3.7	3.6	3.7
22	3.1	3.3	3.2	3.1	3.2	3.1	3.1
23	3.6	3.4	3.5	4.6	4.9	5.2	4.4
24	4.5	4.5	5.4	5.7	4.9	4.0	4.0
25	3.7	4.0	4.4	4.2	4.6	4.8	5.4*
26	3.5	3.9	5.8	5.4	4.9	5.3	5.6
27	3.9	4.0	4.1	5.0	5.4	4.4	3.9

* Denotes animals developing ventricular fibrillation

Group 4
Bilateral thoracic sympathectomy and stellectomy three weeks prior to coronary occlusion

Dog	Time-*minutes* after occlusion 1	3	5	7	9	11	13
28	3.1	3.5	3.5	3.2	3.0	3.0	3.2
29	3.3	3.2	3.6	3.7	3.7	4.2	4.4*
30	3.5	3.9	4.7	4.3	3.9	3.4	3.5
31	3.4	3.4	3.5	3.3	3.4	3.2	3.4
32	3.7	3.8	4.2	4.3	3.6	3.8	3.7
33	4.0	4.6	4.8	4.9	5.4	5.6	4.8*
34	4.2	3.9	4.5	4.7	3.9	3.8	3.7
35	4.1	4.1	3.7	4.0	4.1	4.6	4.7
36	3.5	3.6	3.6	4.2	4.8	4.9	5.9

* Denotes animals developing ventricular fibrillation

da matriz P (definida abaixo).

$$
\begin{array}{cc}
 & \begin{array}{ccccccc} (0) & (1) & (2) & (3) & (4) & (5) & (6) \end{array} \\
P = & \begin{bmatrix}
1 & -3 & 5 & -1 & 3 & -1 & 1 \\
1 & -2 & 0 & 1 & -7 & 4 & -6 \\
1 & -1 & -3 & 1 & 1 & -5 & 15 \\
1 & 0 & -4 & 0 & 6 & 0 & -20 \\
1 & 1 & -3 & -1 & 1 & 5 & 15 \\
1 & 2 & 0 & -1 & -7 & -4 & -6 \\
1 & 3 & 5 & 1 & 3 & 1 & 1
\end{bmatrix}
\end{array}
$$

As colunas da matriz P correspondem aos coeficien-

tes de polinômios ortogonais de graus 0, 1, 2, 3, 4, 5 e 6 para sete pontos igualmente espaçados.

O critério adotado foi o da Razão de Verossimilhança Generalizada de Wilks, e o resultado dos testes está apresentado na tabela 4.2.3.

TABELA 4.2.3

COLUNAS DE P UTILIZADAS NA CONSTRUÇÃO DE G_2	GRAU DO POLINÔMIO CUJO AJUSTE É TESTADO	DISTRIBUIÇÃO DA ESTATÍSTICA DE TESTE	G.L.	VALOR OBSERVADO	NÍVEL DESCRITIVO
(6)	5	F exata	4;32	1,6770	0,1788
(5),(6)	4	χ^2 assintótica	8	7,9785	0,5631
(4),(5),(6)	3	χ^2 assintótica	12	13,3254	0,3470
(3),(4),(5),(6)	2	χ^2 assintótica	16	34,8868	0,0049
(2),(3),(4),(5),(6)	1	χ^2 assintótica	20	41,4583	0,0040
(0),(1),(2),(3),(4),(5),(6)	sem interesse	-	-	-	-

Podemos então concluir que um polinômio de 3º grau se ajusta razoavelmente às observações. Vamos então considerar o seguinte modelo:

$$\beta_{it} = \xi_{i0} + \xi_{i1}t + \xi_{i2}t^2 + \xi_{i3}t^3 \qquad (4.2.2)$$

onde

β_{it} é o valor esperado da variável para o i-ésimo grupo no instante t.

ξ_{i0} é o coeficiente linear para o i-ésimo grupo.

ξ_{i1}, ξ_{i2}, ξ_{i3} são respectivamente os coeficientes dos dos termos de graus 1, 2 e 3 para o i-esimo grupo.

Neste caso o MLMC$(Y, X\xi G, \Sigma)$ seria definido por:

$$X = \begin{bmatrix} 1 & 0 & 0 & 0 \\ \vdots & & & \vdots \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \vdots & & & \vdots \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \vdots & & & \vdots \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ \vdots & & & \vdots \\ 0 & 0 & 0 & 1 \end{bmatrix} ; \xi = \begin{bmatrix} \xi_{10} & \xi_{11} & \xi_{12} & \xi_{13} \\ \vdots & \vdots & \vdots & \vdots \\ \xi_{40} & \xi_{41} & \xi_{42} & \xi_{43} \end{bmatrix} ; G = \begin{bmatrix} 1 & 1 & \dots & 1 \\ 1 & 3 & \dots & 13 \\ 1 & 3^2 & \dots & 13^2 \\ 1 & 3^3 & \dots & 13^3 \end{bmatrix}$$

Com a finalidade de simplificar os cálculos convém reescrever o modelo em função de polinômios ortogonais, ou seja:

$$\beta_{it} = \xi^*_{i0} + \xi^*_{i1}\psi_1(x) + \xi^*_{i2}\psi_2(x) + \xi^*_{i3}\psi_3(x) \qquad (4.2.3)$$

onde $\psi_i(x)$ é o polinômio ortogonal de grau i para sete pontos igualmente espaçados.

Então o MLMC[Y,Xξ*G*,Σ*] seria definido por:

$$\xi^* = \begin{bmatrix} \xi^*_{10} & \xi^*_{11} & \xi^*_{12} & \xi^*_{13} \\ \vdots & \vdots & \vdots & \vdots \\ \xi^*_{40} & \xi^*_{41} & \xi^*_{42} & \xi^*_{43} \end{bmatrix}, \text{ e } G^* = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ -3 & -2 & -1 & 0 & 1 & 2 & 3 \\ 5 & 0 & -3 & -4 & -3 & 0 & 5 \\ -1 & 1 & 1 & 0 & -1 & -1 & 1 \end{bmatrix}$$

A matriz G_1 pode ser definida de tal forma que suas colunas sejam as colunas (0), (1), (2) e (3) de P; a matriz G_2, de tal forma que suas colunas sejam as colunas (4), (5) e (6) de P.

Também como resultado da primeira análise obtemos a matriz abaixo, que é uma estimativa da matriz de correla-

ção entre as variáveis do espaço de estimação e as variáveis do espaço de erros.

		YG_1				YG_2		
		(0)	(1)	(2)	(3)	(4)	(5)	(6)
YG_1	(0)	1,00						
	(1)	0,44	1,00					
	(2)	-0,21	0,22	1,00				
	(3)	-0,24	-0,38	-0,02	1,00			
YG_2	(4)	0,24	-0,25	-0,73*	0,23	1,00		
	(5)	-0,06	0,05	0,05	-0,56*	-0,09	1,00	
	(6)	0,05	-0,04	-0,07	0,35	0,05	-0,01	1,00

* significativamente diferente de zero com nível de significância $\alpha = 0,05$.

Concluímos que somente as colunas (4) e (5) da matriz correspondente ao espaço de erros devem ser incluídas como variáveis auxiliares. *Redefinimos* então a matriz G_2, *desprezando* a coluna (6) e fazemos a transformação $Y_1 = YG_1$ e $Y_2 = YG_2$ prosseguindo na segunda análise com o modelo condicional (2.4.3).

Devemos ressaltar que os coeficientes de polinômios ortogonais encontrados em tabelas, geralmente não são padronizados. Isso implica que $GG_1 = F$ onde F é uma matriz diagonal diferente da identidade. Os cálculos podem ser realizados com a utilização desses coeficientes não padronizados, tomando-se o cuidado de corrigir o estimador obtido, dividindo-se cada coluna pela soma de quadrados dos coeficientes de polinômios ortogonais correspondentes.

A estimativa de ξ^* obtida para este exemplo é dada por:

$$\hat{\xi}^* = \begin{bmatrix} 4,6792 & 0,1505 & -0,0539 & -0,1780 \\ 3,5854 & -0,0236 & -0,0170 & 0,0275 \\ 4,0662 & 0,1385 & -0,0308 & -0,0708 \\ 3,9128 & 0,0672 & -0,0151 & -0,0056 \end{bmatrix}$$

Para obter a estimativa da matriz ξ basta fazer uma transformação de variáveis e substituir em (4.2.3) os polinômios ortogonais obtidos de uma tabela (Pearson & Hartley, 1958, por exemplo).

A tabela 4.2.4 apresenta estimativas de ξ^* e dos desvios-padrão dos elementos de $\hat{\xi}^*$ quando consideramos duas ou nenhuma variável auxiliar.

TABELA 4.2.4

	VARIÁVEIS AUXILIARES							
	NENHUMA				COLUNAS (4) E (5) DE Y_2			
ESTIMATIVA DE ξ^*	4,6222	0,1639	-0,0034	-0,1833	4,6792	0,1505	-0,0539	-0,1780
	3,5529	-0,0164	-0,0081	0,0017	3,5854	-0,0236	-0,0170	0,0275
	4,1446	0,1205	-0,0563	-0,0458	4,0662	0,1385	-0,0308	-0,0708
	3,9397	0,0615	-0,0213	0,0260	3,9128	0,0672	-0,0151	-0,0056
DESVIOS-PADRÃO ESTIMADOS DOS ELEMENTOS DE $\hat{\xi}^*$	0,1733	0,0374	0,0184	0,0485	0,1799	0,0387	0,0133	0,0418
	0,1644	0,0354	0,0174	0,0460	0,1675	0,0360	0,0124	0,0390
	0,1838	0,0396	0,0195	0,0514	0,1933	0,0415	0,0143	0,0450
	0,1733	0,0314	0,0184	0,0485	0,1770	0,0380	0,0131	0,0411

Pode-se observar que há um decréscimo sensível na variância estimada das estimativas dos coeficientes dos termos de segundo e terceiro graus, quando fazemos a análise com as duas variáveis auxiliares escolhidas. Porém há um aumento da variância estimada das estimativas dos coeficientes dos termos de primeiro grau e independente. Isto se justifica pelo fato de as variáveis auxiliares terem uma maior cor

relação com os termos de segundo e terceiro graus do que com os termos independente e de primeiro grau.

EXEMPLO 4.2.3 - Neste exemplo discutimos o seguinte problema apresentado ao SEA para análise estatística.

Uma amostra de 42 ratos foi dividida em 6 grupos de 7 animais e cada grupo foi alimentado com um dos seguintes tipos de ração:

CASEÍNA { ad libitum : grupo G_1
CASEÍNA { com restrição: grupo G_2

SOJA { ad libitum : grupo G_3
SOJA { com restrição: grupo G_4

SOJA + METIONINA { ad libitum : grupo G_5
SOJA + METIONINA { com restrição: grupo G_6

Em cada animal foram realizadas 5 observações da variável peso em gramas (a primeira antes do início do experimento, e as outras quatro a intervalos de uma semana após seu início) e um dos objetivos do experimento era explicar o comportamento da variável através de uma função, e comparar o efeito de cada tipo de ração no comportamento da variável peso no decorrer do tempo (4 semanas).

As médias das observações para cada um dos grupos nos instantes considerados estão representados na tab.4.2.5.

TABELA 4.2.5

	t_0	$t_1=1$	$t_2=2$	$t_3=3$	$t_4=4$
G_1	41,36	58,26	83,37	103,50	125,13
G_2	44,17	49,76	60,07	69,87	84,60
G_3	41,99	48,63	59,11	65,17	72,30
G_4	45,26	43,06	48,76	53,86	59,28
G_5	42,66	56,39	73,96	98,20	124,51
G_6	45,44	47,31	55,79	71,10	88,99

As medidas realizadas no instante t_0 (antes do início do experimento) foram utilizadas para um teste de homogeneidade do peso inicial; não havendo evidência estatística para rejeição da hipótese, os pesos iniciais foram considerados homogêneos e as medidas referentes aos pesos iniciais foram desprezadas.

Para a aplicação da técnica de Análise de Curvas de Crescimento, consideramos inicialmente o $MLM(Y, X\xi^*, \Sigma)$, definimos a matriz P abaixo obtida através de uma tabela de coeficientes de polinômios ortogonais,

$$P = \begin{array}{c} \begin{array}{cccc} (0) & (1) & (2) & (3) \end{array} \\ \begin{bmatrix} 1 & -3 & 1 & -1 \\ 1 & -1 & -1 & 3 \\ 1 & 1 & -1 & -3 \\ 1 & 3 & 1 & 1 \end{bmatrix} \end{array}$$

e testamos o ajuste de curvas polinomiais de graus 2 e 1 considerando hipóteses da forma H: $I\xi^* G_2 = 0$.

O critério adotado foi o da Razão de Verossimilhança Generalizada de Wilks, e o resultado dos testes está apresentado na tabela 4.2.6.

TABELA 4.2.6

COLUNAS DE P UTILIZADAS NA CONSTRUÇÃO DE G_2	GRAU DO POLINÔMIO CUJO AJUSTE É TESTADO	DISTRIBUIÇÃO DA ESTATÍSTICA DE TESTE	G.L.	VALOR OBSERVADO	NÍVEL DESCRITIVO
(3)	2	F exata	(6,34)	6,1976	0,0003
(2),(3)	1	χ^2 assintótica	12	49,5981	0,0001

As hipóteses devem ser rejeitadas nos dois casos, de onde podemos concluir que polinômios do 1º e 2º graus não são convenientes para explicar o comportamento da variável

no tempo (para os seis grupos considerados simultaneamente).

O próximo passo seria tentar ajustar um polinômio de terceiro grau às observações. Entretanto, além de não podermos testar o ajuste desse modelo do terceiro grau, (pois só temos 4 observações por animal), o número de parâmetros a estimar (24, no caso) seria equivalente ao número de parâmetros do modelo usual de Análise de Perfis, discutido por Morrison, (1976), por exemplo. A comparação entre os efeitos dos diversos tipos de ração no comportamento da variável peso pode então ser analisada segundo essa técnica.

- □ -

CAPÍTULO 5

CONCLUSÃO

Os resultados da técnica de análise de curvas de crescimento apresentados neste trabalho estão baseados principalmente nos trabalhos de Potthoff e Roy, (1964),Rao,(1965, 1966 e 1967), Khatri, (1966) e Grizzle e Allen, (1969). Embora a técnica desenvolvida por esses autores possa ser aplicada a uma grande classe de problemas envolvendo diversos esquemas de planejamento de experimentos, sentimos que dois aspectos do problema de curvas de crescimento deverão ser objetos de outras pesquisas:

i) o desenvolvimento de métodos computacionais mais simples para a análise de problemas onde as observações não são igualmente espaçadas;

ii) a extensão da técnica para análise de situações onde o comportamento da variável estudada não pode ser explicado através de modelos polinomiais.

O problema citado em i) foi abordado por Hills, (1968) sob o ponto de vista de diferenças finitas, e o problema de modelos não lineares de crescimentos foi estudado por Allen, (1967). Referências sobre esses autores são apresentadas no final do trabalho, juntamente com citações de outros trabalhos mais recentes na área de curvas de crescimen

to. Esses artigos lidam com aspectos específicos do problema, e não apresentam contribuições substanciais à técnica discutida neste trabalho.

- □ -

APÊNDICE A

O PROBLEMA DE ESTIMAÇÃO EM MODELOS LINEARES MULTIVARIADOS

O problema considerado neste apêndice é o de obter estimadores dos parâmetros β, Σ e de combinações lineares da forma $C\beta U$ a partir de um conjunto de dados para os quais admitimos válidas as suposições do $MLM(Y, X\beta, \Sigma)$. O problema da determinação de estimadores de $C\beta U$ utilizando o $MLM(Y, X\beta, \Sigma)$ fica reduzido ao problema da determinação de estimadores de $C\beta^*$ utilizando o $MLM(Y^*, X\beta^*, \Sigma^*)$ obtido através da transformação $Y^* = YU$, $\beta^* = \beta U$, $\Sigma^* = U'\Sigma U$.

Na seção A.1 determinamos um estimador de $C\beta$ através de uma aplicação do critério de mínimos quadrados. Além disso obtemos um estimador não-viciado de Σ.

Na seção A.2 consideramos uma aplicação do critério de máxima verossimilhança, obtendo estimadores de β e Σ. Um estimador de $C\beta$ é obtido através da aplicação de uma propriedade de estimadores de máxima verossimilhança. Observamos também que esse estimador de Σ é equivalente (a menos de uma constante) ao estimador de Σ obtido na seção A.1.

Na seção A.3 mostramos que os estimadores de $C\beta$ obtidos segundo os critérios de mínimos quadrados e de máxima verossimilhança são equivalentes, e correspondem ao estimador linear não viciado de mínima variância.

Finalmente na seção A.4 apresentamos as matrizes de variância-covariância e a distribuição do estimador de $C\beta$ obtido, além da distribuição do estimador de Σ.

A.1 - UMA APLICAÇÃO DE CRITÉRIO DE MÍNIMOS QUADRADOS AO MODELO LINEAR MULTIVARIADO

O estimador de mínimos quadrados do modelo linear *univariado* é obtido através da minimização da forma quadrática $(y-X\beta)'(y-X\beta)$, que representa a soma de quadrados dos desvios entre os valores observados e os valores esperados. Para o modelo multivariado esse método não pode ser aplicado diretamente, pois nesse caso o produto $(Y-X\beta)'(Y-X\beta)$ corresponde a uma matriz cujos elementos são formas quadráticas ou bilineares, cada um representando uma soma de quadrados ou de produtos dos desvios entre os valores observados e esperados de componentes da variável multivariada.

Uma adaptação do método de mínimos quadrados para o caso multivariado pode ser obtida escrevendo-se o MLM$(Y, X\beta, \Sigma)$ na forma:

$$E(y_i) = X\beta_i \qquad (i=1,2,\ldots,p)$$

$$Cov(y_i, y_j) = \sigma_{ij} I \qquad (i,j=1,2,\ldots,p)$$

Para cada i estamos diante de um modelo linear univariado, e o estimador de mínimos quadrados de $C\beta_i$ é dado por:

$$C\hat{\beta}_i = C(X'X)^{-1}X'y_i$$

Dispondo os vetores $C\hat{\beta}_i$ $(i=1,2,\ldots,p)$ como colunas de uma matriz, obtemos:

$$[C\hat{\beta}_1 \; C\hat{\beta}_2 \ldots C\hat{\beta}_p] = C(X'X)^{-1}X'[y_1 \; y_2 \ldots y_p] = C(X'X)^{-1}X'Y = \widehat{C\beta} \qquad (A.1.1)$$

que definiremos como estimador de mínimos quadrados de $C\beta$.

Tomando $C = I$ obtemos o estimador de mínimos quadrados de β:

$$\hat{\beta} = (X'X)^{-1}X'Y \qquad (A.1.2)$$

A forma do estimador (A.1.1) é semelhante àquela do caso univariado (com o vetor y substituído pela matriz Y), porém neste caso devemos entender (A.1.1) como o estimador de $C\beta$ cujas colunas são estimadores de mínimos quadrados obtidos através de um modelo univariado onde consideramos somente uma das componentes da variável multivariada. Como veremos na seção A.3, o estimador (A.1.1) corresponde ao estimador linear não viciado de mínima variância obtido quando consideramos todas as componentes da variável multivariada em conjunto.

Utilizando a transformação sugerida na introdução deste apêndice obteremos o estimador de mínimos quadrados de $C\beta U$:

$$C\hat{\beta}U = C(X'X)^{-1}X'YU \qquad (A.1.3)$$

Seja $\Sigma = \{\sigma_{ij}\}$ $(i,j=1,2,\ldots,p)$ e determinemos um estimador não viciado de Σ. Consideremos o modelo univariado $y_i = X\beta_i + \varepsilon_i$ onde $\varepsilon_i \sim N_n(0,\sigma_{ii}I)$. Sabemos que neste caso um estimador não viciado de σ_{ii} é dado por:

$$\hat{\sigma}_{ii} = \frac{S_0(i,i)}{n-r} \qquad (A.1.4)$$

onde

$$S_0(i,i) = \min_{\beta_i}(y_i - X\beta_i)'(y_i - X\beta_i) = (y_i - X\hat{\beta}_i)'(y_i - X\hat{\beta}_i)$$

Determinemos então um estimador não viciado de σ_{ij}. Com essa finalidade vamos definir a variável $y^* = y_i + y_j$. Então:

$$E(y^*) = E(y_i + y_j) = X(\beta_i + \beta_j) = X\beta^* \text{ onde } \beta^* = \beta_i + \beta_j$$

$$V(y^*) = V(y_i+y_j) = \sigma_{ii}I+\sigma_{jj}I+2\sigma_{ij}I = \sigma^2 I \text{ onde } \sigma^2=\sigma_{ii}+\sigma_{jj}+2\sigma_{ij}$$

Podemos considerar o modelo univariado $y^* = X\beta^*+\varepsilon^*$ onde $\varepsilon^* \sim N_n(0,\sigma^2 I)$ e o estimador de mínimos quadrados de β^* será $\hat{\beta}^* = (X'X)^{-1}X'y^* = \hat{\beta}_i+\hat{\beta}_j$. Utilizando (A.1.4), obteremos um estimador não viciado de σ^2:

$$\hat{\sigma}^2 = \frac{S_0^{*2}}{n-r}$$

onde:

$$S_0^{*2} = (y^*-X\hat{\beta}^*)'(y^*-X\hat{\beta}^*) = (y_i+y_j-X\hat{\beta}_i-X\hat{\beta}_j)'(y_i+y_j-X\hat{\beta}_i-X\hat{\beta}_j) =$$

$$= S_0(i,i)+S_0(j,j)+2S_0(i,j).$$

Então:

$$E(S_0^{*2}) = (n-r)\sigma^2 = (n-r)(\sigma_{ii}+\sigma_{jj}+2\sigma_{ij}) =$$

$$= E[S_0(i,i)]+E[S_0(j,j)]+2E[S_0(i,j)]$$

Logo:

$$E\left[\frac{S_0(i,j)}{n-r}\right] = \sigma_{ij} \qquad (A.1.5)$$

onde

$$S_0(i,j) = \min_{\beta_i,\beta_j} (y_i-X\beta_i)'(y_j-X\beta_j) = (y_i-X\hat{\beta}_i)'(y_j-X\hat{\beta}_j).$$

Os resultados (A.1.4) e (A.1.5) nos permitem obter estimadores não viciados para todos os elementos de Σ.

A matriz $S = \{S_0(i,j)\}$ $(i,j=1,2,\ldots,p)$ é a matriz de soma de quadrados e produtos cruzados do resíduo e pode ser considerada como uma generalização da soma de quadrados do

resíduo no modelo univariado. Cada elemento de S pode ser escrito na forma:

$$S_0(i,j) = y_i'y_j - y_i'X(X'X)^{-1}X'y_j$$

e então

$$S = Y'[I-X(X'X)^{-1}X']Y \qquad (A.1.6)$$

Concluímos que um estimador não viciado de Σ é dado por:

$$\hat{\Sigma} = \frac{1}{n-r} Y'[I-X(X'X)^{-1}X']Y = \frac{1}{n-r} S \qquad (A.1.7)$$

Devemos observar que na determinação dos estimadores através da aplicação da técnica de mínimos quadrados não utilizamos nenhuma suposição sobre a distribuição das variáveis aleatórias consideradas, isto é, das linhas da matriz ε.

A.2 - UMA APLICAÇÃO DO CRITÉRIO DE MÁXIMA VEROSSIMILHANÇA AO MODELO LINEAR MULTIVARIADO

Consideremos o MLM($Y, X\beta, \Sigma$). Com a finalidade de preservar a notação usual para vetores e matrizes (onde x representa um vetor coluna e x' um vetor linha), adotamos para esta seção uma formulação equivalente do modelo acima, obtida pela transposição das matrizes envolvidas. Sejam então:

$$\begin{aligned} Y' &= U \quad (p\times n) \\ \beta' &= \theta \quad (p\times r) \\ X' &= Z \quad (r\times n) \\ \varepsilon' &= \varepsilon^* \quad (p\times n) \end{aligned} \qquad (A.2.1)$$

O modelo transposto pode ser escrito na forma $U = \theta Z+\varepsilon^*$ onde $E(U) = \theta Z$ e cada *coluna* da matriz ε^* tem distribuição normal

multivariada com valor esperado 0 e matriz de variância-covariância Σ (p×p). Vamos utilizar para este modelo a notação MLM'$(U,\theta Z,\Sigma)$. Além disso sejam $\Sigma = \{\sigma_{ij}\}$ e $\Sigma^{-1} = \{\sigma^{ij}\}$.

A função de verossimilhança para uma amostra de tamanho n de uma distribuição normal multivariada com valor esperado θZ e matriz de variância-covariância Σ pode ser escrita como:

$$L = (2\pi)^{-\frac{1}{2}np} \left|\Sigma^{-1}\right|^{\frac{1}{2}n} \exp\left[-\frac{1}{2}\sum_{\alpha=1}^{n}(u_\alpha-\theta z_\alpha)'\Sigma^{-1}(u_\alpha-\theta z_\alpha)\right] \quad (A.2.2)$$

onde u_α $(\alpha=1,2,\ldots,n)$ corresponde a α-ésima coluna da matriz U.

Nosso objetivo é determinar as estimativas $\hat{\theta}$ e $\hat{\Sigma}$ de θ e Σ que maximizam L (considerada como função de θ e Σ).

Primeiramente vamos escrever cada uma das formas quadráticas do expoente de L em função dos elementos de cada vetor u_α.

$$(u_\alpha-\theta z_\alpha)'\Sigma^{-1}(u_\alpha-\theta z_\alpha) = \sum_{i=1}^{p}\sum_{j=1}^{p}(u_{i\alpha} - \sum_{k=1}^{r}\theta_{ik}z_{k\alpha})\sigma^{ij}(u_{j\alpha} - \sum_{k=1}^{r}\theta_{jk}z_{k\alpha}).$$

Tomando o logaritmo da função de verossimilhança, obtemos:

$$\log L = -\frac{1}{2}np\log 2\pi + \frac{1}{2}n\log\left|\Sigma^{-1}\right| +$$

$$-\frac{1}{2}\sum_{\alpha=1}^{n}\left[\sum_{i=1}^{p}\sum_{j=1}^{p}(u_{i\alpha} - \sum_{k=1}^{r}\theta_{ik}z_{k\alpha})\sigma^{ij}(u_{j\alpha} - \sum_{k=1}^{r}\theta_{jk}z_{k\alpha}\right] \quad (A.2.3)$$

Utilizando os resultados apresentados por Anderson (1958-seção 3.2), podemos concluir que o máximo de log L em relação a θ e Σ é igual ao máximo de log L em relação a θ e Σ^{-1}. Além disso, se $\hat{\Sigma}^{-1}$ maximiza log L, então $\hat{\Sigma} = (\hat{\Sigma}^{-1})^{-1}$ também maximiza log L. Pode-se demonstrar ainda, que o máximo de log L é atingido quando $\frac{\partial \log L}{\partial \theta_{ij}} = 0$ $(i=1,2,\ldots p;\ j=1,2,\ldots,r)$. Pa

ra obter o máximo de log L utilizando as propriedades acima, demonstraremos o seguinte lema (Anderson, 1958 - seção 8.2):

<u>LEMA A.2.1</u> - Se $f(t_1,\ldots,t_m) = \sum_{i=1}^{m}\sum_{j=1}^{m} d_{ij}t_it_j$ onde $d_{ij} = d_{ji}$, então:

A) $$\frac{\partial f}{\partial v} = 2\sum_{i=1}^{m}\sum_{j=1}^{m} d_{ij}\frac{\partial t_i}{\partial v}t_j$$

B) $$\frac{\partial}{\partial v}\sum_{\alpha=1}^{n} f(t_{1\alpha},\ldots,t_{m\alpha}) = 2\sum_{\alpha=1}^{n}\left[\sum_{i=1}^{m}\sum_{j=1}^{m} d_{ij}\frac{\partial t_{i\alpha}}{\partial v}t_{j\alpha}\right]$$

$$\frac{\partial}{\partial v}\sum_{i=1}^{m}\sum_{j=1}^{m} d_{ij}t_it_j = \sum_{i=1}^{m}\sum_{j=1}^{m} d_{ij}\frac{\partial t_i}{\partial v}t_j + \sum_{i=1}^{m}\sum_{j=1}^{m} d_{ij}\frac{\partial t_i}{\partial v}t_i =$$

$$= \sum_{i=1}^{m}\sum_{j=1}^{m}(d_{ij}+d_{ji})\frac{\partial t_i}{\partial v}t_j = 2\sum_{i=1}^{m}\sum_{j=1}^{m} d_{ij}\frac{\partial t_i}{\partial v}t_j$$

O resultado A) está demonstrado e o resultado B) é obtido através de uma somatória em relação a α.

Vamos calcular $\frac{\partial \log L}{\partial \theta_{gh}}$ $(g=1,2,\ldots,p;\ h=1,2,\ldots,r)$.

Consideremos a transformação:

$$t_{i\alpha} = u_{i\alpha} - \sum_{k=1}^{r}\theta_{ik}z_{k\alpha} = u_{i\alpha} - (\theta_{i1}z_{i\alpha}+\ldots\theta_{ir}z_{r\alpha})$$

Então podemos escrever:

$$\log L = -\frac{1}{2}np\log 2\pi + \frac{1}{2}n\log|\{\sigma^{ij}\}| - \frac{1}{2}\sum_{\alpha=1}^{n}\left[\sum_{i=1}^{p}\sum_{j=1}^{p} t_{i\alpha}\sigma^{ij}t_{j\alpha}\right]$$

Utilizando o resultado B) do Lema A.2.1, temos:

$$\frac{\partial \log L}{\partial \theta_{gh}} = -\sum_{\alpha=1}^{n}\left[\sum_{i=1}^{p}\sum_{j=1}^{p}\sigma^{ij}\frac{\partial t_{i\alpha}}{\partial \theta_{gh}}t_{j\alpha}\right] = -\sum_{\alpha=1}^{n}\left[\sum_{j=1}^{p}\sigma^{hj}z_{h\alpha}t_{j\alpha}\right] =$$

$$= -\sum_{\alpha=1}^{n}\left[\sum_{j=1}^{p}\sigma^{ij}z_{h\alpha}(u_{j\alpha}-\sum_{k=1}^{r}\theta_{jk}z_{k\alpha}\right] = -\sum_{j=1}^{p}\sigma^{ij}\left[\sum_{\alpha=1}^{n}z_{h\alpha}u_{j\alpha}-\sum_{k=1}^{r}\theta_{jk}\left(\sum_{\alpha=1}^{n}z_{h\alpha}z_{k\alpha}\right)\right]$$

Fazendo:

$$c_{hj} = \sum_{\alpha=1}^{n} z_{h\alpha} u_{j\alpha}$$

$$a_{hk} = \sum_{\alpha=1}^{n} z_{h\alpha} z_{k\alpha} \qquad (A.2.4)$$

e igualando $\frac{\partial \log L}{\partial \theta_{gh}}$ a zero, obtemos:

$$\sum_{j=1}^{p} \sigma^{hj} \left[c_{hj} - \sum_{k=1}^{r} \theta_{jk} a_{hk} \right] = 0 \qquad (A.2.5)$$

Fazendo

$$b_j = c_{hj} - \sum_{k=1}^{r} \theta_{jk} a_{hk}$$

e escrevendo as equações (A.2.5) para todos os valores de h, temos:

$$\begin{aligned} \sigma^{11} b_1 + \sigma^{12} b_2 + \ldots + \sigma^{1p} b_p &= 0 \\ \sigma^{21} b_1 + \sigma^{22} b_2 + \ldots + \sigma^{2p} b_p &= 0 \\ &\vdots \\ \sigma^{p1} b_1 + \sigma^{p2} b_2 + \ldots + \sigma^{pp} b_p &= 0 \end{aligned} \qquad (A.2.6)$$

Como Σ^{-1} é positiva definida por hipótese (em quase toda parte), suas colunas são linearmente independentes, e para que sejam válidas as relações (A.2.6) devemos ter:

$$b_j = 0 \quad (j=1,2,\ldots,p),$$

que é equivalente a:

$$\sum_{k=1}^{r} \theta_{jk} a_{hk} = c_{hj} \quad (j=1,2,\ldots,p;\ h=1,2,\ldots,r).$$

Utilizando as relações (A.2.4), podermos escrever:

$$\sum_{k=1}^{r} \theta_{jk}\left(\sum_{\alpha=1}^{n} z_{h\alpha} z_{k\alpha}\right) = \sum_{\alpha=1}^{n} z_{h\alpha} u_{j\alpha} \Longrightarrow \sum_{\alpha=1}^{n}\left(\sum_{k=1}^{r} \theta_{jk} z_{h\alpha} z_{k\alpha}\right) = \sum_{\alpha=1}^{n} z_{h\alpha} u_{j\alpha}$$

que em notação matricial assume a forma $\theta ZZ' = UZ'$. Como Z tem posto-linha completo ZZ' é inversível e o estimador de máxima verossimilhança de θ será:

$$\hat{\theta} = UZ'(ZZ')^{-1} \qquad (A.2.7)$$

Através das relações (A.2.1) podemos escrever o resultado obtido utilizando a notação usual da $MLM(Y, X\beta, \Sigma)$. Então o estimador de máxima verossimilhança de β será:

$$\hat{\beta} = (X'X)^{-1}X'Y \qquad (A.2.8)$$

Para obter o estimador de máxima verossimilhança de $C\beta$ utilizaremos o resultado do teorema abaixo, cuja demonstração pode ser encontrada, por exemplo, em Mood, Graybill, Boes, (1974).

<u>TEOREMA A.2.2</u> - Seja $\hat{\theta} = (\hat{\theta}_1, \ldots, \hat{\theta}_k)$ onde $\hat{\theta}_j = \hat{\theta}_j(X_1, X_2, \ldots, X_n)$, um estimador de máxima verossimilhança de $\theta = (\theta_1, \ldots, \theta_k)$ na densidade $f(\cdot; \theta_1, \ldots, \theta_k)$. Se $\tau(\theta) = [\tau_1(\theta), \ldots, \tau_r(\theta)]$ $(1 \leq r \leq k)$ é uma transformação do espaço paramétrico, então um estimador de máxima verossimilhança de $\tau(\theta)$ será:

$$\tau(\hat{\theta}) = [\tau_1(\hat{\theta}), \ldots, \tau_r(\hat{\theta})].$$

Logo o estimador de máxima verossimilhança de $C\beta$ será:

$$C\hat{\beta} = C(X'X)^{-1}X'Y \qquad (A.2.9)$$

Utilizando a sugestão apresentada na introdução deste apêndice, podemos concluir que o estimador de máxima verossimilhança de $C\beta U$ é dado por:

$$C\hat{\beta}U = C(X'X)^{-1}X'YU \qquad (A.2.10)$$

Utilizando o resultado do Lema 3.2.2 de Anderson, (1958) podemos mostrar que, em relação a σ^{ij}, log L atinge um máximo para:

$$\hat{\Sigma}^{-1} = n\left[\sum_{\alpha=1}^{n}(u_\alpha-\hat{\theta}z_\alpha)(u_\alpha-\hat{\theta}z_\alpha)'\right]^{-1}$$

Então o estimador de máxima verossimilhança de Σ será:

$$\hat{\Sigma} = \left[\hat{\Sigma}^{-1}\right]^{-1} = \frac{1}{n}\sum_{\alpha=1}^{n}(u_\alpha\hat{\theta}z_\alpha)(u_\alpha-\hat{\theta}z_\alpha)' = \frac{1}{n}\sum_{\alpha=1}^{n}\left[u_\alpha u'_\alpha - u_\alpha z'_\alpha\hat{\theta}' - \hat{\theta}z_\alpha u'_\alpha + \hat{\theta}z_\alpha z'_\alpha\theta'\right] =$$

$$= \frac{1}{n}\left[\sum_{\alpha=1}^{n}u_\alpha u'_\alpha - \sum_{\alpha=1}^{n}u_\alpha z'_\alpha\hat{\theta}' - \hat{\theta}\sum_{\alpha=1}^{n}z_\alpha u'_\alpha + \hat{\theta}\sum_{\alpha=1}^{n}z_\alpha z'_\alpha\hat{\theta}'\right] =$$

$$= \frac{1}{n}[UU'-UZ'(ZZ')^{-1}ZU'-UZ'(ZZ')^{-1}ZU'+UZ'(ZZ')^{-1}(ZZ')(ZZ')^{-1}ZU'] =$$

$$= \frac{1}{n}[UU'-UZ'(ZZ')^{-1}ZU'] = \frac{1}{n}U[I-Z'(ZZ')^{-1}Z']U$$

Através das relações (A.2.1) podemos escrever o resultado acima utilizando a notação usual do MLM($Y, X\beta, \Sigma$), e então:

$$\hat{\Sigma} = \frac{1}{n}Y'[I-X(X'X)^{-1}X']Y = \frac{1}{n}S \qquad (A.2.11)$$

Observando os resultados (A.1.7) e (A.2.11) concluímos que os estimadores de máxima verossimilhança e mínimos quadrados de Σ são equivalentes a menos das constantes $\frac{1}{n}$ e $\frac{1}{n-r}$, pois o estimador de mínimos quadrados de Σ é dado por

$$\hat{\Sigma} = \frac{1}{n-r}S.$$

A.3 - DETERMINAÇÃO DOS ESTIMADORES LINEARES NÃO VICIADOS DE MÍNIMA VARIÂNCIA

Consideremos o MLM($Y, X\beta, \Sigma$) escrito na forma:

$$E(y_i) = X\beta_i \qquad (i=1,2,\ldots,p)$$

$$Cov(y_i, y_j) = \sigma_{ij} I \qquad (i,j=1,2,\ldots,p)$$

Vaamos primeiramente estudar a estimação de funções paramétricas da forma $c'\beta_i$ (onde c é um vetor r×1) que envolvem somente os parâmetros relativos à i-ésima componente da variável multivariada. O estimador procurado deve ser função de *todas as observações* e deve apresentar as seguintes características:

i) ser linear
ii) ser não viciado
iii) apresentar variância mínima na classe dos estimadores lineares não viciados.

Uma função linear de todas as observações pode ser escrita na forma:

$$a_1'y_1 + \ldots + a_p'y_p = \underset{(1\times np)}{[a_p' \ldots a_p']} \underset{(np\times 1)}{\begin{bmatrix} y_1 \\ \vdots \\ y_p \end{bmatrix}}$$

Como o estimador procurado deve ser não viciado, então:

$$E(a_1'y_1) + \ldots + E(a_p'y_p) = a_1'X\beta_1 + \ldots + a_p'X\beta_p = c'\beta_i$$

o que implica: $a_i'X = c'$ e $a_j'X = 0$, $\forall j \neq i$.

As variedades lineares geradas pelas colunas de X'

e $X'X$, $M(X')$ e $M(X'X)$, respectivamente, são tais que

$$M(X') = M(X'X).$$

Então, como $c = X'a_i$, existe um vetor λ $(r\times 1)$ tal que:

$$c = X'X\lambda \Longrightarrow \lambda = (X'X)^{-1}c$$

Tomemos a função linear $c'(X'X)^{-1}X'y_i$. Então:

$$E[c'(X'X)^{-1}X'y_i] = c'(X'X)^{-1}X'\beta_i.$$

Mostremos que essa função linear corresponde ao estimador de mínima variância na classe dos estimadores lineares não viciados.

$$V\left[\sum_{i=1}^{p} a_i'y_i\right] = V\left\{\left[\sum_{i=1}^{p} a_i'y_i - c'(X'X)^{-1}X'y_i\right] + c'(X'X)^{-1}X'y_i\right\} =$$

$$= V\left[\sum_{i=1}^{p} a_i'y_i - c'(X'X)^{-1}X'y_i\right] + V\left[c'(X'X)^{-1}X'y_i\right] +$$

$$+ 2\ \mathrm{cov}\left[\sum_{i=1}^{p} a_i'y_i - c'(X'X)^{-1}X'y_i ; c'(X'X)^{-1}X'y_i\right]$$

Mas

$$\mathrm{cov}\left[\sum_{i=1}^{p} a_i'y_i - c'(X'X)^{-1}X'y_i ;\ c'(X'X)^{-1}X'y_i\right] =$$

$$= \sum_{j=1}^{p} c'(X'X)^{-1}X'\mathrm{cov}(y_i,y_j)a_j - c'(X'X)^{-1}X'\mathrm{cov}(y_i,y_i)X(X'X)^{-1}c =$$

$$= \sum_{j=1}^{p} c'(X'X)^{-1}X'a_j\sigma_{ij} - c'(X'X)^{-1}X'X(X'X)^{-1}c\sigma_{ii} =$$

$$= c'(X'X)^{-1}c\sigma_{ii} - c'(X'X)^{-1}c\,\sigma_{ii} = 0, \text{ pois } X'a_j = 0,\ \forall j\neq i.$$

Logo

$$V\left[c'(X'X)^{-1}X'y_i\right] = V\left[\sum_{i=1}^{p} a_i'y_i\right] - V\left[\sum_{i=1}^{p} a_i'y_i - c'(X'X)^{-1}X'y_i\right] \leq V\left[\sum_{i=1}^{p} a_i'y_i\right]$$

Então o estimador desejado é dado por:

$$c'\hat{\beta}_i = c'(X'X)^{-1}X'y_i \qquad (A.3.1)$$

Para obter o estimador linear não viciado de mínima variância de $C\beta$, basta considerar as linhas c_i' $(i=1,\ldots,c)$ da matriz C e obter os estimadores de $c_i'\beta_j$ $(i=1,\ldots,c;j=1,\ldots,p)$ através de (A.3.1). O estimador determinado dessa forma pode ser escrito como:

$$C\hat{\beta} = C(X'X)^{-1}X'Y \qquad (A.3.2)$$

Tomando $C = I$ obtemos o estimador linear não viciado de mínima variância de β:

$$\hat{\beta} = (X'X)^{-1}X'Y \qquad (A.3.3)$$

Utilizando a sugestão apresentada na introdução deste apêndice obtemos o estimador linear não viciado de mínima-variância de $C\beta U$:

$$C\hat{\beta}U = C(X'X)^{-1}X'YU \qquad (A.3.4)$$

Analisando as expressões (A.1.3), (A.2.10) e (A.3.4) concluímos que os estimadores de mínimos quadrados, de máxima verossimilhança e o estimador linear não viciado de mínima variância de $C\beta U$ são obtidos através da mesma função das observações.

A distribuição dos estimadores obtidos nas seções A.1, A.2 e A.3 será estudada na seção seguinte.

A.4 - DISTRIBUIÇÃO DOS ESTIMADORES OBTIDOS

Consideremos primeiramente estimadores de $c_g'\beta_i$ e $c_h'\beta_j$, onde c_g e c_h são vetores $(r\times 1)$, segundo qualquer dos métodos considerados neste apêndice. Conforme o resultado (A.3.1) esses estimadores podem ser escritos na forma:

$$c_g'\hat{\beta}_i = c_g'(X'X)^{-1}X'y_i \quad \text{e} \quad c_h'\hat{\beta}_j = c_h'(X'X)^{-1}X'y_j .$$

Utilizando resultados da teoria univariada podemos obter as variâncias desses estimadores, bem como a covariância entre os mesmos:

$$\begin{aligned} V(c_g'\hat{\beta}_i) &= V[c_g'(X'X)^{-1}X'y_i] = c_g'(X'X)^{-1}c_g\sigma_{ii} \\ V'c_h'\hat{\beta}_j) &= V[c_h'(X'X)^{-1}X'y_j] = c_h'(X'X)^{-1}c_h\sigma_{jj} \qquad (A.4.1) \\ Cov(c_g'\hat{\beta}_i, c_h'\hat{\beta}_j) &= Cov[c_g'(X'X)^{-1}X'y_i; c_h'(X'X)^{-1}X'y_j] = c_g'(X'X)^{-1}c_h\sigma_{ij} \end{aligned}$$

Consideremos agora o estimador $C\hat{\beta} = C(X'X)^{-1}X'Y$ de $C\beta$, escrito na forma de um vetor $(cp\times 1)$. Cada elmento desse vetor é da forma $c_g'(X'X)^{-1}X'y_i$, $(g=1,\ldots,p;\ i=1,\ldots,p)$. Utilizando (A.4.1) podemos escrever simbolicamente:

$$V(C\hat{\beta}) = \Sigma\otimes[C(X'X)^{-1}C'] \qquad (A.4.2)$$

onde $\otimes$ representa o produto de Kronecker, definido como:

$$\underset{(p\times q)}{A} \otimes \underset{(m\times n)}{B} = \begin{bmatrix} a_{11}B & \ldots & a_{1q}B \\ \vdots & & \vdots \\ a_{p1}B & \ldots & a_{pq}B \end{bmatrix}$$

Tomando $C = I$, podemos escrever:

$$V(\hat{\beta}) = \Sigma\otimes(X'X)^{-1} \qquad (A.4.3)$$

E novamente utilizando a sugestão apresentada na introdução deste apêndice, teremos:

$$V(C\hat{\beta}U) = (U'\Sigma U)\otimes[C(X'X)^{-1}C'] \qquad (A.4.4)$$

Consideremos o vetor (np×1) obtido quando escrevemos as colunas de Y uma seguida da outra. Esse vetor tem distribuição normal multivariada com valor esperado $X\beta$ — considerada com um vetor (np×1) obtido quando escrevemos suas colunas uma em seguida da outra — e matriz de variância covariância $\Sigma\otimes I$. Como os estimadores de $C\beta$ e $C\beta U$ correspondem a combinações lineares dos elementos de Y, podemos concluir que esses estimadores têm distribuição normal multivariada aplicando resultados clássicos sobre essa distribuição (Rao, 1973, por exemplo).

Pode-se mostrar, também, que a matriz

$$S = Y'[I-X(X'X)^{-1}X']Y$$

tem distribuição de Wishart, central, com parâmetros (n - r) e Σ,[$S \sim W_p(n-r,\Sigma,\cdot)$]. Então qualquer estimador $\hat{\Sigma}$ da forma kS terá distribuição proporcional a distribuição de Wishart (Rao, 1973, por exemplo).

- □ -

APÊNDICE B

O PROBLEMA DE TESTES DE HIPÓTESES EM MODELOS LINEARES MULTIVARIADOS

Consideramos neste apêndice o problema de determinação de testes de significância para a hipótese linear geral H: $C\beta U = \Gamma$ a partir de um conjunto de dados para os quais admitimos válidas as suposições do $MLM(Y, X\beta, \Sigma)$.

Fazendo a transformação $\beta U = \beta^*$, o problema de obtenção de um teste para a hipótese H: $C\beta U = \Gamma$ considerado acima é equivalente ao problema da obtenção de um teste para a hipótese H: $C\beta^* = \Gamma$ admitindo o $MLM(Y^*, X\beta^*, \Sigma^*)$ onde $Y^* = YU$ e $\Sigma^* = U'\Sigma U$. Dessa forma analisamos somente hipóteses da forma H: $C\beta = \Gamma$ utilizando o $MLM(Y, X\beta, \Sigma)$.

Nas seções B.1 e B.2 obtemos testes de significância para a hipótese considerada, através da aplicação de dois critérios heurísticos: o critério da União-Intersecção de Roy e o critério da Razão de Verossimilhança Generalizada (estudado por Wilks, no caso multivariado). Obtemos as estatísticas utilizadas nos testes e fazemos considerações gerais sobre a distribuição das mesmas, apresentando os principais resultados obtidos por diversos pesquisadores.

Na seção B.3 observamos que as estatísticas de teste obtidas através dos dois critérios são funções das raízes da equação $|H-\lambda S| = 0$, onde H é a matriz de somas de quadrados e produtos cruzados devida ao desvio em relação à hipótese nula e S é a matriz de somas de quadrados e produtos cruzados devida ao resíduo. São apresentadas outras estatísticas de teste e algumas considerações gerais sobre suas distribuições e propriedades, sendo citados os resultados mais importantes.

B.1 - O PRINCÍPIO DA UNIÃO-INTERSECÇÃO DE ROY

Consideremos o MLM$(Y, X\beta, \Sigma)$. Nosso objetivo é determinar um teste de significância para a hipótese $H: C\beta = \Gamma$, que pode ser escrita na forma equivalente $H: C\beta_i = \Gamma_i^*$ $(i=1,\ldots,p)$.

Vamos primeiramente salientar as semelhanças existentes na estrutura do problema de testes de hipóteses para os casos univariado e multivariado. Tomemos um vetor ℓ $(p\times 1)$ arbitrário, porém fixo, e façamos uma redução do problema proposto para o caso univariado, considerando o problema de obtenção de um teste de significância para a hipótese $H': C\beta^* = \Gamma^*$ admitindo o modelo univariado $y^* = X\beta^*+\varepsilon^*$, onde $y^* = Y\ell$, $\beta^* = \beta\ell$, $\Gamma^* = \Gamma\ell$ e ε^* tem distribuição normal multivariada com valor esperado 0 e matriz de variância - covariância $\sigma_\ell^2 I$, $\sigma_\ell^2 = \ell'\Sigma\ell$.

Utilizando os resultados da teoria univariada (Searle, 1971, cap. 3, por exemplo) sabemos que um teste para a hipótese H' é construído com base nas somas de quadrados devida ao resíduo (S_0^2) e devida ao desvio em relação à hipótese nula $(R_1^2 - S_0^2)$, onde:

$$S_0^2 = \min_{\beta^*}(y^*-X\beta^*)'(y^*-X\beta^*)$$

$$R_1^2 = \min_{C\beta^*=\Gamma^*}(y^*-X\beta^*)'(y^*-X\beta^*)$$

que podem ser expressas na forma:

$$S_0^2 = y^{*\prime}[I-X(X'X)^{-1}X']y^*$$

$$R_1^2 = y^{*\prime}[I-X(X'X)^{-1}X']y^* + \qquad (B.1.1)$$

$$+ [C(X'X)^{-1}X'y^*-\Gamma^*]'[C(X'X)^{-1}C']^{-1}[C(X'X)^{-1}X'y^*-\Gamma^*]$$

Com a finalidade de determinar expressões análogas para o caso multivariado, vamos definir:

$$R_1(i,j) = \min_{\substack{C\beta_i=\Gamma_i \\ C\beta_j=\Gamma_j}} (y_i-X\beta_i)'(y_j-X\beta_j) = y_i'[I-X(X'X)^{-1}X']y_j +$$

$$+ [C(X'X)^{-1}X'y_i-\Gamma_i]'[C(X'X)^{-1}C']^{-1}[C(X'X)^{-1}X'y_j-\Gamma_j]$$

e considerar a matriz $R = \{R_1(i,j)\}$, $(i,j=1,2,\ldots,p)$, que pode ser escrita na forma:

$$R = Y'[I-X(X'X)^{-1}X']Y + [C(X'X)^{-1}X'Y-\Gamma]'[C(X'X)^{-1}C']^{-1}[C(X'X)^{-1}X'Y-\Gamma] \qquad (B.1.2)$$

Substituindo $y^*=Y\ell$ e $\Gamma^*=\Gamma\ell$ nas expressões (B.1.1), obtemos:

$$S_0^2 = \ell'Y'[I-X(X'X)^{-1}X']Y\ell$$

$$R_1^2 = \ell'\{Y'[I-X(X'X)^{-1}X']Y +$$

$$+ [C(X'X)^{-1}X'Y-\Gamma]'[C(X'X)^{-1}C']^{-1}[C(X'X)^{-1}X'Y-\Gamma]\}\ell$$

Então utilizando as relações (A.1.6) e (B.1.2) concluímos que: $S_0^2 = \ell'S\ell$ e $R_1^2 = \ell'R\ell$. Logo $R_1^2 - S_0^2 = \ell'(R-S)\ell = \ell'H\ell$, e a matriz H é definida como a matriz de somas de quadrados e produtos devida ao desvio em relação à hipótese multivariada H: $C\beta = \Gamma$.

A decomposição matricial $R = S+(R-S) = S+H$ é uma ge-

neralização da partição de somas de quadrados $R_1^2=S_0^2+(R_1^2-S_0^2)$ obtida através da teoria de Análise de Variância Univariada.

Vejamos agora alguns resultados sobre a distribuição dessas matrizes (S e H) de formas quadráticas e bilineares.

Primeiramente mostremos que a matriz S é positiva definida (p.d.) em quase toda parte (q.t.p.). Tomemos um vetor $\ell \in \mathbb{R}^p$, $\ell \neq 0$, arbitrário, porém fixo, e consideremos a seguinte forma quadrática:

$$\ell' S \ell = \ell'(Y-X\hat{\beta})'(Y-X\hat{\beta})\ell.$$

Fazendo a transformação $(Y-X\hat{\beta})\ell = z$, temos $\ell' S \ell = z'z \geq 0, \forall \ell \in \mathbb{R}^p$. Para mostrar que S é p.d. em q.t.p. basta mostrar que

$$\ell' S \ell = 0 \Longrightarrow \ell = 0 \text{ em q.t.p.}$$

$$\ell' S \ell = 0 \Longrightarrow z'z = 0 \Longrightarrow z = 0 \Longrightarrow (Y-X\hat{\beta})\ell = 0 \qquad \text{(B.1.3)}$$

Como Y é uma matriz cujas colunas são variáveis aleatórias multivariadas contínuas, e como a matriz $[I-X(X'X)^{-1}X']$ tem posto $n-r \geq p$ (por hipótese), então a matriz $[I-X(X'X)^{-1}X']Y$ tem posto p e suas colunas são linearmente independentes em q.t.p. Mas $(Y-X\hat{\beta})\ell = [I-X(X'X)^{-1}X']Y\ell$ é uma combinação linear das colunas matriz $[I-X(X'X)^{-1}X']Y$; logo, para que (B.1.3) seja válido devemos ter $\ell = 0$ em q.t.p. Então

$$\ell' S \ell > 0, \ \forall \ell \in \mathbb{R}^p, \ \ell \neq 0 \text{ (em q.t.p.)}$$

e S p.d. e tem posto p em q.t.p.

Mostremos agora que a matriz H é pelo menos positiva semi-definida (p.s.d.) e se $p \leq c$, H é p.d. em q.t.p.

A matriz $[C(X'X)^{-1}C']^{-1}$ é simétrica e não singular. Então ela é p.d. e existe uma matriz T não singular, tal que

$[C(X'X)^{-1}C']^{-1} = T'T$. Então podemos escrever:

$$H = [C(X'X)^{-1}X'Y-\Gamma]'T'T[C(X'X)^{-1}X'Y-\Gamma]$$

Tomemos um vetor $\ell \in \mathbb{R}^p$, $\ell \neq 0$, arbitrário, porém fixo, e consideremos a seguinte forma quadrática:

$$\ell'H\ell = \ell'[C(X'X)^{-1}X'Y-\Gamma]'T'T[C(X'X)^{-1}X'Y-\Gamma]\ell$$

Fazendo a transformação $T[C(X'X)^{-1}X'Y-\Gamma]\ell = w$, temos

$$\ell'H\ell = w'w \geq 0, \ \forall \ell \in \mathbb{R}^p$$

$$\ell'H\ell = 0 \Rightarrow w'w = 0 \Rightarrow w = 0 \Rightarrow T[C(X'X)^{-1}X'Y-\Gamma]\ell = 0 \Rightarrow [C(X'X)^{-1}X'Y-\Gamma]\ell = 0 \qquad (B.1.4)$$

Sabemos que a matriz $[C(X'X)^{-1}X'Y-\Gamma]$ $(c \times p)$ tem posto

$$s = \min(c,p).$$

Se $p \leq c$, então $s = p$ e as colunas de $[C(X'X)^{-1}X'Y-\Gamma]$ são linearmente independentes em q.t.p. Logo para que (B.1.4) seja válido, devemos ter $\ell = 0$ em q.t.p. Então

$$\ell'H\ell > 0, \ \forall \ell \in \mathbb{R}^p, \ \ell \neq 0 \ \text{(em q.t.p.)}$$

e H é p.d. em q.t.p.

Se $p > c$, então $s = c$ e as colunas de $[C(X'X)^{-1}X'Y-\Gamma]$ são linearmente dependentes. Logo existe $\ell \in \mathbb{R}^p$, $\ell \neq 0$, tal que $[C(X'X)^{-1}X'Y-\Gamma]\ell = 0$. Então H é p.s.d.

Se a hipótese H: $C\beta = \Gamma$ for verdadeira, para todo vetor $\ell \in \mathbb{R}^p$, $\ell \neq 0$, $S_0^2 = \ell'S\ell$ e $R_1^2 - S_0^2 = \ell'H\ell$ são independentes e têm as seguintes distribuições:

$$S_0^2 = \ell'S\ell \sim \sigma_\ell^2 \chi^2_{n-r}$$

$$R_1^2 - S_0^2 = \ell'H\ell \sim \sigma_\ell^2 \chi^2_c$$

(Searle, 1971, cap. 3, por exemplo).

Então, utilizando resultados conhecidos sobre a distribuição de Wishart (Rao, 1973, por exemplo), as matrizes S e H são independentes e têm as seguintes distribuições:

$$S \sim W_p(n-r, \Sigma, \cdot) \text{ central}$$
$$H \sim W_p(c, \Sigma \cdot) \quad \text{central}$$

Se a hipótese H: $C\beta = \Gamma$ não for verdadeira, para todo vetor $\ell \in \mathbb{R}^p$, $\ell \neq 0$, S_0^2 e $R_1^2 - S_0^2$ são independentes e têm as seguintes distribuições:

$$S_0^2 \sim \sigma_\ell^2 \chi^2_{n-r} \quad \text{central}$$
$$R_1^2 - S_0^2 \sim \sigma_\ell^2 \chi^2_c \text{ central ou não, dependendo de } \ell.$$

Utilizando resultados sobre a distribuição de Wishart, concluímos que as matrizes S e H são independentes e têm as seguintes distribuições:

$$S \sim W_p(n-r, \Sigma, \cdot) \text{ central}$$
$$H \sim W_p\{c, \Sigma, [C\beta-\Gamma]'[C(X'X)^{-1}C']^{-1}[C\beta-\Gamma]\} \text{ não central}$$

Como no caso univariado, podemos obter um teste de significância para a hipótese H: $C\beta = \Gamma$ através da comparação das matrizes S e H. Vamos obter esse teste através da aplicação do princípio da União-Intersecção de Roy, que descreveremos a seguir (Roy, 1957).

A hipótese multivariada H: $C\beta = \Gamma$ será verdadeira se e somente se a hipótese univariada H': $C\beta\ell = \Gamma\ell$ for verdadeira para todo vetor $\ell \in \mathbb{R}^p$, $\ell \neq 0$. Da teoria univariada sabemos que a estatística utilizada para testar H' é dada por:

$$F(\ell) = \frac{(R_1^2 - S_0^2)/c}{S_0^2/(n-r)} = \frac{\ell' H \ell / c}{\ell' S \ell / (n-r)}$$

que sob a hipótese nula tem distribuição F central com parâ

metros c e (n-r).

Para um teste univariado com nível de significância α', a hipótese $H': C\beta\ell = \Gamma\ell$ é aceita se $F(\ell) \leq F_{\alpha',c,n-r}$. A aceitação da hipótese multivariada $H: C\beta = \Gamma$ (considerando um nível de significância α) é equivalente à aceitação da hipótese univariada $H': C\beta\ell = \Gamma\ell$ para todo vetor $\ell \in \mathbb{R}^p$, $\ell \neq 0$. A região de aceitação de $H: C\beta = \Gamma$ corresponde à intersecção:

$$\bigcap_{\ell \neq 0} \left\{ F(\ell) \leq F_{\alpha' c,n-r} \right\} \qquad (B.1.5)$$

de todas as regiões de aceitação das hipóteses univariadas $H' : C\beta\ell = \Gamma\ell$. Então a região de aceitação definida em (B.1.5) é equivalente a:

$$\max_{\ell \neq 0} F(\ell) \leq F_{\alpha,c,n-r} \qquad (B.1.6)$$

pois se o máximo $F(\ell)$ pertencer à região de aceitação da hipótese nula, os outros valores também pertencerão. Devemos então determinar

$$\max_{\ell \neq 0} \left[F(\ell) = \frac{\ell' H \ell / c}{\ell' S \ell / (n-r)} \right],$$

que é proporcional à

$$\max_{\ell \neq 0} \left[F^*(\ell) = \frac{\ell' H \ell}{\ell' S \ell} \right].$$

Com essa finalidade demonstremos os dois teoremas seguintes:

<u>TEOREMA B.1.1</u> - Seja A (m×m) uma matriz simétrica e B (m×m) uma matriz simétrica p.d. Então existe uma matriz R não sin

gular, tal que $R^{-1}{}'\Lambda R^{-1} = A$ e $R^{-1}{}'R^{-1} = B$ onde Λ é a matriz diagonal formada pelas raízes da equação $|A-\lambda B| = 0$.

Como B é p.d., existe uma matriz não singular T, tal que $B = T'T$. Considerando a equação $|A-\lambda B| = 0$, podemos escrever:

$$|A-\lambda B| = 0 \Longleftrightarrow |A-\lambda T'T| = 0 \Longleftrightarrow |T^{-1}{}'||A-\lambda T'T||T^{-1}| = 0 \Longleftrightarrow |T^{-1}{}'AT^{-1}-\lambda I| = 0$$

As raízes da equação $|A-\lambda B| = 0$ são as mesmas raízes da equação $|T^{-1}{}'AT^{-1}-\lambda I| = 0$.

Seja $C = T^{-1}{}'AT^{-1}$, que é simétrica. Então existe uma matriz P não singular, tal que: $P'CP = \Lambda \Longleftrightarrow P'T^{-1}{}'AT^{-1}P = \Lambda$ e $P'P = I$, onde Λ é uma matriz diagonal formada pelas raízes características de C (que são as raízes de $|A-\lambda B| = 0$).

Tomando $R = T^{-1}P$ podemos escrever:

i) $R'AR = \Lambda \Longrightarrow A = R^{-1}{}'\Lambda R^{-1}$

ii) $R'T'TR = P'T^{-1}{}'T'TT^{-1}P = I \Longrightarrow R'BR = I \Longrightarrow B = R^{-1}{}'R^{-1}$

<u>TEOREMA B.1.2</u> - Se A (m×m) é uma matriz simétrica e B (m×m) é uma matriz simétrica p.d. então o máximo e o mínimo de $\frac{x'Ax}{x'Bx}$, onde x é um vetor (m×1), são respectivamente a maior e a menor raiz da equação $|A-\lambda B| = 0$.

Consideremos a transformação $x' = y'R'$ onde R' é uma matriz tal que $R'AR = \Lambda$ e $R'BR = I$ (utilizando o resultado anterior). Então:

$$\frac{x'Ax}{x'Bx} = \frac{y'R'ARy}{y'R'BRy} = \frac{y'\Lambda y}{y'y} = \frac{\sum_{i=1}^{m} \lambda_i y_i^2}{\sum_{i=1}^{m} y_i^2} = \frac{\lambda_1 y_1^2}{\sum_{i=1}^{m} y_i^2} + \ldots + \frac{\lambda_m y_m^2}{\sum_{i=1}^{m} y_i^2}$$

Vamos supor, sem perda de generalidade que $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_m$. Então devemos determinar:

$$\max_{\rho} \sum_{i=1}^{m} \lambda_i \rho_i \text{ onde } \rho_i = \frac{y_i^2}{\sum_{i=1}^{m} y_i^2} \text{ e } \sum_{i=1}^{m} \rho_i = 1,\ 0 \le \rho_i \le 1,\ \forall i.$$

O máximo será atingido quando $\rho_1 = 1$, ou seja quando

$$\frac{y_1^2}{\sum_{i=1}^{m} y_i^2} = 1$$

o que implica $y_1 = 1$ e $y_i = 0$, $\forall i \neq 1$. Logo $\max_x \frac{x'Ax}{x'Bx} = \lambda_1$. Analogamente mostramos que $\min_x \frac{x'Ax}{x'Bx} = \lambda_m$.

Aplicando o resultado do Teorema B.1.2, concluímos que

$$\max_{\ell \neq 0} \left[F(\ell) = \frac{\ell' H \ell / c}{\ell' S \ell / (n-r)} \right]$$

é proporcional à maior raiz da equação $|H-\lambda S| = 0$. Para obter um teste de significância para a hipótese H: $C\beta = \Gamma$ precisamos estudar a distribuição das raízes características da equação $|H-\lambda S| = 0$. Nesse sentido, consideremos os dois teoremas seguintes, dos quais demonstraremos o segundo.

<u>TEOREMA B.1.3</u> - Seja A $(p \times p)$ uma matriz simétrica. Então:

i) Se $\rho(A) = s$ a equação $|A-\lambda I| = 0$ tem uma raiz nula com multiplicidade $(p-s)$.

ii) Todas as raízes características e vetores característicos de A são reais.

iii) Se A é p.d. todas suas raízes características são positivas; se A é p.s.d. todas suas raízes características são não-negativas.

(A demonstração pode ser encontrada em Roy, 1957, por exemplo).

TEOREMA B.1.4 - Seja A (p×p) uma matriz simétrica pelo menos p.s.d. de posto s e B (p×p) uma matriz simétrica p.d. Então a equação $|A-\lambda B| = 0$ tem exatamente s raízes positivas e (p-s) raízes nulas.

Como já vimos na demonstração do Teorema B.1.1 as raízes de $|A-\lambda B| = 0$ são as mesmas de $|T^{-1'}AT^{-1}-\lambda I| = 0$. Mas $T^{-1'}AT^{-1}$ é uma matriz simétrica pelo menos p.s.d de posto s. Então pelo Teorema B.1.3, $T^{-1'}AT^{-1}$ tem s raízes características positivas e (p-s) raízes características nulas.

Chamemos a maior raiz característica da equação $|H-\lambda S| = 0$ de λ_s onde $s = \min(c,p)$. Aceitaremos a hipótese H: $C\beta = \Gamma$ com nível de significância α se $\lambda_s \leq \lambda(\alpha)$ onde $\lambda(\alpha)$ é o $100\alpha^{\circ}$ percentil da distribuição da maior raiz característica de $|H-\lambda S| = 0$, quando a hipótese nula é verdadeira. Vários autores (citados em Morrison, 1976) tabelaram os percentis superiores da distribuição da maior raiz característica da equação: $|H-\theta(H+S)| = 0$. Vamos então demonstrar que a hipótese H: $C\beta = \Gamma$ pode ser testada através da estatística $\theta_s = \frac{\lambda_s}{1+\lambda_s}$ onde θ_s representa a maior raiz da equação

$$|H-\theta(H+S)| = 0.$$

TEOREMA B.1.5 - Uma condição ncessária e suficiente para que λ_s seja a maior raiz de $|H-\lambda S| = 0$ é que $\theta_s = \frac{\lambda_s}{1+\lambda_s}$ seja a maior raiz de $|H-\theta(H+S)| = 0$.

Sejam $\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_s$ as s raízes positivas de $|H-\lambda S| = 0$. Tomemos a raiz λ_i; então:

$$|H-\lambda_i S| = 0 \Longrightarrow |H-\lambda_i S+\lambda_i H-\lambda_i H| = 0 \Longrightarrow |(1+\lambda_i)H-\lambda_i(H+S)| = 0 \Longrightarrow$$

$$\Longrightarrow |H - \frac{\lambda_i}{1+\lambda_i}(H+S)| = 0.$$

Logo $\theta_i = \frac{\lambda_i}{1+\lambda_i}$ é raiz de $|H-\theta(H+S)| = 0$.

Como H + S é simétrica e p.d., pelo Teorema B.1.4 a

equação $|H-\theta(H+S)| = 0$ tem exatamente s raízes positivas e (p-s) raízes nulas. Sejam $\theta_1 \leq \theta_2 \leq \ldots \leq \theta_s$ as raízes positivas. Tomemos a raiz θ_j e façamos $\lambda_j = \frac{\theta_j}{1-\theta_j} \Longrightarrow \theta_j = \frac{\lambda_j}{1+\lambda_j}$. Então:

$$|H-\theta_j(H+S)| = 0 \Longrightarrow |H - \frac{\lambda_j}{1+\lambda_j}(H+S)| = 0 \Longrightarrow |(1+\lambda_j)H - \lambda_j(H+S)| = 0$$

$$\Longrightarrow |H+\lambda_j H - \lambda_j H - \lambda_j S| = 0 \Longrightarrow |H-\lambda_j S| = 0.$$

Logo λ_j é raiz de $|H-\lambda S| = 0$.

Tomemos $\lambda_i < \lambda_j$. Então $\lambda_i + \lambda_i\lambda_j < \lambda_j + \lambda_i\lambda_j \Longrightarrow$ $\Longrightarrow \lambda_i(1+\lambda_j) < \lambda_j(1+\lambda_i) \Longrightarrow \frac{\lambda_i}{1+\lambda_i} < \frac{\lambda_j}{1+\lambda_j} \Longrightarrow \theta_i < \theta_j$. Logo λ_s e $\theta_s = \frac{\lambda_s}{1+\lambda_s}$ são respectivamente as maiores raízes de

$$|H-\lambda S| = 0 \text{ e } |H-\theta(H+S)| = 0.$$

Os parâmetros da distribuição da maior raiz da equação $|H-\theta(H+S)| \doteq 0$ sob a hipótese nula são:

$$s = \min(c,p)$$

$$m = \frac{|c-p|-1}{2} \qquad \text{(B.1.7)}$$

$$n^* = \frac{n-r-p-1}{2}$$

e a região de aceitação da hipótese H: $C\beta = \Gamma$ (com nível de significância α) será dada por: $\theta_s \leq \theta_{s,m,n^*}(\alpha)$ onde $\theta_{s,m,n^*}(\alpha)$ é o $100\alpha^{\circ}$ percentil da distribuição da maior raiz da equação

$$|H-\theta(H+S)| = 0,$$

quando a hipótese é verdadeira.

$\theta_{s,m,n^*}(\alpha)$ pode ser obtido de tabelas convenientes; Morrison, (1976) apresenta tabelas para alguns valores dos parâmetros s, m, n* e cita os principais trabalhos nessa área.

As raizes da equação $|H-\theta(H+S)| = 0$ são iguais às raízes características de $H(H+S)^{-1}$ e na prática geralmente

é mais conveniente obter as raízes características de

$$H(H+S)^{-1}$$

através de um programa convencional. A distribuição conjunta das raízes características de matrizes da forma $H(H+S)^{-1}$ foi determinada simultaneamente por vários autores citados em Morrison (1976), que também apresenta indicações de outros trabalhos importantes envolvendo as distribuições necessárias para obtenção do teste proposto.

É comum encontrarmos esquemas de planejamento onde $s = 1$, e então a única raiz positiva de $H(H+S)^{-1}$ tem distribuição Beta com parâmetros $(m+1, n^*+1)$. Se considerarmos a transformação $F = \frac{n^*+1}{m+1} \frac{\theta s}{1-\theta s}$, a nova variável terá distribuição F com parâmetros $(2m+2)$, $(2n^*+2)$.

B.2 - O PRINCÍPIO DA RAZÃO DE VEROSSIMILHANÇA GENERALIZADA DE WILKS

Consideremos o $MLM(Y, X\beta, \Sigma)$. Nosso objetivo nesta seção é obter um teste de significância para a hipótese $H: C\beta = \Gamma$ através da aplicação do princípio da Razão de Verossimilhança Generalizada.

Com base nos mesmos argumentos da introdução da seção A.2 utilizaremos a forma equivalente $MLM'(U, \theta Z, \Sigma)$. Obter um teste de significância para a hipótese $H: C\beta = \Gamma$ utilizando o $MLM(Y, X\beta, \Sigma)$ é equivalente a obter um teste para a hipótese $H: \theta C^* = \Gamma^*$ utilizando o $MLM'(U, \theta Z, \Sigma)$ onde $C^* = C'$ e $\Gamma^* = \Gamma'$.

Primeiramente vamos reduzir o $MLM'(U, \theta Z, \Sigma)$ a sua forma canônica, segundo sugestão de Roy, J. (1966).

DEFINIÇÃO B.2.1 - Uma matriz G $(g \times h)$ é dita semi-ortogonal se $g > h$ e $G'G = I$

Como $\rho(Z) = r$, podemos obter uma matriz T_1 $(n \times n - r)$ semi-ortogonal, tal que $ZT_1 = 0$. Além disso, podemos escre-

ver cada coluna da matriz C^* em função de dois componentes: um pertencendo à variedade linear gerada pelas colunas de ZZ' e outra pertencendo a uma variedade linear ortogonal a esta. Então podemos escrever:

$$C^* = ZZ'D + R, \text{ onde } RD'ZZ' = 0 \qquad (B.2.1)$$

Como $RD'ZZ' = 0 \Longrightarrow RD'Z = 0$, a matriz D $(r \times c)$, deve ser tal que:

$$(C^*-ZZ'D)D'Z = 0 \Longrightarrow C^*D'Z = ZZ'DD'Z \Longrightarrow (ZZ')^{-1}C^*D'Z = DD'Z \qquad (B.2.2)$$

Como $\rho(Z'DD'Z) = c$, podemos obter uma matriz E $(n \times c)$ tal que $T_2 = Z'DD'ZE$ seja semi-ortogonal e $T_2'T_1 = 0$. Finalmente podemos construir T_3 $(n \times r-c)$ tal que $T = [T_1 T_2 T_3]$ $(n \times n)$ seja uma matriz ortogonal.

Consideremos a seguinte transformação ortogonal:

$$V = UT \Longleftrightarrow [V_1 V_2 V_3] = U[T_1 T_2 T_3] \qquad (B.2.3)$$

Então temos:

$$E(V_1) = E(U)T_1 = \theta Z T_1 = 0$$
$$E(V_2) = E(U)T_2 = \theta ZZ'DD'ZE = \theta C^*D'ZE = M_2^*$$
$$E(V_3) = E(U)T_3 = \theta Z T_3 = M_3$$

Fazendo $\varepsilon^{**} = \varepsilon^*[T_1\ T_2\ T_3]$, o modelo transformado pode ser escrito na forma $V = [V_1\ V_2\ V_3] = [0\ M_2^*\ M_3] + \varepsilon^{**}$ ou $V = M^* + \varepsilon^{**}$ onde $M^* = [0\ M_2^*\ M_3]$. Como os momentos de segunda ordem são invariantes em relação a transformações ortogonais, o modelo descrito acima corresponde ao $MLM'(V, M^*, \Sigma)$.

Se a hipótese H: $\theta C^* = \Gamma^*$ for verdadeira, teremos:

$$M_2^* = \theta C^*D'ZE = \Gamma^*D'ZE = M_{20}^*.$$

Tomando então:

$$W = [W_1\ W_2\ W_3] = [V_1\ V_2\ V_3] - [0\ M_{20}^*\ 0] = [0\ M_2^* - M_{20}^*\ M_3] + \varepsilon^{**} \Longrightarrow$$

$$\Longrightarrow W = [W_1\ W_2\ W_3] = [0\ M_2\ M_3] + \varepsilon^{**} = M + \varepsilon^{**}$$

o problema da obtenção de um teste de significância para a hipótese H: $\theta C^* = \Gamma^*$ utilizando o MLM'$(U, \theta Z, \Sigma)$ é equivalente ao problema de obtenção de um teste para a hipótese H': $M_2 = 0$ utilizando o MLM'(W, M, Σ).

Fazendo $\Sigma^{-1} = \{\sigma^{ij}\}$ a função de verossimilhança pode ser escrita como:

$$L = (2\pi)^{-\frac{1}{2}np} |\Sigma^{-1}|^{\frac{1}{2}n} \exp\left\{-\frac{1}{2} \sum_{\alpha=1}^{n} (w_\alpha - M_\alpha)' \Sigma^{-1} (w_\alpha - M_\alpha)\right\} =$$

$$= (2\pi)^{-\frac{1}{2}np} |\Sigma^{-1}|^{\frac{1}{2}n} \exp\left\{-\frac{1}{2} \sum_{\alpha=1}^{n} \left[\sum_{i=1}^{p} \sum_{j=1}^{p} \sigma^{ij} (w_{i\alpha} - M_{i\alpha})(w_{j\alpha} - M_{j\alpha})\right]\right\}$$

$$= (2\pi)^{-\frac{1}{2}np} |\Sigma^{-1}|^{\frac{1}{2}n} \exp\left\{-\frac{1}{2} \sum_{i=1}^{p} \sum_{j=1}^{p} \sigma^{ij} \left[\sum_{\alpha=1}^{n-r} w_{i\alpha} w_{j\alpha} + \sum_{\alpha=n-r+1}^{n-r+c} (w_{i\alpha} - M_{i\alpha})(w_{j\alpha} - M_{j\alpha}) + \right.\right.$$

$$\left.\left. + \sum_{\alpha=n-r-+c+1}^{n} (w_{i\alpha} - M_{i\alpha})(w_{j\alpha} - M_{j\alpha})\right]\right\} \qquad (B.2.4)$$

A estatística de teste segundo o princípio da razão de verossimilhança generalizada é definida como:

$$\Lambda^* = \frac{\sup\limits_{(M_2 = 0, \Sigma)} L(M, \Sigma)}{\sup\limits_{(M, \Sigma)} L(M, \Sigma)} \qquad (B.2.5)$$

Utilizando os resultados obtidos na seção A.2, podemos determinar os estimadores de máxima verossimilhança de M e Σ nos seguintes casos:

i) Sem restrição:

$$\hat{M} = [0 \; W_2 \; W_3]$$

$$\hat{\Sigma} = \frac{1}{n} \sum_{i=1}^{p} \sum_{j=1}^{p} \left[\sum_{\alpha=1}^{n-r} w_{i\alpha} w_{j\alpha} \right] = \frac{1}{n} W_1 W_1' \tag{B.2.6}$$

ii) Sob a restrição $M_2 = 0$

$$\hat{M}_H = [0 \; 0 \; W_3]$$

$$\hat{\Sigma}_H = \frac{1}{n} \sum_{i=1}^{p} \sum_{j=1}^{p} \left[\sum_{\alpha=1}^{n-r} w_{i\alpha} w_{j\alpha} + \sum_{\alpha=n-r+1}^{n-r+c} w_{i\alpha} w_{j\alpha} \right] = \frac{1}{n} [W_1 W_1' + W_2 W_2'] \tag{B.2.7}$$

Então:

$$\sup_{(M,\Sigma)} L(M,\Sigma) = (2\pi)^{-\frac{1}{2}np} |\hat{\Sigma}^{-1}|^{\frac{1}{2}n} \exp\left\{ -\frac{1}{2} \sum_{\alpha=1}^{n} (w_\alpha - M_\alpha)' \hat{\Sigma}^{-1} (w_\alpha - M_\alpha) \right\} =$$

$$= (2\pi)^{-\frac{1}{2}np} |\hat{\Sigma}^{-1}|^{\frac{1}{2}n} \exp\left\{ -\frac{1}{2} \sum_{\alpha=1}^{n} tr[(w_\alpha - M_\alpha)' \hat{\Sigma}^{-1} (w_\alpha - M_\alpha)] \right\} =$$

$$= (2\pi)^{-\frac{1}{2}np} |\hat{\Sigma}^{-1}|^{\frac{1}{2}n} \exp\left\{ -\frac{1}{2} tr \hat{\Sigma}^{-1} \left[\sum_{\alpha=1}^{n} (w_\alpha - M_\alpha)' (w_\alpha - M_\alpha) \right] \right\} =$$

$$= (2\pi)^{-\frac{1}{2}np} |\hat{\Sigma}^{-1}|^{\frac{1}{2}n} \exp\left\{ -\frac{1}{2} n \; tr \; \hat{\Sigma}^{-1} \hat{\Sigma} \right\} =$$

$$= (2\pi)^{-\frac{1}{2}np} |\hat{\Sigma}^{-1}|^{\frac{1}{2}n} e^{-\frac{1}{2}n} \tag{B.2.8}$$

De maneira análoga, podemos mostrar que:

$$\sup_{(M_{20}=0,\Sigma)} L(M,\Sigma) = (2\pi)^{-\frac{1}{2}np} |\hat{\Sigma}_H^{-1}|^{\frac{1}{2}n} e^{-\frac{1}{2}n} \tag{B.2.9}$$

Utilizando os resultados (B.2.5), (B.2.8) e (B.2.9) concluímos que:

$$\Lambda^* = \frac{|\hat{\Sigma}|^{\frac{1}{2}n}}{|\hat{\Sigma}_H|^{\frac{1}{2}n}} \qquad (B.2.10)$$

Nós rejeitaremos a hipótese $H: \theta C^* = \Gamma^*$ (com nível de significância α) se o valor observado (λ^*) da estatística Λ^* que $\lambda^* < \lambda^*(\alpha)$ onde $\lambda^*(\alpha)$ é obtido a partir de

$$P[\Lambda^* < \lambda^*(\alpha) | H \text{ é verdadeira}] = \alpha.$$

Podemos ainda considerar a transformação:

$$\Lambda = \Lambda^{*\frac{2}{n}} = \frac{|\hat{\Sigma}|}{|\hat{\Sigma}_H|} \qquad (B.2.11)$$

Nesse caso rejeitaremos a hipótese $H: \theta C^* = \Gamma^*$ (com nível de significância α) se o valor observado (λ) da estatística Λ for tal que $\lambda < \lambda(\alpha)$ onde $\lambda(\alpha)$ é obtido a partir de

$$P[\Lambda < \lambda(\alpha) | H \text{ é verdadeira}] = \alpha.$$

Vamos agora obter a expressão de Λ em função dos elementos do $MLM(Y, X\beta, \Sigma)$.

Como $ZT_1 = 0$, temos

$$ZT_1T_1'Z = 0 = ZZ' - ZZ'(ZZ')^{-1}ZZ' = Z[I - Z'(ZZ')^{-1}Z]Z' \Longrightarrow T_1T_1' = I - Z'(ZZ')^{-1}Z$$

Então:

$$W_1W_1' = V_1V_1' = UT_1T_1'U' = U[I - Z'(ZZ')^{-1}Z]U' = Y'[I - X(X'X)^{-1}X']Y \qquad (B.2.12)$$

De (B.2.6) e (B.2.12) concluímos que:

$$\hat{\Sigma} = \frac{1}{n} Y'[I - X(X'X)^{-1}X']Y = \frac{1}{n} S \qquad (B.2.13)$$

Por outro lado:

$$W_2W_2' = [V_2-\Gamma^*D'ZE][V_2-\Gamma^*D'ZE]' = [UT_2-\Gamma^*D'ZE][UT_2-\Gamma^*D'ZE]' =$$

$$= UT_2T_2'U'-UT_2E'Z'D\Gamma^{*\prime}-\Gamma^*D'ZET_2'U'+\Gamma^*D'ZEE'Z'D\Gamma^{*\prime} =$$

$$= UZ'DD'ZEE'Z'DD'ZU' - UZ'DD'ZEE'Z'D\Gamma^{*\prime} - \Gamma^*D'ZEE'Z'DD'ZU' +$$

$$+ \Gamma^*D'ZEE'Z'D\Gamma^{*\prime}$$

Utilizando o resultado (B.2.2), obtemos:

$$W_2W_2' = UZ'(ZZ')^{-1}C^*D'ZEE'Z'DC^{*\prime}(ZZ')^{-1}ZU'-UZ'(ZZ')^{-1}C^*D'ZEE'Z'D\Gamma^{*\prime} -$$

$$- \Gamma^*D'ZEE'ZDC^{*\prime}(ZZ')^{-1}ZU' + \Gamma^*D'ZEE'Z'D\Gamma^{*\prime} =$$

$$= [UZ'(ZZ')^{-1}C^*-\Gamma^*][D'ZEE'Z'D][C^{*\prime}(ZZ')^{-1}ZU'-\Gamma^{*\prime}] \qquad (B.2.14)$$

Como T_2 é semi-ortogonal, temos:

$$T_2'T_2 = I \Longrightarrow E'Z'DD'ZZ'DD'ZE = I$$

e utilizando o resultado (B.2.2) vem:

$$E'Z'DC^{*\prime}(ZZ')^{-1}ZZ'(ZZ')^{-1}C^*D'ZE = I \Longrightarrow E'Z'DC^{*\prime}(ZZ')^{-1}C^*D'ZE = I \Longrightarrow$$

$$\Longrightarrow D'ZEE'Z'DC^{*\prime}(ZZ')^{-1}C^*D'ZE(D'ZE)^{-1} = I \Longrightarrow D'ZEE'Z'D[C^{*\prime}(ZZ')^{-1}C^*] =$$

$$= I \Longrightarrow D'ZEE'Z'D = [C^{*\prime}(ZZ')^{-1}C^*]^{-1} \qquad (B.2.15)$$

De (B.2.14) e (B.2.15) concluímos que:

$$W_2W_2' = [UZ'(ZZ')^{-1}C^*-\Gamma^*][C^{*\prime}(ZZ')^{-1}C^*]^{-1}[UZ'(ZZ')^{-1}C^*-\Gamma^*]' \Longrightarrow W_2W_2' =$$

$$= [C(X'X)^{-1}X'Y-\Gamma]'[C(X'X)^{-1}C']^{-1}[C(X'X)^{-1}X'Y-\Gamma] \qquad (B.2.16)$$

De (B.2.7), (B.2.12) e (B.2.16) concluímos que:

$$\hat{\Sigma}_H = \frac{1}{n}\left\{Y'[I-X(X'X)^{-1}X']Y+[C(X'X)^{-1}X'Y-\Gamma]'[C(X'X)^{-1}C']^{-1}[C(X'X)^{-1}X'Y-\Gamma]\right\} =$$

$$= \frac{1}{n}(S+H) \qquad (B.2.17)$$

Logo, utilizando os resultados (B.2.11), (B.2.13) e (B.2.17), temos:

$$\Lambda = \frac{|S|}{|S+H|} \qquad (B.2.18)$$

Pelo Teorema B.1.1 existe uma matriz R não singular tal que $R'HR = \Lambda$ e $R'SR = I$ onde Λ é uma matriz diagonal formada pelas raízes da equação $|H-\lambda S| = 0$. Então:

$$\Lambda = \frac{|S|}{|S+H|} = \frac{|R'||S||R|}{|R'||S+H||R|} = \frac{|R'SR|}{|R'SR+R'HR|} = \frac{|I|}{|I+\Lambda|} = \prod_{i=1}^{s} \left(\frac{1}{1+\lambda_1}\right) \qquad (B.2.19)$$

Pelo Teorema B.1.5, as raízes características positivas de $H(S+H)^{-1}$ são dadas por $\theta_i = \frac{1}{1+\lambda_i}$ onde λ_i $(i=1,2,\ldots,s)$ são as raízes da equação $|H-\lambda S| = 0$. Então de (B.2.19) segue que:

$$\Lambda = \prod_{i=1}^{s} (1-\theta_i) \qquad (B.2.20)$$

Wilks (1932) obteve os momentos da distribuição de Λ sob a hipótese nula, além de expressões da distribuição de Λ para alguns valores dos parâmetros p e c. Esses resultados, com algumas correções são apresentados por Andersen (1958). A estatística $-[n-r-\frac{1}{2}(p-c+1)]\log \Lambda$ tem distribuição assintótica de χ^2 com cp graus de liberdade. Outras aproximações e transformações da distribuição de Λ são discutidas em Anderson (1958), que cita as principais contribuições ao estudo dessa distribuição.

Para alguns valores de p e c é possível transformar a estatística Λ numa outra estatística com distribuição F. Por exemplo, para $s = \min(c,p) = 1$, a estatística

$$F = \frac{1-\Lambda}{\Lambda} \frac{n^*+1}{m+1} \text{ onde } m = \frac{|c-p|-1}{2}, \; n^* = \frac{n-r-p-1}{2}$$

tem distribuição F com $(2m+2)$, $(2n^*+2)$ graus de liberdade sob a hipótese nula.

3.3 - OUTROS CRITÉRIOS DE TESTE; CONSIDERAÇÕES GERAIS

Nas seções B.1 e B.2 discutimos dois tipos de testes de significância para a hipótese H: $C\beta = \Gamma$ quando consideramos o MLM$(Y, X\beta, \Sigma)$. Como vimos, esses testes dependem das raízes da equação $|H-\lambda S| = 0$. Explorando essa característica comum aos dois testes, Roy (1953, 1957) demonstrou os seguintes resultados:

i) As raizes da equação $|H-\lambda S| = 0$ são os únicos elementos invariantes sob certas transformações lineares, que aplicadas à matriz de observações ou à matriz de planejamento, mantêm invariante a hipótese original.

ii) A distribuição conjunta dessas raízes, quando a hipótese nula é verdadeira, envolve somente parâmetros conhecidos (graus de liberdade).

Com base nesses resultados, parece razoável admitir que outras possíveis estatísticas de teste sejam funções dessas raízes. Pillai (1955) afirma que todas as estatísticas de teste apresentadas possuem essa característica; como exemplos, cita: $\theta_s = \max(\theta_i)$, devida a Roy,

$$\Lambda = \prod_{i=1}^{s} (1-\theta_i), \text{ devida a Wilks e}$$

$$V = \sum_{i=1}^{s} \frac{\theta_i}{1-\theta_i} = tr(HS^{-1}) \text{ devida a Lawley e Hotelling.}$$

Além disso, Pillai apresenta mais três sugestões:

$$H = s\left[\sum_{i=1}^{s} (1-\theta_i)\right]^{-1}; \quad R = s\left[\sum_{i=1}^{s} \theta_i^{-1}\right] \text{ e } T = s\left[\sum_{i=1}^{s} \lambda_i^{-1}\right]^{-1}.$$

A escolha de uma determinada função como estatísti

ca de teste foi sempre baseada em considerações heurísticas, cujas principais justificativas são fundamentadas nas seguintes propriedades:

I) Possibilidade de obtenção de uma forma explícita e tabulável para a distribuição (exata ou assintótica) da estatística de teste sob a hipótese nula.

II) O teste deve apresentar boas características operacionais, como:

IIA) ser não viciado;

IIB) ser tal que o poder seja uma função não-decrescente de cada um dos parâmetros indicadores de desvio em relação à hipótese nula;

IIC) existência de algum limite inferior razoável para o poder do teste;

IID) possibilidade da obtenção de intervalos de confiança;

IIE) admissibilidade.

Embora muitos pesquisadores tenham trabalhado no estudo das distribuições das estatísticas de teste propostas, foram obtidos poucos resultados que favorecem o cálculo das probabilidades exatas, mesmo quando a hipótese nula é verdadeira. Uma exceção é o critério proposto por Wilks, para o qual existem alguns resultados úteis, que são discutidos em Anderson (1958), por exemplo. Também não há resultados abrangentes no estudo das propriedades dos testes apresentados. Anderson (1958), resumiu claramente o problema:

"... Unfortunately, there is no theory to tell us what is a good class of tests (a complete class), nor how to choose among the proposed tests, when we want power against certain kinds of alternatives."

Pillai (1955) cita vários autores cujos trabalhos envolvem alguns resultados específicos. Por exemplo: a es-

tatística V de Lawley Hotelling obedece satisfatoriamente às propriedades ii) e I); a estatística Λ de Wilks, às propriedades ii), I), IIA) e IIB); a estatística θ_s de Roy, às propriedades ii), I), IIA) e IID). No entanto, parece que o trabalho que apresenta resultados práticos mais úteis é o de Schatzoff (1966), em que são feitas comparações entre diversas estatísticas de testes através da introdução do conceito de "nível de significância esperado". As comparações foram obtidas através de simulação, e a principal conclusão é que não há um domínio de uma estatística sobre as outras, o que não soluciona o problema de seleção de um "melhor" teste. Porém são apresentadas condições estruturais das populações envolvidas, que favorecem um ou outro critério.

Segundo Schatzoff, para grandes amostras, a utilização das estatísticas de Wilks ou Lawley-Hotelling é indiferente em virtude de sua equivalência assintótica. No caso de pequenas amostras o critério de Wilks parece aconselhável em função da existência de aproximações para o poder relativo a algumas alternativas.

Finalizando este apêndice, apresentamos alguns comentários sobre testes envolvendo hipóteses da forma

$$H: C\beta U = \Gamma.$$

Utilizando a sugestão apresentada na introdução do apêndice, podemos obter um teste para uma hipótese dessa forma, considerando a transformação $Y^* = YU$. Nesse caso as matrizes S e H podem ser escritas como:

$$S = Y^{*\prime}[I-X(X'X)^{-1}X']Y^* = U'Y'[I-X(X'X)^{-1}X']YU \qquad (B.3.1)$$

$$H = [C(X'X)^{-1}X'YU-\Gamma]'[C(X'X)^{-1}C']^{-1}[C(X'X)^{-1}X'YU-\Gamma] =$$

$$= [C\hat{\beta}U-\Gamma]'[C(X'X)^{-1}C']^{-1}[C\hat{\beta}U-\Gamma] \qquad (B.3.2)$$

Além disso, conforme os resultados obtidos na seção B.1, po

demos concluir que se a hipótese nula é verdadeira, S e H são independentes e têm distribuição

$$S \sim W_u[n-r, U'\Sigma U, \cdot] \text{ central}$$

$$H \sim W_u[c, U'\Sigma U, \cdot] \text{ central.}$$

Se a hipótese nula não for verdadeira, S e H são independentes e têm distribuição:

$$S \sim W_u[n-r, U'\Sigma U, \cdot] \text{ central}$$

$$H \sim W_u\left\{c, U'\Sigma U, [C\beta U-\Gamma]'[C(X'X)^{-1}C']^{-1}[C\beta U-\Gamma]\right\} \text{ não central.}$$

- □ -

BIBLIOGRAFIA

[1] - Anderson,T.W.(1958), Introduction to Multivariate Statistical Analysis, John Wiley and Sons, New York.

[2] - Bock,R.D.(1963), "Multivariate analysis of variance of repeated measurements". Problems in Measuring Change, University of Wisconsin Press, Wisconsin, pp. 85-103.

[3] - Box,G.E.P.(1950), "Problems in the analysis of growth and wear curves". Biometrics, 6, pp.362-389.

[4] - Elston,R.C. & Grizzle,J.E. (1962), "Estimation of time-response curves and their confidence bands". Biometrics,18,pp.148-159.

[5] - Gleser,L. & Olkin,I. (1966), "A k-sample regression model with covariance". Multivariate Analysis, Academic Press,New York, pp. 59-72.

[6] - Grizzle,J.E. & Allen,D.M.(1969),"Analysis of growth and dose response curves." Biometrics, 25, pp. 357-381.

[7] - Healy,M.J.R.(1961), "Experiments for comparing growth curves". Abstract nº 752, Biometrics, 17, pp. 333.

[8] - Khatri,C.G. (1966), "A note on a MANOVA medel applied to problems in growth curves". Ann.Inst.Statis.Math.,18, pp.75-86.

[9] - Leech,F. & Healy,M.J.R. (1959), "The analysis of experiments on growth rate". Biometrics, 15, pp. 98-106.

[10] - Mood,A.M.,Graybill,F.A. & Boes,D.C. (1974), Introduction to the Theory of Statistics, McGraw-Hill Kogakusha Ltd.

[11] - Morrison,D.F. (1976), Multivariate Statistical Methods, McGraw-Hill, Inc.

[12] - Pearson,E.S. & Hartley,H.O. (1958), Biometrika, Tables for Statisticians, Volume 1, Cambridge University Press.

[13] - Pillai,K.C.S. (1955), "Some new test criteria in multivariate analysis". Ann.Math.Statist.,26, pp. 117-121.

[14] - Potthoff,R.F. & Roy,S.N., (1964), "A generalized multivariate analysis" of variance model useful especially for growth curves problems". Biometrika, 51, pp. 313-326.

[15] - Rao,C.R. (1958), "Some statistical methods for comparison of growth curves". Biometrics, 14, pp. 1-17.

[16] - Rao,C.R. (1959), "Some problems involving linear hypothesis in multivariate analysis". Biometrika, 46, pp. 49-58.

[17] - 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, pp. 447-458.

[18] - Rao,C.R. (1966), "Covariance adjustment and related problems in multivariate analysis". Multivariate Analysis,Academic Press, New York, pp. 87-103.

[19] - Rao,C.R. (1967), "Least squares theory using an estimated dispersion matrix and its application to measurement of signals". Proc.Fifth Berkeley Symp.Math.Statist.Prob.,Vol.I,pp.355-372.

[20] - Rao,C.R.,(1973),Linear Statistical Inference and its Applications, 2nd edition, John Wiley and Sons, New York.

[21] - Roy,J. (1966), "Power of the Likehood Ratio Test Used in Analysis of Dispersion". Multivariate Analysis, Academic Press, New York, pp. 105-127.

[22] - Roy,S.N. (1953), "On a heuristic method of test construction and its use in multivariate analysis". Ann.Math.Statist., 24,pp. 220-238.

[23] - Roy,S.N. (1957), Some Aspects of Multivariate Analysis,John Wiley and Sons, New York.

[24] - Schatzoff,M. (1966), "Sensitivity comparisons among tests of the general linear hipothesis". American Statistical Association Journal, June 1966, pp. 415-435.

[25] - Searle, S.R., (1971), Linear Models, John Wiley and Sons, New York.

[26] - Wilks,S.S., (1932), "Certain generalizations in the analysis of variance". Biometrika, 24, pp. 471-494.

[27] - Wishart,J., (1938), "Growth rate determinations in nutrition studies with the bacon pig, and their analysis". Biometrika, 30, pp. 16-28.

- □ -

REFERÊNCIAS

[1] - Allen,D.M.(1967) "Multivariate Analysis of Nonlinear Models",North Carolina Institute of Statistics Mimeo Series Nº 558.

[2] - Bouden,D.C. & Skinhorst,R.K.,(1973),"Tolerance Bands for Growth Curves". Biometrics, 28, pp.361-371.

[3] - Hills,M.,(1968),"A note on the analysis of growth curves" Biometrics, 24, pp. 189-196.

[4] - Morrison,D.F.,(1970),"The optimal spacing of repeated measurements", Biometrics, 26, pp. 281-290.

[5] - Morrison,D.F.,(1972),"The analysis of a single sample of repeated measurements", Biometrics,28, pp. 55-71.

[6] - Rodriguez,G.,(1975),"A multivariate linear model with latent factor structure", North Carolina Institute of Statistics Mimeo Series Nº 1014.

[7] - Snee,R.D.,(1972),"On the analysis of response curve data", Tech. 14, pp. 47-62.

[8] - Schwertman,N.C.,(1974),"The analysis of growth curve date with missing observations", California State University, tecnical report.

- □ -